

FALSEHOODS AND THE FIRST AMENDMENT

*Cass R. Sunstein**

TABLE OF CONTENTS

I. LIES AND MORE LIES	388
II. PROTECTING FALSEHOODS	396
<i>A. Official Fallibility</i>	398
<i>B. Chilling Truth and Public Goods</i>	400
<i>C. Learning from Falsity</i>	402
<i>D. Learning What Others Think</i>	403
<i>E. Counterspeech > Bans</i>	404
<i>F. Harm in the Balance</i>	405
III. FALSE STATEMENTS AND REPUTATIONAL HARM.....	406
<i>A. Old Law</i>	407
<i>B. Foundations</i>	410
<i>C. Well Beyond Libel</i>	415
IV. DEEPPAKES AND DOCTORED VIDEOS.....	418
<i>A. The Horizon</i>	419
<i>B. Facebook's Incomplete Solution</i>	423
V. CONCLUSION	425

* Robert Walmsley University Professor, Harvard Law School. I am grateful to Eric Posner, Geoffrey Stone, and David Strauss for valuable comments on a previous draft and to Ethan Lowens for superb research assistance.

I. LIES AND MORE LIES

SEEN FROM THE VIEWPOINT OF POLITICS, TRUTH HAS A DESPOTIC CHARACTER. IT IS THEREFORE HATED BY TYRANTS, WHO RIGHTLY FEAR THE COMPETITION OF A COERCIVE FORCE THEY CANNOT MONOPOLIZE, AND IT ENJOYS A RATHER PRECARIOUS STATUS IN THE EYES OF GOVERNMENTS THAT REST ON CONSENT AND ABHOR COERCION. FACTS ARE BEYOND AGREEMENT AND CONSENT, AND ALL TALK ABOUT THEM — ALL EXCHANGES OF OPINION BASED ON CORRECT INFORMATION — WILL CONTRIBUTE NOTHING TO THEIR ESTABLISHMENT. UNWELCOME OPINION CAN BE ARGUED WITH, REJECTED, OR COMPROMISED UPON, BUT UNWELCOME FACTS POSSESS AN INFURIATING STUBBORNNESS THAT NOTHING CAN MOVE EXCEPT PLAIN LIES.

— HANNAH ARENDT¹

What is the constitutional status of falsehoods? From the standpoint of the First Amendment, does truth or falsity matter? These questions have become especially pressing with the increasing power of social media, the frequent contestation of established facts, and the current focus on “fake news,” disseminated by both foreign and domestic agents in an effort to drive politics in the United States and elsewhere in particular directions. In 2012, the Supreme Court ruled for the first time that intentional falsehoods are protected by the First Amendment, at least when they do not cause serious harm.²

But in important ways, 2012 seems like a generation ago, and the Court has yet to give an adequate explanation for its conclusion. Such an explanation must begin with the risk of a “chilling effect,” by which an effort to punish or deter falsehoods might also in the process chill truth. But that is hardly the only reason to protect falsehoods, intentional or otherwise; there are several others.³ Even so, the various arguments suffer from abstraction and high-mindedness; they do not amount to decisive reasons to protect falsehoods. These propositions bear on old questions involving defamation and on new questions involving fake news, deepfakes, and doctored videos.

For orientation: some things are true. Dropped objects fall. The Earth goes around the sun. The Holocaust happened. Barack Obama was born in the United States. Elvis Presley is dead. Cigarette smoking causes cancer. Some falsehoods are harmful. They ruin lives. They lead people to take unnecessary risks or fail to protect themselves against serious dangers. They might undermine or even destroy the process of self-government.

1. HANNAH ARENDT, *Truth and Politics*, in BETWEEN PAST AND FUTURE 227, 241 (2006).

2. *United States v. Alvarez*, 567 U.S. 709, 723 (2012) (plurality opinion).

3. See *infra* Part II.

Suppose that in response to the spread of falsehoods, including “fake news,” Congress or a state enacts a new law: the False Information Act. The law makes it a civil wrong to circulate or publish false information. It applies not only to newspapers and television stations but also to social media platforms, such as Facebook and Twitter. The penalty is \$1, alongside an order to cease and desist. Would the False Information Act violate the First Amendment?

The answer is clearly yes,⁴ but it is not clear why.⁵ One of my main goals here is to offer an account of why democracies protect both innocent and deliberate falsehoods. With that account, I will offer something like two and a half cheers for current constitutional understandings. In the process, however, I shall suggest that narrower, more targeted versions of the False Information Act, suitable to the modern era, should be upheld. Among other things, I shall argue that public officials should be able to allow redress of libelous statements that do not meet the standards of *New York Times Co. v. Sullivan*,⁶ which looks increasingly anachronistic, a bit of a dinosaur in light of what is happening online and improved understandings about how information spreads.⁷ I shall also argue that officials should have (considerable) authority to regulate false statements of fact, deepfakes, and doctored videos. Even more clearly, Facebook, Twitter, YouTube, and other social media platforms should be doing significantly more than they are doing now to control the spread of falsehoods with the goal of protecting democratic processes, the reputations of individuals and institutions, and most broadly, the social norm in favor of respect for, and recognition of, what is true — a matter of uncontestable fact. As Hannah Arendt warned:

The chances of factual truth surviving the onslaught of power are very slim indeed; it is always in danger of being maneuvered out of the world not only for a time but, potentially, forever. Facts and events are infinitely more fragile things than axioms, discoveries, theories — even the most wildly speculative ones — produced by the human mind; they occur in the field

4. See *Alvarez*, 567 U.S. at 723 (plurality opinion) (striking down the Stolen Valor Act, Pub. L. No. 109-437, 120 Stat. 3266 (2006), which criminalized false claims about having military medals).

5. Professor Frederick Schauer rightly notes that “we have, perhaps surprisingly, arrived at a point in history in which an extremely important social issue about the proliferation of demonstrable factual falsity in public debate is one as to which the venerable and inspiring history of freedom of expression has virtually nothing to say.” Frederick Schauer, *Facts and the First Amendment*, 57 UCLA L. REV. 897, 908 (2010).

6. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254 (1964).

7. See Soroush Vosoughi et al., *The Spread of True and False News Online*, 359 SCIENCE 1146, 1146 (2018).

of the ever-changing affairs of men, in whose flux there is nothing more permanent than the admittedly relative permanence of the human mind's structure.⁸

With the help of social media, falsehoods are increasingly credible, and they pose a serious threat to democratic aspirations.⁹ Russian interference in the 2016 presidential election is a prominent example.¹⁰ Domestic political actors are mimicking Russia's tactics in response.¹¹ To take just one example from 2019, a rumor that Senator Elizabeth Warren displayed a doll in blackface in her kitchen began on 4chan, but rapidly spread to mainstream platforms, such as Facebook and Twitter.¹²

Lies can go viral in a shockingly short time,¹³ and for reasons that remain imperfectly understood, false statements appear to spread more quickly than true ones.¹⁴ Deepfakes — videos appearing to depict people saying or doing things they never said or did — are not merely on the horizon; they are here.¹⁵ Doctored videos are less technologically advanced, but their effects on viewers are similar.¹⁶ Neither deepfakes nor doctored videos make a literal statement that is false. They do not literally say, “up is down” or “two plus two equals six,” but their effects are similar to those of false statements: they display something with respect to people or events that is false.

On many occasions, the Supreme Court has suggested that false statements lack constitutional protection.¹⁷ For most of American history, observers might well have concluded that the government had

8. ARENDT, *supra* note 1, at 227.

9. See generally JOHAN FARKAS & JANNICK SCHOU, *POST-TRUTH, FAKE NEWS AND DEMOCRACY* (2020) (exploring how fake news, post-truthism, and alternative facts are shaping society and arguing that truth must be at the heart of democracy, even at the expense of popular sovereignty and the possibility of chilling).

10. See ROBERT S. MUELLER, III, *REPORT ON THE INVESTIGATION INTO RUSSIAN INTERFERENCE IN THE 2016 PRESIDENTIAL ELECTION* 14 (2019), <https://www.justice.gov/storage/report.pdf> [<https://perma.cc/PPD7-96ZC>].

11. See Emily Birnbaum & Olivia Beavers, *Americans Mimic Russian Disinformation Tactics Ahead of 2020*, THE HILL (May 8, 2019, 6:00 AM), <https://thehill.com/policy/cybersecurity/442620-americans-mimic-russian-disinformation-tactics-ahead-of-2020> [<https://perma.cc/944D-ZM43>].

12. *Id.*

13. See, e.g., Sapna Maheshwari, *How Fake News Goes Viral: A Case Study*, N.Y. TIMES (Nov. 20, 2016), <https://nyti.ms/2fdzLBD> [<https://perma.cc/X5AA-QU96>].

14. See David M. J. Lazer et al., *The Science of Fake News*, 359 SCIENCE 1094, 1095 (2018); Vosoughi et al., *supra* note 7, at 1149 (explaining that the quick spread of falsehoods may be due to their novelty).

15. See Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1757 (2019).

16. *Cf.* Corp. Training Unlimited, Inc. v. Nat'l Broad. Co., 868 F. Supp. 501, 507–08 (E.D.N.Y. 1994).

17. See *Illinois ex rel. Madigan v. Telemarketing Assocs., Inc.*, 538 U.S. 600, 612 (2003) (“Like other forms of public deception, fraudulent charitable solicitation is unprotected

broad authority to punish such statements. As noted, it was not until 2012 that the Court made it clear that false statements fall within the ambit of the First Amendment. In *United States v. Alvarez*,¹⁸ a badly divided Court held that the First Amendment prohibited the criminal prosecution of a person who falsely claimed that he was a recipient of the Congressional Medal of Honor.¹⁹ Among other things, the Justices who agreed with the result emphasized that punishing false speech would deter free debate²⁰ and that less restrictive alternatives, such as counterspeech, could promote the state's legitimate interests²¹ (such as not wanting to dilute the effect of actually receiving the Congressional Medal of Honor).

In explaining why this intentional falsehood was protected, the plurality spoke grandly:

Permitting the government to decree this speech to be a criminal offense, whether shouted from the rooftops or made in a barely audible whisper, would endorse government authority to compile a list of subjects about which false statements are punishable. That governmental power has no clear limiting principle. Our constitutional tradition stands against the idea that we need Oceania's Ministry of Truth. Were this law to be sustained, there could be an endless list of

speech."); *BE&K Constr. Co. v. NLRB*, 536 U.S. 516, 531 (2002) ("[F]alse statements may be unprotected for their own sake."); *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 52 (1988) ("False statements of fact are particularly valueless; they interfere with the truth-seeking function of the marketplace of ideas, and they cause damage to an individual's reputation that cannot easily be repaired by counterspeech, however persuasive or effective."); *Keeton v. Hustler Magazine, Inc.*, 465 U.S. 770, 776 (1984) ("There is 'no constitutional value in false statements of fact.'" (quoting *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 340 (1974))); *Bill Johnson's Rests., Inc. v. NLRB*, 461 U.S. 731, 743 (1983) ("[F]alse statements are not immunized by the First Amendment right to freedom of speech."); *Herbert v. Lando*, 441 U.S. 153, 171 (1979) ("Spreading false information in and of itself carries no First Amendment credentials."); *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc.*, 425 U.S. 748, 771 (1976) ("Untruthful speech, commercial or otherwise, has never been protected for its own sake."); *Gertz*, 418 U.S. at 340 ("[T]he erroneous statement of fact is not worthy of constitutional protection."); *Time, Inc. v. Hill*, 385 U.S. 374, 389 (1967) ("[T]he constitutional guarantees [of the First Amendment] can tolerate sanctions against *calculated* falsehood without significant impairment of their essential function."); *Garrison v. Louisiana*, 379 U.S. 64, 75 (1964) ("[T]he knowingly false statement and the false statement made with reckless disregard of the truth, do not enjoy constitutional protection.").

18. *United States v. Alvarez*, 567 U.S. 709 (2012).

19. *Id.* at 730 (plurality opinion).

20. *Id.* at 718 ("[S]ome false statements are inevitable if there is to be an open and vigorous expression of views.").

21. *Id.* at 710, 726.

subjects the National Government or the States could single out.²²

The plurality said that content-based restrictions on false statements must “satisfy exacting scrutiny,”²³ which looks something like traditional “strict scrutiny.” Speaking only for himself and Justice Kagan, Justice Breyer called for a somewhat-lower standard: “intermediate scrutiny” or a proportionality test.²⁴ As he put it, the goal “is to offer proper protection in the many instances in which a statute adversely affects constitutionally protected interests but warrants neither near-automatic condemnation . . . nor near-automatic approval.”²⁵

In *Alvarez*, six members of the Court expressed an enthusiastic commitment to the “marketplace of ideas.”²⁶ In one of the most important opinions in all of American law, Justice Holmes gave voice to that commitment in arguing that “the ultimate good desired is better reached by free trade in ideas — that the best test of truth is the power of the thought to get itself accepted in the competition of the market.”²⁷ In *Alvarez*, the plurality invoked those words when making it plain that the Court “has never endorsed the categorical rule the Government advances: that false statements receive no First Amendment protection.”²⁸ But in the relatively few years since *Alvarez*, the world has changed dramatically, not least because of the increasing role of social media and the spread of falsehoods on it. As we shall see, the plurality in *Alvarez* was myopic in focusing largely on established categories of cases, such as libel, in which false statements of fact can sometimes be regulated or sanctioned.²⁹ In the modern era, false statements falling short of libel can still cause serious problems for individuals and the

22. *Id.* at 723 (citing GEORGE ORWELL, *NINETEEN EIGHTY-FOUR* (Centennial ed. 2003) (1949)).

23. *Id.* at 724.

24. *Id.* at 730–31 (Breyer, J., concurring).

25. *Id.* at 731.

26. For objections, see generally Derek E. Bambauer, *Shopping Badly: Cognitive Biases, Communications, and the Fallacy of the Marketplace of Ideas*, 77 U. COLO. L. REV. 649 (2006); Vincent Blasi, *Holmes and the Marketplace of Ideas*, 2004 SUP. CT. REV. 1; Paul H. Brietzke, *How and Why the Marketplace of Ideas Fails*, 31 VAL. U. L. REV. 951 (1997); R. H. Coase, *Advertising and Free Speech*, 6 J. LEGAL STUD. 1 (1977); R. H. Coase, *The Market for Goods and the Market for Ideas*, 64 AM. ECON. REV. 384 (1974); Aaron Director, *The Parity of the Economic Market Place*, 7 J.L. & ECON. 1 (1964); Stanley Ingber, *The Marketplace of Ideas: A Legitimizing Myth*, 1984 DUKE L.J. 1; William P. Marshall, *In Defense of the Search for Truth as a First Amendment Justification*, 30 GA. L. REV. 1 (1995).

27. *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

28. *Alvarez*, 567 U.S. at 719 (plurality opinion).

29. As the *Alvarez* Court pointed out, there are a few “historic and traditional categories” of speech where content-based restrictions are permissible. *Id.* at 717 (quoting *United States v. Stevens*, 559 U.S. 460, 468 (2010)). These categories include, but are not limited to, speech “intended, and likely, to incite imminent lawless action, obscenity, defamation, speech integral to criminal conduct, [and] so-called ‘fighting words.’” *Id.*

society, even if they do not fit within established categories. If they cause such problems, there is a legitimate argument that they should be regulable.

No one should doubt that for some falsehoods, the marketplace works exceedingly poorly;³⁰ it can be the problem, not the solution, perhaps especially online.³¹ Far from being the best test of truth, the marketplace ensures that many people accept falsehoods or take mere fragments of lives or small events as representative of some alarming or despicable whole.³² Behavioral science makes this point entirely clear³³: scientific research has almost uniformly rejected the idea that “the truth of a proposition is the dominant factor in determining which propositions will be accepted.”³⁴

Suppose, for example, that on Twitter, an emotionally gripping falsehood is starting to spread about a high-level political official or the leader of a large company. Or consider the potential consequences of a statement on Facebook about criminal behavior by a neighbor of yours, someone with no access to the media and without credibility online. The problem is serious and pervasive, and it seems to be mounting.³⁵ On occasion, it results in serious harm to people’s lives,³⁶ damages the

30. See, e.g., Jeff Zeleny, *Persistent “Birthers” Fringe Disorients Strategists*, N.Y. TIMES (Aug. 4, 2009), <https://www.nytimes.com/2009/08/05/us/politics/05zeleny.html> [<https://perma.cc/HVJ4-THHD>] (noting that cable TV and talk radio hosts have “dedicat[ed] countless hours” to a theory that President Obama was born in Kenya).

31. See generally Lazer et al., *supra* note 14; Vosoughi et al., *supra* note 7. The dramatic rise of online news has introduced an assortment of problems that have challenged the marketplace of ideas metaphor in new ways. For discussion of such problems, including news deserts, echo chambers, troll armies, and flooding, see Martha Minow, *The Changing Ecosystem of News and Challenges for Freedom of the Press*, 64 LOY. L. REV. 499, 503–18 (2018); and TIM WU, IS THE FIRST AMENDMENT OBSOLETE? 11–17 (2017), <https://s3.amazonaws.com/kfai-documents/documents/5d8a0f848d/Is-the-First-Amendment-Obsolete-.pdf> [<https://perma.cc/P724-SRUJ>].

32. See Schauer, *supra* note 5, at 911–12.

33. See Lazer et al., *supra* note 14; Schauer, *supra* note 5, at 911; Vosoughi et al., *supra* note 7, at 1147–48.

34. Schauer, *supra* note 5, at 911 (explaining that despite such research, “free speech claimants . . . trot out the tired old clichés that are little more than modern variants on Milton’s now-legendary but almost certainly inaccurate paean to the pervasiveness and power of human rationality”).

35. See *id.* at 911–12.

36. See, e.g., Nina Berman, *The Victims of Fake News*, COLUM. JOURNALISM REV. (Fall 2017), https://www.cjr.org/special_report/fake-news-pizzagate-seth-rich-newtown-sandy-hook.php [<https://perma.cc/J3RN-6UAH>] (interviewing victims and subjects of fake news).

prospects of businesses,³⁷ hurts investors,³⁸ and undermines democracy itself.³⁹

It is important to underline the last point in particular. As Justice Thomas noted, “the common law deemed libels against public figures to be, if anything, *more* serious and injurious than ordinary libels.”⁴⁰ In 1808, a court stressed that “the people may be deceived, and reject the best citizens, to their great injury.”⁴¹ Free speech is meant, in large part, to promote self-government; a well-functioning democracy cannot exist unless people are able to say what they think, even if what they think is false.⁴² But if people spread false statements — most obviously about public officials and institutions — democracy itself will suffer. For no good reason, citizens might lose faith in particular leaders and policies, and even in their government itself. Acting strategically, candidates, parties, outsiders, or others can try to make that happen.⁴³ At the same time, false statements impede our ability to think well, as citizens, about those who do or might lead, or about what to do about a crisis, whether large or small.

We can better understand the problem if we note that in ordinary life, many human beings seem to follow a simple rule: *people generally do not say things unless they are true, or at least substantially true*. We tend to be “credulous Bayesians,” in the sense that we update on the basis of what we hear, but do not sufficiently discount the motivations or limited information of the source of what we hear.⁴⁴ If someone says that a doctor is a criminal, that some student or professor has engaged

37. See, e.g., Matthew F. Ferraro & Jason C. Chipman, *Fake News Threatens Our Businesses, Not Just Our Politics*, WASH. POST (Feb. 8, 2019, 1:33 PM), https://www.washingtonpost.com/outlook/fake-news-threatens-our-businesses-not-just-our-politics/2019/02/08/f669b62c-2b1f-11e9-984d-9b8fba003e81_story.html [<https://perma.cc/AWF3-ZBTH>].

38. See, e.g., Claire Atkinson, *Fake News Can Cause “Irreversible Damage” to Companies — And Sink Their Stock Price*, NBC NEWS (Apr. 25, 2019, 12:54 PM), <https://www.nbcnews.com/business/business-news/fake-news-can-cause-irreversible-damage-companies-sink-their-stock-n995436> [<https://perma.cc/CQ4W-PHUP>].

39. See Minow, *supra* note 31, at 543–44.

40. *McKee v. Cosby*, 139 S. Ct. 675, 679 (2019) (Thomas, J., concurring in denial of certiorari).

41. *Commonwealth v. Clap*, 4 Mass. (3 Tyng) 163, 169–70 (1808).

42. See Lee C. Bollinger & Geoffrey R. Stone, *Dialogue*, in *THE FREE SPEECH CENTURY* 1, 5 (Geoffrey R. Stone & Lee C. Bollinger eds., 2018).

43. See, e.g., Drew Harwell, *Doctored Images Have Become a Fact of Life for Political Campaigns. When They’re Disproved, Believers “Just Don’t Care.”*, WASH. POST (Jan. 14, 2020, 7:00 AM), <https://www.washingtonpost.com/technology/2020/01/14/doctored-political-images/> [<https://perma.cc/8RQ3-M3VE>]; Mike Isaac & Cecilia Kang, *Facebook Says It Won’t Back Down from Allowing Lies in Political Ads*, N.Y. TIMES (Jan. 9, 2020), <https://nyti.ms/2QFvehH> [<https://perma.cc/H9TC-8RGQ>].

44. Edward L. Glaeser & Cass R. Sunstein, *Extremism and Social Learning*, 1 J. LEGAL ANALYSIS 263, 265 (2009). Of course, it is also true that the source of the information matters, especially if it is drawn to people’s attention. If cigarette companies say that the risks of cigarette smoking are small, people are not likely to be much moved.

in some terrible misconduct, or that a candidate for public office is corrupt, many people will think that the statement would not have been made unless it had some basis in fact. On this view, there is fire wherever there is smoke. And even if most of us are not so credulous and do not adhere to such a rule, the mere presence of a false statement can leave a cloud of suspicion, a kind of negative feeling or affective after-effect that can ultimately influence our beliefs and behavior.⁴⁵ If we hear, “Politician Jones never committed assault,” our minds will nonetheless associate Jones with assault. In the terms of a metaphor often used in behavioral science, the parts of our mind that make fast and intuitive choices and that are sometimes driven by emotion (often called “System 1”) might credit the false statement, even if the slower, more deliberative, and more calculative parts (“System 2”) do not.⁴⁶

It is true and important that any effort to regulate speech will create a chilling effect.⁴⁷ Punish people for spreading falsehoods, and you will find yourself chilling truth. In *Alvarez*, Justice Breyer noted that “the threat of criminal prosecution for making a false statement can inhibit the speaker from making true statements, thereby ‘chilling’ a kind of speech that lies at the First Amendment’s heart.”⁴⁸ Suppose, for example, that a law will hold people accountable if they circulate a false statement about a presidential candidate. To be sure, it is good if people — voters as well as the candidate personally — are not injured as a result of that false statement. But that very law will discourage others from disclosing, on the basis of credible evidence, the fact that a candidate has done something wrong or even terrible.⁴⁹

But there is a countervailing consideration. Sometimes a chilling effect can be an excellent safeguard. Without such an effect, the marketplace of ideas will allow many people to spread damaging falsehoods about both individuals and institutions.⁵⁰ If false statements create serious problems, it is important to ensure that the fear of a chilling effect does not itself have a chilling effect on public discussion or social practices. Some falsehoods can hurt or even ruin individual lives. For all these reasons, it is sensible to hope that social norms and

45. See Danielle Polage, *The Effect of Telling Lies on Belief in the Truth*, 13 EUROPE’S J. PSYCHOL. 633, 639 (2017).

46. See DANIEL KAHNEMAN, THINKING, FAST AND SLOW 20 (2011).

47. A classic discussion is Frederick Schauer, *Fear, Risk and the First Amendment: Unraveling the “Chilling Effect,”* 58 B.U. L. REV. 685 (1978).

48. *United States v. Alvarez*, 567 U.S. 709, 733 (2012) (Breyer, J., concurring).

49. In *Alvarez*, the dissenting opinion insisted: “The lies covered by the Stolen Valor Act have no intrinsic value and thus merit no First Amendment protection unless their prohibition would chill other expression that falls within the Amendment’s scope.” *Id.* at 750 (Alito, J., dissenting).

50. See Vosoughi et al., *supra* note 7, at 1147.

even laws will chill them. We need, in short, to find ways to discourage the spread of statements that are at once false and damaging.⁵¹

The remainder of this Article comes as follows. Part II, the heart of the Article, offers a general discussion of why and when falsehoods deserve constitutional protection. To state the central conclusion too simply: In general and for multiple reasons, falsehoods can contribute to public debate, which makes it reasonable to adopt a rebuttable presumption that they may not be prohibited unless the government is able to make a strong demonstration of harm. But it is important to emphasize that singly or together, the arguments in favor of protecting falsehoods are not conclusive, and with respect to some falsehoods, they are palpably inadequate.

Part III turns to the specific issue of reputational injury, with particular reference to libel, perhaps the most important domain in which the Supreme Court has grappled with the constitutional status of falsehoods under the First Amendment. I argue that *New York Times Co. v. Sullivan* badly overshot the mark and that it is ill-suited to the current era. Those who have been injured by negligent falsehoods, including public figures, should not be remediless. The real question is: what remedies should they have? Part III also explores possible responses to intentional falsehoods that are not libelous, including positive statements that give people or institutions credit for things they never did.

Part IV turns to deepfakes and doctored videos, which present novel threats to both individual lives and democratic self-government. Deepfakes and doctored videos are of interest in themselves, and they also put a bright spotlight on some of the challenges posed by falsehoods. The basic claim is that if deepfakes and doctored videos are damaging in a relevant sense, and if large numbers of observers are likely to think that they are real, these videos should not be immune from regulations on constitutional grounds. Part V offers a brief conclusion.

II. PROTECTING FALSEHOODS

Why are (some) falsehoods protected by the First Amendment? Consider a few examples:

“The moon landing was faked.” “Pigs really can fly.”
 “The United States military carried out the 9/11 attacks.” “The stock market is at an all-time low.” “Ruth Bader Ginsburg is Chief Justice of the United States

51. *See, e.g.*, *BE&K Constr. Co. v. NLRB*, 536 U.S. 516, 531 (2002) (“[F]alse statements may be unprotected for their own sake.”); *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 52 (1988) (“False statements of fact are particularly valueless.”); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 340 (1974) (“[T]here is no constitutional value in false statements of fact.”).

Supreme Court.” “Dropped objects don’t fall.” “Bob Dylan did not write any songs.” “The Holocaust never happened.” “The real unemployment rate in the United States is at least seventy percent.” “Senator Elizabeth Warren is a Russian agent — bought and paid.” “Plastic is a kind of gold.” “Dogs are descended from coyotes, not wolves.” “The Earth is flat.” “The Easter Bunny is real.” “The U.S. Constitution was ratified in 1727.” “The plays of William Shakespeare were written by Albert Camus.”

Many people believe some of these propositions. Let us stipulate that all of them are demonstrably false. Why should they be protected? Perhaps surprisingly, the Supreme Court has never given a systematic answer to that question. The most famous discussion, and in some ways still the best, comes from John Stuart Mill.⁵² I shall draw on that discussion here. But as we shall see, Mill’s arguments run into serious problems. They are too abstract and general to resolve hard questions, including those posed by libel, deepfakes, and doctored videos.⁵³

Before exploring that objection, it is important to emphasize that in well-functioning societies, restrictions on lying and safeguards against spreading falsehoods come mostly from social norms, not from law.⁵⁴ Those who do not tell the truth are likely to face social sanctions of various kinds.⁵⁵ People are aware of that risk in advance, and so generally internalize the norm in favor of honesty, such that they feel shame or guilt if they violate it. In addition, a norm of truth-telling can be seen, in some contexts, as a solution to a prisoner’s dilemma: people know that if they lie, their rivals or opponents might lie as well, and a norm in favor of truthfulness avoids a kind of mutual destruction.⁵⁶ If social norms punish lying or the spread of falsehoods, there is much less pressure on the legal system to intervene.

52. See generally JOHN STUART MILL, *ON LIBERTY AND THE SUBJECTION OF WOMEN* 22–63 (Alan Ryan ed., Penguin Books 2006) (1859).

53. I am bracketing here the fact that some falsehoods, specifically those that are not typically subject to regulation, have unmistakable value. As Justice Breyer put it in *Alvarez*: “False factual statements can serve useful human objectives[.] . . . [T]hey may prevent embarrassment, protect privacy, shield a person from prejudice, provide the sick with comfort, or preserve a child’s innocence; . . . they may stop a panic or otherwise preserve calm in the face of danger; and even . . . can promote a form of thought that ultimately helps realize the truth.” *Alvarez*, 567 U.S. at 733 (Breyer, J., concurring).

54. For relevant discussion, see Eric Talley, *Disclosure Norms*, 149 U. PA. L. REV. 1955, 2010–11 (2001).

55. See, e.g., Heiko Rauhut, *Beliefs About Lying and Spreading of Dishonesty: Undetected Lies and Their Constructive and Destructive Social Dynamics in Dice Experiments*, PLOS ONE, Nov. 2013, at 1, 5 (2013).

56. See EDNA ULLMANN-MARGALIT, *THE EMERGENCE OF NORMS* 14 (1977).

In this light, we can see the use of the common law and regulatory interventions as a response to the imperfect power of norms in certain contexts. When norms break down, there are likely to be increasing pleas for a more aggressive legal response. Suppose, for example, that in certain sectors, norms only weakly favor truth-telling, and that reputational, economic, political, or other incentives favor lying. Or suppose that it is easy for people to speak anonymously or otherwise to avoid disclosure of their identities. If so, many people will want to invoke longstanding legal sanctions on falsehoods (such as libel law) or perhaps invent new ones. Where norms break down or prove inadequate, the demand for legal intervention grows.

A. Official Fallibility

If the government is allowed to punish or censor what it characterizes as false, it might actually end up punishing or censoring truth. The reason is that its own judgments may not be reliable. In defending the right to say what is false, Mill made much of this point, arguing that those who seek to suppress speech “of course deny its truth; but they are not infallible All silencing of discussion is an assumption of infallibility.”⁵⁷ Public officials might be foolish or ignorant. However confident, they might be wrong. Their judgments might be self-serving. On Mill’s view, official fallibility is a sufficient reason to allow what officials deem to be falsehoods — and to allow public discussion and counterspeech to provide a corrective, if a corrective is what is needed.

Taken by itself, Mill’s argument is unconvincing. To be sure, it suggests a serious cautionary note, especially about the proposition that officials should be allowed to censor or punish whatever they claim or deem to be falsehoods. But with appropriate safeguards, a well-functioning legal system ought to be able to reduce the risk of official error without giving free rein to lies and liars, and even to sincerely held falsehoods. One safeguard is an independent tribunal; a court, and not merely the executive branch, should be the institution to resolve the question of truth or falsity. It is true and important that courts are not infallible. Their factfinding tools are hardly perfect, and they might well have biases of their own. Nonetheless, an independent tribunal can serve as an important check on the potentially self-serving judgments of others, including executive officials.⁵⁸

57. MILL, *supra* note 52, at 24.

58. Insofar as we are dealing with purely private claims, as when one citizen sues a newspaper or broadcaster for damages, fears about executive bias or overreaching are not implicated.

Another safeguard is the burden of proof. With what clarity must it be shown that a statement is in fact false?⁵⁹ With a high burden of proof, the risk of unjustified restrictions on speech is diminished. If we are very concerned about the risk of error, we might say that falsehoods may be punished or censured only when an independent tribunal has concluded that there is no reasonable doubt about the matter — as, for example, when someone has clearly been libeled with a false claim or accusation. It should be clear that as the burden of proof grows increasingly difficult to meet, Mill’s argument becomes increasingly weak. At the same time, a high burden of proof will also ensure that false and harmful statements are not chilled.

In direct response to Mill’s claim, it is important to note that the legal system has long forbidden various kinds of falsehoods, including perjury,⁶⁰ false advertising,⁶¹ and fraud.⁶² People are not allowed to say that they are government agents unless they actually are.⁶³ In these cases, Mill’s argument has been rightly and thoroughly rejected. It is true that in such cases, there is usually demonstrable harm,⁶⁴ and we might want to say that false statements are protected unless there actually is such harm. But the question here is whether Mill’s argument about official fallibility is, by itself, a convincing reason to protect falsehoods. It is not.

The most plausible argument in Mill’s defense would be rule-consequentialist. Perhaps the risk of official error (including judicial error) is high — not always, but much of the time. Perhaps case-by-case inquiries into the question of truth or falsity would be burdensome and time-consuming. Perhaps we do best to avoid those inquiries and conclude that officials cannot punish or censor falsehoods unless they can make a powerful showing of harm. That argument is not implausible in the abstract, but it too is unconvincing, at least in well-functioning legal systems. We should be able to agree that there is no point in regulating speech unless it is harmful. But if falsehoods are a little bit harmful, should they be protected merely because officials are fallible? If people lie to the Federal Bureau of Investigation about their employment history, or falsely claim to be a government agent when trying to frighten

59. *Compare* *Phila. Newspapers, Inc. v. Hepps*, 475 U.S. 767, 777 (1986) (explaining that a “private figure plaintiff must bear the burden of showing that the speech at issue is false before recovering damages for defamation from a media defendant”), *with* *Consol. Edison Co. v. Pub. Serv. Comm’n*, 447 U.S. 530, 540 (1980) (explaining that when a state actor restricts speech, it is the state’s burden to show that such restriction is justified).

60. 18 U.S.C. § 1621 (2018).

61. 15 U.S.C. § 54 (2018).

62. 18 U.S.C. § 1001.

63. 18 U.S.C. § 912 (“Whoever falsely assumes or pretends to be an officer or employee acting under the authority of the United States or any department, agency or officer thereof, and acts as such . . . shall be fined under this title or imprisoned.”).

64. *See* *United States v. Alvarez*, 567 U.S. 709, 717–19 (2012) (plurality opinion).

a neighbor, or say in a job application for state employment that they competed in the Olympics, is it so clear that the First Amendment should be taken to prohibit punishment, even if the harm done by those lies is relatively modest?

B. Chilling Truth and Public Goods

A different reason to protect falsehoods has nothing to do with official fallibility. It is that allowing government to punish or censor what is false might deter people from saying what is true. As four Justices put it in *Alvarez*:

Were the Court to hold that the interest in truthful discourse alone is sufficient to sustain a ban on speech, . . . it would give government a broad censorial power unprecedented in this Court's cases or in our constitutional tradition. The mere potential for the exercise of that power casts a chill, a chill the First Amendment cannot permit if free speech, thought, and discourse are to remain a foundation of our freedom.⁶⁵

So much, in this light, for the False Information Act.

No one should doubt that if people could be punished for saying something false, they might silence themselves. The mere possibility of a criminal or civil proceeding might induce self-silencing.⁶⁶ To be sure, this problem could be reduced if the legal system had a perfect technology for detecting falsehoods; people could then be confident that so long as they told the truth, they could not be punished. But with many of their statements, people have *degrees of certainty*. They might be fifty-one percent confident, sixty percent confident, eighty percent confident, or ninety-five percent confident. If falsehoods are punishable, people might not speak out unless they are essentially certain — which would be a significant loss to speakers and society as a whole. In the face of potential punishment, loss aversion might lead people to shut up.⁶⁷ What kind of democracy, and what kind of society, insists that people shut up unless they know that they are right?

Justice Breyer may well have something like this point in mind in insisting that “[l]aws restricting false statements about philosophy, religion, history, the social sciences, the arts, and the like raise [serious]

65. *Id.* at 723.

66. See Note, *The First Amendment Overbreadth Doctrine*, 83 HARV. L. REV. 844, 853–54 (1970).

67. See EYAL ZAMIR, LAW, PSYCHOLOGY, AND MORALITY: THE ROLE OF LOSS AVERSION 15 (2014).

concerns, and in many contexts have called for strict scrutiny.⁶⁸ Justice Alito combined that point with Mill's concern about institutional fallibility:

[T]here are broad areas in which any attempt by the state to penalize purportedly false speech would present a grave and unacceptable danger of suppressing truthful speech. Laws restricting false statements about philosophy, religion, history, the social sciences, the arts, and other matters of public concern would present such a threat. The point is not that there is no such thing as truth or falsity in these areas or that the truth is always impossible to ascertain, but rather that it is perilous to permit the state to be the arbiter of truth.⁶⁹

This is an important claim. We can fortify it by emphasizing that (true) information confers benefits on society as a whole, not merely on speakers.⁷⁰ If someone discloses something of importance, many people will benefit. The point is nicely captured by Joseph Raz:

If I were to choose between living in a society which enjoys freedom of expression, but not having the right myself, or enjoying the right in a society which does not have it, I would have no hesitation in judging that my own personal interest is better served by the first option.⁷¹

When free speech principles are designed to protect against a chilling effect, speakers benefit of course. But the larger goal is to help countless others.

On social welfare grounds, however, we should be cautious before accepting the view that a chilling effect is a sufficient reason to protect lies and other falsehoods. Why should a chilling effect be a trump card? Suppose that what is chilled contains a great deal of falsity and a small amount of truth. If so, it is hardly clear that a ban on falsity is unjustified. At a minimum, we would need to know the magnitude of the chilling effect, and also the harm produced by chilling truth, along with the benefit produced by chilling falsehood. No one thinks that the ban on perjury should be lifted because it also chills truthful testimony

68. *Alvarez*, 567 U.S. at 731–32 (Breyer, J., concurring).

69. *Id.* at 751–52 (Alito, J., dissenting).

70. See Alexander Tsesis, *Free Speech Constitutionalism*, 2015 U. ILL. L. REV. 1015, 1034–35.

71. JOSEPH RAZ, *ETHICS IN THE PUBLIC DOMAIN* 39 (1994).

(though it undoubtedly does at least a little of that). It is not reasonable to think that a ban on false commercial speech should be lifted because it also chills truthful commercial advertising (though it undoubtedly does at least a little of that). Finally, the Supreme Court has allowed a construction contractor to bring a defamation action against a credit reporting agency for falsely reporting that the contractor had filed for bankruptcy (even though the Court's decision is highly likely to deter true statements as well).⁷² The fact that banning falsity chills truth is a relevant consideration, but it is hardly a decisive point in favor of constitutional protection of false statements.

C. Learning from Falsity

If the government could be trusted and if the chilling effect was modest or inconsequential, ought we tolerate false statements of fact? That question raises this one: what is the social value of unquestionably false statements? A strong possibility, signaled by Mill, is that they can improve understanding:

However unwillingly a person who has a strong opinion may admit the possibility that his opinion may be false, he ought to be moved by the consideration that however true it may be, if it is not fully, frequently, and fearlessly discussed, it will be held as a dead dogma, not a living truth If the cultivation of the understanding consists in one thing more than in another, it is surely in learning the grounds of one's own opinions He who knows only his own side of the case, knows little of that. His reasons may be good, and no one may have been able to refute them. But if he is equally unable to refute the reasons on the opposite side; if he does not so much as know what they are, he has no ground for preferring either opinion.⁷³

Mill added that false statements can bring about "the clearer perception and livelier impression of truth, produced by its collision of error."⁷⁴ Mill's point holds for a large number of false statements.⁷⁵ If people are told that the moon landing was faked or that the Holocaust never happened, they can learn more about the truth of these matters —

72. *See* *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 763 (1985).

73. MILL, *supra* note 52, at 42–44.

74. *Id.* at 23.

75. *But see* *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 52 (1988) ("False statements of fact are particularly valueless [because] they interfere with the truth-seeking function of the marketplace of ideas.").

but only if the statements are not censored. That point is sufficient to suggest that falsity, by itself, should not be taken as a *general* reason to allow punishment or censorship.

But for many questions, Mill's argument is too abstract and high-minded. If someone falsely tells a federal investigator that an applicant for federal employment is a cocaine addict or proclaims that she won the Congressional Medal of Honor when she actually did not, should we say that people have been enabled to discover a "living truth"? If someone impersonates a police officer, is it a good idea to force people to find out that she is doing that?⁷⁶ The problem is amplified by the fact that human beings have limited attention.⁷⁷ The effort required to find out whether a statement is true or false may mean that people will simply believe it, certainly if it fits with their antecedent convictions.⁷⁸ The existence of online echo chambers, in which many people sort themselves into groups of the like-minded, further compounds the problem.⁷⁹

Mill was right to say that people can learn from false statements and that living truths are better than dead dogmas. But here again, his argument does not provide a convincing basis for general protection of falsehoods, intentional or otherwise, even in cases in which they do not produce clear and imminent harm.

D. Learning What Others Think

There is a related point, not emphasized by Mill. When people hear falsehoods, they can learn more about what other people think, and why. If people find out that many of their fellow citizens believe that President Barack Obama was not born in the United States or that the United States was responsible for the attacks of 9/11, they will learn something that is important to know. "Pluralistic ignorance," understood as ignorance about what other people actually think, can be a serious problem.⁸⁰ If falsehoods can be spoken and written, people will be better able to obtain a sense of the distribution of views within their society. That can be a large benefit. For one thing, it can give people a sense that their own views might be wrong; it can crack a wall of certainty. For another thing, it can provide people with information that is

76. See *United States v. Chappell*, 691 F.3d 388 (4th Cir. 2012) (ruling that the First Amendment does not ban the "Virginia police impersonation statute, . . . [which] prohibits individuals from falsely assuming or pretending to be a law enforcement officer").

77. See Xavier Gabaix, *Behavioral Inattention 5* (Nat'l Bureau of Econ. Research, Working Paper No. 24096, 2018).

78. See Cass R. Sunstein et al., *How People Update Beliefs About Climate Change: Good News and Bad News*, 102 CORNELL L. REV. 1431, 1433 (2017).

79. See CASS R. SUNSTEIN, #REPUBLIC: DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA 114–15 (2017).

80. See TIMUR KURAN, PRIVATE TRUTHS, PUBLIC LIES 78 (1997).

important to know. If many of one's fellow citizens believe that the Earth is flat, that vaccines cause autism, or that climate change is a hoax invented by the Chinese government, citizens benefit from obtaining that knowledge.

Again, this is not a decisive argument in favor of allowing falsehoods. It works best for false beliefs that are sincerely held; it is much harder to understand the argument as a reason to protect lies. And even for sincerely held false beliefs, it is inadequate. The benefit of learning what others think might be outweighed by the cost of allowing falsehoods to spread. The only point is that it is a benefit.

E. Counterspeech > Bans

The final point is pragmatic and in some ways a generalization. Banning or punishing falsehoods might simply drive beliefs underground.⁸¹ They might not be exposed to the light. They can be fueled by the very fact that they are forbidden.⁸² If the goal is to reduce their power, allowing falsehoods to have some oxygen and forcing people to meet them with counterarguments might be best.⁸³ A law might forbid denial of the Holocaust, but from the standpoint of the very people who support that the premise of the law, freedom of speech might be best. One reason is that suppression of speech might intensify people's commitment to the very falsehoods that the speech contains. Another reason is that suppression might create a kind of forbidden fruit, broadening the appeal of those falsehoods. Yet another reason is that suppression might be taken as an attack on individual autonomy. Isn't it better to convince people, rather than to shut them up?

The now-familiar response is that these points are far too abstract to resolve concrete cases. With respect to Holocaust denial, is oxygen a good idea? In every nation? That is hardly clear. Nor is it clear that it applies to lies about what people have said or done. Consider *Alvarez* itself: If someone falsely claims to have won the Congressional Medal of Honor, is it really such a bad idea to drive that claim underground?

81. See Kent Greenfield, *The Limits of Free Speech*, THE ATLANTIC (Mar. 13, 2015), <https://www.theatlantic.com/politics/archive/2015/03/the-limits-of-free-speech/387718> [<https://perma.cc/DW7Z-8T2M>].

82. See Joanna Plucinska, *Hate Speech Thrives Underground*, POLITICO (Feb. 14, 2018, 2:59 PM), <https://www.politico.eu/article/hate-speech-and-terrorist-content-proliferate-on-web-beyond-eu-reach-experts> [<https://perma.cc/FK9N-ED7Z>].

83. See, e.g., *United States v. Alvarez*, 567 U.S. 709, 726 (2012) (plurality opinion) (explaining that "the dynamics of free speech, of counterspeech, of refutation, can overcome" falsehoods); see also Jeffery C. Barnum, *Encouraging Congress to Encourage Speech: Reflections on United States v. Alvarez*, 76 ALB. L. REV. 527, 543–45 (2012); Louis W. Tompros et al., *The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in a Post-Alvarez, Social Media-Obsessed World*, 31 HARV. J.L. & TECH. 65, 106–07 (2017) (describing useful features of counterspeech).

F. Harm in the Balance

When falsehoods are banned, it is not only because they are falsehoods, but also because they threaten to produce real harm. Consider the canonical example of an unprotected falsehood: a false cry of “fire!” in a crowded theater.⁸⁴ Such a cry is not merely false. It also threatens, with a sufficiently high probability, to cause serious harm — and under plausible assumptions, little or nothing can be done, in time, to prevent that harm. Some falsehoods create a clear and present danger.⁸⁵ If someone sells a product with a false claim that it cures some illness,⁸⁶ there is an immediate risk that people will waste their money on it (and fail to take appropriate steps to address that illness). If someone commits perjury⁸⁷ — by saying, for example, that she saw a defendant at the scene of a crime, when she saw no such thing — the likelihood of an unjust conviction is immediately increased. If someone claims to be a government agent,⁸⁸ and so induces some action or statement from someone, there is immediate harm. All of these acts are unlawful. In fact, many falsehoods that the law forbids create something like a clear and present danger. But it is difficult to say something like that about other false statements of fact, which are more innocuous. (Consider, for example, a claim from one neighbor to another, that her dog — who is actually a mixed breed — is a pure-bred Labrador Retriever, or an exaggerated report about an alleged achievement on the tennis court over the weekend.)

To be sure, many falsehoods can be harmful even if they do not fall in the traditional categories. Holocaust denials might fuel anti-Semitism.⁸⁹ Claims about U.S. responsibility for the 9/11 attacks can discredit efforts to counteract terrorism. False claims about presidents and presidential candidates can be delegitimizing or worse. One more time: If social welfare is the goal, we would want to measure the benefits against the costs of allowing the false statement in question, or perhaps the category of statements of which it is a part.

84. *Schenck v. United States*, 249 U.S. 47, 52 (1919) (“The most stringent protection of free speech would not protect a man in falsely shouting fire in a theatre and causing a panic.”).

85. *See* *Dennis v. United States*, 341 U.S. 494, 503 (1951). The clear and present danger test as understood in *Dennis* is generally understood to have been replaced by the far more stringent test in *Brandenburg v. Ohio*, 395 U.S. 444 (1969), which requires that speech be intended to produce, and is likely to produce, “imminent lawless action,” *id.* at 447. Interestingly, however, the Court has yet to confront a fact pattern akin to that in *Dennis*, and while *Brandenburg* is the more recent pronouncement, it is not entirely clear how the Court would proceed, if it did so.

86. *See* 15 U.S.C. § 54 (2018).

87. *See* 18 U.S.C. § 1621 (2018).

88. *See* 18 U.S.C. §§ 709, 912.

89. Under German law, for example, denial of the Holocaust is forbidden, and the same prohibition can be found in some other nations. *See* ROBERT A. KAHN, HOLOCAUST DENIAL AND THE LAW 6 (2004).

But lacking available evidence, we might reasonably collect the foregoing arguments in support of a rebuttable presumption in favor of protecting false statements⁹⁰: *False statements are protected unless the government can show that allowing them will cause serious harms that cannot be avoided through a more speech-protective route.* In the abstract, it cannot easily be shown that this approach is preferable to more case-by-case alternatives. The answer depends largely on whether institutions are capable of making reliable case-by-case judgments, and if they can do so at low cost. If they are, and if they can, case-by-case judgments are probably better. But in the real world, the presumption is certainly a sensible place to start.

III. FALSE STATEMENTS AND REPUTATIONAL HARM

Let us now turn to more concrete issues. In the United States and many other nations, the law has long attempted to balance the interest in reputation with the interest in free speech. When lawyers and judges discuss that balance, they usually speak of, and deplore, the chilling effect that is created by the prospect of civil or criminal penalties for any kind of speech.⁹¹ Fearing the threat of damage actions, penalties, and lawsuits, whistleblowers, experts, journalists, and bloggers might keep their thoughts and opinions to themselves.⁹² A legal system ought not to discourage questions, objections, and dissent, which can promote accountability and uncover error or corruption.⁹³ Strict rules of libel law, for example, can chill speech about public figures and public issues in a way that could damage democratic debate.⁹⁴ To the degree that there is something like a marketplace of ideas, we should be especially concerned about the chilling effect because it will undermine processes that ultimately produce the truth.⁹⁵

But let us be careful about an undue emphasis on only one side of the equation. There are two points here. *First*, for increasingly clear reasons, the marketplace of ideas can fail, ensuring that false statements will spread and become entrenched. A great deal of behavioral science helps explain why this is the case. There is evidence that falsehoods spread more quickly than truth.⁹⁶ People often engage in “motivated reasoning”; they credit false statements because they *like* believing that

90. See EDNA ULLMANN-MARGALIT, *NORMAL RATIONALITY* 27–31 (Avishai Margalit & Cass R. Sunstein eds., 2017) (discussing rebuttable presumptions generally).

91. See, e.g., Schauer, *supra* note 47, at 697.

92. See, e.g., JULIANNE SCHULTZ, *REVIVING THE FOURTH ESTATE: DEMOCRACY, ACCOUNTABILITY AND THE MEDIA* 47–68 (1998).

93. See, e.g., *id.* at 196.

94. See Schauer, *supra* note 47, at 709–10.

95. See *supra* Section II.B.

96. Vosoughi et al., *supra* note 7, at 1147.

they are true.⁹⁷ Because of social influences, people will follow those whom they trust, even when those whom they trust are not reliable.⁹⁸ As Madison put it, “[t]he reason of man, like man himself, is timid and cautious when left alone, and acquires firmness and confidence in proportion to the number with which it is associated.”⁹⁹ Under certain conditions, correcting falsehoods can be exceptionally difficult.¹⁰⁰ Corrections of false statements might even backfire, entrenching and intensifying people’s commitments to them.¹⁰¹

Second, on occasion, the chilling effect should be welcomed, especially if it comes from social norms that encourage truth telling and that discourage lies.¹⁰² The chilling effect can reduce damaging and destructive falsehoods, including falsehoods about individuals, whether or not famous, and institutions, whether public or private.¹⁰³ As we have seen, some falsehoods are helpful ways of producing the truth in the long run.¹⁰⁴ But many falsehoods are not only damaging but also entirely useless to those who seek to know what is true. A society without any chilling effect, imposed by social norms or by law, would be a singularly ugly place. What societies need is not the absence of “chill,” but an optimal level of it. The question is: how do we get there?

A. Old Law

With respect to reputational harm, current constitutional law offers one possible route. The basic principles are, of course, laid down in *New York Times Co. v. Sullivan*.¹⁰⁵ The decision is familiar, but with recent objections in mind,¹⁰⁶ it will be useful to approach it anew, to understand the context, and to see what the Court actually said.

97. Sunstein et al., *supra* note 78, at 1440; *see also* Michael Thaler, The “Fake News” Effect: An Experiment on Motivated Reasoning and Trust in News 33 (Nov. 1, 2019) (unpublished manuscript) (on file with author). An important finding, specifying motivated reasoning, involves “desirability bias” in belief updating. Ben M. Tappin et al., *The Heart Trumps the Head: Desirability Bias in Political Belief Revision*, 146 J. EXPERIMENTAL PSYCHOL.: GEN. 1143, 1147 (2017).

98. *See* CASS R. SUNSTEIN, CONFORMITY 11–35 (2019).

99. THE FEDERALIST NO. 49 (James Madison).

100. Lazer et al., *supra* note 14, at 1095.

101. *See* Brendan Nyhan & Jason Reifler, *When Corrections Fail: The Persistence of Political Misperceptions*, 32 POL. BEHAV. 303, 308–09 (2010).

102. *See supra* Section II.B.

103. *See* CASS R. SUNSTEIN, ON RUMORS: HOW FALSEHOODS SPREAD, WHY WE BELIEVE THEM, WHAT CAN BE DONE 75 (2014).

104. *See supra* Sections II.C–D.

105. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254 (1964). On the history of libel, *see generally* William L. Prosser, *Libel Per Quod*, 46 VA. L. REV. 839 (1960).

106. *McKee v. Cosby*, 139 S. Ct. 675, 678 (2019) (Thomas, J., concurring in denial of certiorari). The literature on *Sullivan* is voluminous. *See generally* KERMIT L. HALL & MELVIN I. UROFSKY, *NEW YORK TIMES V. SULLIVAN: CIVIL RIGHTS, LIBEL LAW, AND THE FREE PRESS* (2011); ANTHONY LEWIS, *MAKE NO LAW: THE SULLIVAN CASE AND THE FIRST AMENDMENT* (1991); Richard A. Epstein, *Was New York Times v. Sullivan Wrong?*, 53 U.

For present purposes, the bare facts of the case were simple. In the early 1960s, civil rights organizations ran an advertisement in the *New York Times* complaining of brutal police responses to civil rights protests in Montgomery, Alabama.¹⁰⁷ L.B. Sullivan, a Montgomery commissioner with authority over the police, brought a suit for libel.¹⁰⁸ There is no question that the advertisement contained falsehoods about Sullivan, and while some were trivial, those falsehoods could be seen to be injurious to his reputation.¹⁰⁹ There is also no question that the case was, in important respects, a civil rights case, not just a case about freedom of expression. In important respects, the libel action was directed against the civil rights movement.¹¹⁰ It was an effort to make some kind of statement and to exact some kind of revenge.

The Court famously ruled that when a public official is involved, the Constitution allows recovery only if the speaker had “actual malice.”¹¹¹ This means that speakers cannot be held liable unless (a) they are actually aware that the statement was false, or (b) they acted “with reckless disregard” to the question of truth or falsity.¹¹² It follows that speakers are free from liability if they spread falsehoods innocently and in good faith, or even if they acted unreasonably in saying what they did, in the sense that they had reason to know that what they were saying was false.¹¹³ Negligence is not enough for liability. The puzzle here is that while the Court was clear during this period that false statements, as such, did not have First Amendment value, the Court was nonetheless providing a great deal of protection to those very statements.

In explaining its conclusions, the Court stressed that the Constitution limits the government’s power even when the government is attempting to control unquestionably false statements.¹¹⁴ The Court said that “erroneous statement is inevitable in free debate, and . . . it must be protected if the freedoms of expressions are to have the ‘breathing space’ that they ‘need . . . to survive.’”¹¹⁵ Thus neither “factual error”

CHI. L. REV. 782 (1986). For a valuable historical perspective, see generally James Maxwell Koffler, *The Pre-Sullivan Common Law Web of Protection Against Political Defamation Suits*, 47 HOFSTRA L. REV. 153 (2018).

107. *Sullivan*, 376 U.S. at 256–57.

108. *Id.* at 256.

109. *Id.* at 258–59.

110. See HARRY KALVEN, *THE NEGRO AND THE FIRST AMENDMENT* 55 (1965).

111. *Sullivan*, 376 U.S. at 279–80.

112. *Id.* at 280.

113. *Id.* at 280, 286–87; see also *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 776 (1985) (Brennan, J., dissenting).

114. *Sullivan*, 376 U.S. at 271.

115. *Id.* at 271–72 (quoting *NAACP v. Button*, 371 U.S. 415, 433 (1963)).

nor “defamatory content” is enough to remove the constitutional protection accorded to “criticism of official conduct.”¹¹⁶ The Court emphasized that the free speech principle broadly protects speech that bears on public affairs.¹¹⁷

For public officials, the Court ruled two approaches out of bounds. It said that strict liability is constitutionally unacceptable.¹¹⁸ Under the First Amendment, speakers cannot be held liable simply because they spread falsehoods.¹¹⁹ The Court also ruled out the negligence standard for public officials.¹²⁰ Imagine that a falsehood seriously injures a public official and that the speaker should have known (in light of the evidence she had) that she was speaking falsely. Even if so, the newspaper publishing the falsehood is free from liability so long as it did not know that the statement was false and so long as it was not recklessly indifferent to the question of truth or falsity.¹²¹

Because *New York Times Co. v. Sullivan* involved public officials, it left some key questions open. What if someone libels a private individual, someone who lacks fame or notoriety? What if a newspaper publishes some damaging falsehood about Joe Smith, accusing him of corruption, bribery, theft, or other misconduct? Under longstanding principles in Anglo-American law, Smith may recover damages, and he need not even establish fault.¹²² The very facts of falsehood and harm are enough to give Smith a right to sue.¹²³ The Court’s analysis in *New York Times Co. v. Sullivan*, focusing on the risk of chilling effect and the need for breathing space in the context of “criticism of official conduct,” did not by itself raise doubts about Smith’s ability to invoke the courts to protect his reputation.¹²⁴

Nonetheless, the Court eventually concluded that the free speech principle imposes restrictions on Smith’s libel action too — a conclusion that has implications for what is said on Facebook, YouTube, and other online platforms else. In *Gertz v. Robert Welch, Inc.*,¹²⁵ the Court ruled that ordinary people could recover damages for libel only if they

116. *Id.* at 273.

117. *Id.* It is important to note that the Court did not really speak in originalist terms. For one view of the original understanding, see *McKee v. Cosby*, 139 S. Ct. 675, 681 (2019) (Thomas, J., concurring in the denial of certiorari) (“[P]rotections for free speech . . . did not abrogate the common law of libel . . . Public officers and public figures continued to be able to bring civil libel suits for unprivileged statements without showing proof of actual malice as a condition for liability As of 1952, ‘every American jurisdiction . . . punish[ed] libels directed at individuals.’” (citations omitted)).

118. *Sullivan*, 376 U.S. at 278–79.

119. *Id.* at 279.

120. *Id.* at 279–80.

121. *Id.* at 280.

122. See Epstein, *supra* note 106, at 806.

123. *Id.*

124. *Sullivan*, 376 U.S. at 273.

125. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323 (1974).

could prove negligence.¹²⁶ What this means is that if someone says something false and damaging, it is not enough that the statement was false and that the subject was badly harmed. The subject must also show that the speaker did not exercise proper care.

While it is exceedingly difficult to prove “actual malice,” it is not exactly easy to establish negligence either. Suppose that a reporter learns, from an apparently credible source, that a lawyer or a banker has engaged in some corrupt conduct, or that a high school teacher was sexually involved with a student. Suppose that the allegation is false. Perhaps the reporter can be deemed negligent for failing to ensure that his source was right or for failing to ask alternative sources. But it will not be easy for Smith to demonstrate that the reporter was negligent as a matter of law.

To explain its conclusion in *Gertz*, the Court again invoked the idea of a chilling effect and said that free speech “requires that we protect some falsehood in order to protect speech that matters.”¹²⁷ In *New York Times Co. v. Sullivan*, after all, the Court had contended that a “defense for erroneous statements honestly made” is “essential.”¹²⁸ What was said there applies in *Gertz* as well: “[a] rule compelling the critic . . . to guarantee the truth of all his factual assertions — and to do so on pain of libel judgments virtually unlimited in amount — leads to . . . ‘self-censorship.’”¹²⁹ A constitutional ban on liability without fault, along with a requirement that negligence be shown, operates as a safeguard against journalistic or speaker self-silencing. In short, the Court continued the enterprise, started in *New York Times Co. v. Sullivan*, of attempting to regulate the extent of the “chill” on free speech.

In *Alvarez*, the plurality also emphasized that false statements of fact are sometimes protected in order to prevent the chilling effect: “[S]ome false statements are inevitable if there is to be an open and vigorous expression of views in public and private conversation, expression the First Amendment seeks to guarantee.”¹³⁰ In some sense, that is true. But the question remains: how many false statements are optimal?

B. Foundations

Constitutional limits on libel law can be assessed from many different directions. The first question, of course, is the choice of method. Notwithstanding a brief historical discussion, *New York Times Co. v. Sullivan* was not very grounded in the original public meaning of the

126. *Id.* at 350.

127. *Id.* at 341.

128. *Sullivan*, 376 U.S. at 278.

129. *Id.* at 279.

130. *United States v. Alvarez*, 567 U.S. 709, 718 (2012) (plurality opinion).

First Amendment, and if we are originalists, the decision is exceedingly difficult to defend.¹³¹ For the sake of simplicity, and without taking a stand on contested questions about method in constitutional law, let us explore the issue in the broadly pragmatic terms generally used in the Court's opinions.

A central point in *New York Times Co. v. Sullivan*, and a continuing concern for the modern era, involves two risks: political bias and official power.¹³² Either of these is dangerous in itself; the two can be a potent combination. Political bias can be found if a judge, jury, or even a well-funded plaintiff is acting in order to deter and punish speech of a particular kind.¹³³ We can easily imagine juries that would be sympathetic to one or another side in political debates. It is important to emphasize that the dangers posed by well-funded or ideological plaintiffs, whether they win or not, can essentially destroy the economic security of individuals and institutions.¹³⁴ Official power might be exercised by prosecutors — for example, the Department of Justice — not in a neutral effort to protect people's reputations but in an effort to respond to political interests or the will of a particular person or institution (say, the President).¹³⁵ Strong constitutional safeguards against the use of libel law might be seen, in the 1960s or now, as a kind of prophylactic against these risks.

Even so, we need to make some distinctions. Some false statements involve public officials.¹³⁶ Others involve celebrities — actors or dancers or singers — whose connection to the domain of self-government is usually obscure.¹³⁷ Still others involve not public officials but public

131. See *McKee v. Cosby*, 139 S. Ct. 675, 676 (2019) (Thomas, J., concurring in the denial of certiorari).

132. See *Sullivan*, 376 U.S. at 269, 272–73.

133. See Cassandra Burke Robertson, *Judicial Impartiality in a Partisan Era*, 70 FLA. L. REV. 739, 772–74 (2018); cf. Adam Liptak, *For Justices, Free Speech Often Means "Speech I Agree With,"* N.Y. TIMES (May 5, 2014), <https://nyti.ms/Rh3pKI> [<https://perma.cc/EAJ3-N5SG>].

134. See RYAN HOLIDAY, *CONSPIRACY: PETER THIEL, HULK HOGAN, GAWKER, AND THE ANATOMY OF INTRIGUE* 292–93 (2018).

135. See, e.g., Mark Mazzetti & Katie Benner, *Trump Pressed Australian Leader to Help Barr Investigate Mueller Inquiry's Origins*, N.Y. TIMES (Feb. 13, 2020), <https://nyti.ms/2oJL1jT> [<https://perma.cc/Z8CF-Y46K>] (originally published Sept. 30, 2019); Matt Zapotosky, *Federal Prosecutors Explore Years-Old Media Disclosure, Raising Fears Trump Is Using Justice Dept. for Political Gain*, WASH. POST (Jan. 16, 2020, 9:12 PM), https://www.washingtonpost.com/national-security/federal-prosecutors-explore-years-old-media-disclosure-raising-fears-trump-is-using-justice-dept-for-political-gain/2020/01/16/6bc56b68-38c2-11ea-bb7b-265f4554af6d_story.html [<https://perma.cc/35YE-KUUV>].

136. See, e.g., *Sullivan*, 376 U.S. at 256.

137. See, e.g., Laura M. Holson, *Why Is Sean Penn Suing Lee Daniels?*, N.Y. TIMES (Mar. 18, 2016), <https://nyti.ms/1VkcSc9I> [<https://perma.cc/3HZH-YZ9H>].

issues¹³⁸ — as, for example, if an ordinary person were accused of attempting to bribe an important executive at a local bank. Still others involve victims of sexual harassment.¹³⁹ For those who fall within each of these categories, the law is generally clear. Celebrities are treated the same as public officials. Public issues are not given any kind of special status; the question turns on the status of the person who is bringing suit. Public figures cannot recover for libel unless they can show actual malice.¹⁴⁰ Ordinary people must show negligence.¹⁴¹

It is not at all clear these rules strike the right balance. Consider those involved in public life: because actual malice is so difficult to establish, good people are subject to real damage, and those who do the damage cannot be held accountable in any way. As we have seen, the problem is not restricted to those who are damaged; it extends to self-government itself, which suffers if citizens cannot make fair evaluations.¹⁴² Consider entertainers: those who have decided to act or to sing are at increased risk of public ridicule or even cruelty, even if they have absolutely no role in politics. Consider ordinary people: it is not easy to demonstrate negligence, and if people spread a false and damaging rumor, it will be difficult to hold them accountable. The problem is especially serious in light of what is said on social media. The question of compensation is less important than the question of deterrence. With the law as it now stands, most false statements simply cannot be deterred.

It is far from clear that this is ideal, or even acceptable, from the standpoint of a well-functioning system of either freedom of expression or individual autonomy, which includes one's ability to control her reputation. Is it so clear that if an individual is acting negligently, she should be able to spread falsehoods about candidates for public office? On what assumptions? Do we really want to allow people to be able to spread negligent falsehoods about movie and television stars? True, famous people have a distinctive ability to reach large audiences and thus to correct errors, but among many viewers and readers, the truth will not prevail. Is it so important to provide breathing space for damaging falsehoods about actors, musicians, and other entertainers? In any case, is it clear that ordinary people should not be able to sue when they have been harmed by falsehoods? Any marketplace requires standards and

138. See, e.g., Jamie Tarabay, *Billionaire Wins Defamation Case Against Australian Media Group*, N.Y. TIMES (Feb. 22, 2019), <https://nyti.ms/2Nkrz5n> [<https://perma.cc/DN6P-YMPD>].

139. See, e.g., Julia Jacobs, *#MeToo Cases' New Legal Battleground: Defamation Lawsuits*, N.Y. TIMES (Jan. 12, 2020), <https://nyti.ms/2QKE6lQ> [<https://perma.cc/AM6H-W96R>].

140. *Sullivan*, 376 U.S. at 283.

141. See *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974).

142. See *supra* Sections II.C–D.

ground rules; no market can operate as a free-for-all. It is not at all obvious that the current regulatory system for free speech — the current setting of “chill” — is the one that we would or should choose for the Internet era.

It may or may not be too late to suggest a fundamental rethinking of basic principles. But it is hardly too late to adapt those principles to the modern era. It is worth noting that *New York Times Co. v. Sullivan* was decided in 1964, which might as well be a century ago, or maybe a millennium, in light of the massive technological advances that followed it. Spreading libelous or otherwise damaging statements is far easier today than ever before — but so is correcting them.¹⁴³ Because of technological changes, it would be a miracle if the ruling captured the best approach to accommodate today’s relevant values. Part of the force of the argument about chilling effects points to the expense of lawsuits, including high damage awards.¹⁴⁴ If the law could find ways to protect people against falsehoods without producing the excessive deterrence that comes from costly lawsuits, we might be able to accommodate the conflicting interests. Consider, then, three modest ideas meant to do exactly that.

First, damage caps and schedules could do a great deal to promote free speech values while also ensuring a measure of deterrence. Suppose, for example, that libel awards were usually bounded at a specified level (at the extreme, \$1). A very low limit would, of course, reduce the deterrent effect. But speakers have reputations to protect as well. If they are subject to liability, and if it is determined that they did not tell the truth, their reputations will suffer. From the standpoint of the system of freedom of expression, speakers’ concerns for their reputations are not exactly disasters; from the standpoint of ensuring against harms to individuals, they are an extremely good thing. A cap on damages, alongside liability to establish what is actually true, could work to leverage the speakers’ concerns for their reputations to good effect.

Second, there should be a general right to demand correction or retraction after a clear demonstration that a statement is both false and damaging — in other words, libelous under traditional standards. If a newspaper, broadcaster, or social media platform refuses to provide a prominent correction or retraction after a reasonable period of time, it might be liable for at least modest or nominal damages.

Third, on the Internet in particular, people might have a right to “notice and takedown.” Under this approach, modeled on the copyright

143. See Megan Garber, *How to Catch a Liar on the Internet*, THE ATLANTIC (Sept. 2013), <https://www.theatlantic.com/magazine/archive/2013/09/the-way-we-lie-now/309431> [https://perma.cc/93BV-26NM].

144. See David J. Acheson & Ansgar Wohlschlegel, *The Economics of Weaponized Defamation Lawsuits*, 47 SW. L. REV. 335, 355–56 (2018).

provisions of the Digital Millennium Copyright Act of 1998,¹⁴⁵ website operators, and perhaps social media providers, would be obliged to take down statements that are shown to be libelous under traditional common law standards upon proper notice. Such an obligation would require an amendment of § 230,¹⁴⁶ which immunizes social media platforms from most forms of liability.¹⁴⁷ Many courts have upheld immunity for online platforms against claims that platforms have unreasonably delayed removing defamatory or libelous posts.¹⁴⁸

It is true that consideration of any reform of existing practice would create extended discussion and careful consideration of counterarguments. In some forms, this approach would be burdensome; social media platforms are not courts, and they are hardly in the best position to judge what is libelous. It does not seem reasonable to ask, for example, Twitter or Facebook to engage in anything like adjudication. The simplest approach would be to say that if a statement has been held to be libelous by a competent tribunal, social media providers must take it down and will be held liable if they do not. The problem with this approach is that it would do too little; most victims of libel do not sue at all. The question is whether in very clear cases, where there is no reasonable doubt, social media providers should also have an obligation to take material down. The answer is almost certainly yes.

It is also true that because of the nature of the Internet, notice and takedown cannot provide a complete solution. Once material is posted, it might effectively be there forever.¹⁴⁹ But if it is taken down, it will not be in quite so many places, and at least the victim of the libel will be able to say that it was taken down.

Before we embrace any of these proposals, we would, of course, be required to undertake some sustained analysis of the likely consequences. I mention them not to offer a final verdict, or to endorse any of them in particular, but to sketch some possible approaches that might protect the legitimate rights of speakers while offering safeguards not only to those whose reputations might be damaged by falsehoods, but also to the many others who are harmed when they are misinformed about people, places, and things.

145. Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended in scattered sections of 17 and 28 U.S.C.).

146. 47 U.S.C. § 230 (2012).

147. For an interesting perspective, see Adi Robertson, *Why the Internet's Most Important Law Exists and How People Are Still Getting It Wrong*, THE VERGE (June 21, 2019, 1:02 PM), <https://www.theverge.com/2019/6/21/18700605/section-230-internet-law-twenty-six-words-that-created-the-internet-jeff-kosseff-interview> [<https://perma.cc/WT9S-6Q3C>].

148. See, e.g., *Batzel v. Smith*, 333 F.3d 1018, 1034 (9th Cir. 2003); *Zeran v. Am. Online Inc.*, 129 F.3d 327, 331 (4th Cir. 1997); *Blumenthal v. Drudge*, 992 F. Supp. 44, 49–53 (D.D.C. 1998); *Barrett v. Rosenthal*, 146 P.3d 510, 513 (Cal. 2006).

149. See Patricia Sanchez Abril & Jacqueline D. Lipton, *The Right to Be Forgotten: Who Decides What the World Forgets?*, 103 KY. L.J. 363, 385 (2014).

What about newspapers, magazines, and social media providers? None of them are subject to the First Amendment. There is a good argument that upon notice, and a sufficient factual demonstration, all of them should take down statements that would be libelous under traditional standards, even if no actual malice is involved. Why would that not be the appropriate (voluntary) practice, supposing that a sufficient factual demonstration has been made? To its credit, Facebook's Community Standards are directed against "bullying and harassment."¹⁵⁰ Why shouldn't they also be directed against libel?

C. Well Beyond Libel

Should lawmakers and regulators, concerned about truth, go beyond libel? Or does libel have some unique status, such that any efforts to focus on falsehoods and lies, as such, would raise insuperable constitutional objections? It is easy to answer "no" and "yes," respectively, to these questions. But there is good reason to hesitate before doing that. The easy answers are too easy.

The issue extends quite broadly. In the modern era, a pervasive concern is the dissemination of falsehoods about actual or potential public officials.¹⁵¹ The falsehoods might be innocent, negligent, or intentional. They might not expose people to ridicule or contempt, or otherwise count as libel, but they might be false, embarrassing, disruptive, or damaging even so. Importantly, they might be *positive*, such as, for example, a statement that a particular candidate served with great distinction in the military, competed in the Olympics, performed heroic actions, or invented some technology. We could easily imagine a law that would target such falsehoods, called the Dirty Tricks Act. It might say, for example, that speakers must pay a fine for knowingly spreading lies about candidates for public office; it might go further and target

150. See *Bullying and Harassment*, FACEBOOK: COMMUNITY STANDARDS, <https://www.facebook.com/communitystandards/bullying> [<https://perma.cc/DD63-GZQY>].

151. See generally Gerald G. Ashdown, *Distorting Democracy: Campaign Lies in the 21st Century*, 20 WM. & MARY BILL RTS. J. 1085 (2012) (exploring difficulties with and different methods of approaching the regulation of false speech in political campaigns); William P. Marshall, *False Campaign Speech and the First Amendment*, 153 U. PA. L. REV. 285 (2004) (discussing whether false political campaign speech can and should be regulated).

negligence as well.¹⁵² Analogues can be found in existing law, and they have sometimes been upheld.¹⁵³

If libel itself is involved, of course, *Sullivan* and its progeny set out the governing standards. But suppose that no libel is involved, perhaps because the technical requirements are not quite met, or perhaps because the falsehood casts a person or an institution in an appealing light. In addition to the examples above, it might be said that a presidential candidate opposed the Affordable Care Act, when she strongly supported it; that she wanted to repeal the Second Amendment, when she did not; that she spent six months studying Karl Marx in Moscow, which she never visited; that she wanted to ban hunting, which she did not; that she had a torrid romance with a famous movie star, when she did nothing of the kind; that she never went to law school, when she graduated with honors. Let us simply stipulate that none of these statements are libelous. Cases of this kind are at the intersection of *Sullivan* and *Alvarez*, and it is not clear how they should be handled.¹⁵⁴

Under *Sullivan*, we might think that the right question is whether there is actual malice.¹⁵⁵ Though *Sullivan* applied the actual malice standard for libel, it may be extrapolated to other false statements: if the speaker knew that they were false or acted with reckless indiffer-

152. I am bracketing here the possibility that some imaginable statutes regulating falsehoods could turn out to be a form of content discrimination or viewpoint discrimination. For example, a ban on falsehoods about Republican candidates, but not Democratic candidates, could easily be seen as impermissible content discrimination. A ban on falsehoods about the President of the United States might well fall in the same category. There are some tricky issues here. A ban on threats against the President is generally taken to be constitutional, evidently on the ground that such threats have a content-neutral justification; threats against the Commander in Chief are uniquely harmful. Can the same thing be said about falsehoods? In my view, that conclusion would be hard to justify. A ban on falsehoods involving a particular official, even the President, would be best understood as an effort to protect her from scrutiny.

153. See *Pestak v. Ohio Elections Comm'n*, 926 F.2d 573, 577 (6th Cir. 1991) (upholding an Ohio statute banning false statements in political campaigns). *But see* *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 476 (6th Cir. 2016) (striking down an Ohio election-lies law as a content-based restriction of “core political speech” that lacked sufficient tailoring); 281 Care Comm. v. Arneson, 766 F.3d 774, 785 (8th Cir. 2014) (“[N]o amount of narrow tailoring succeeds because [Minnesota’s political false-statements law] is not necessary, is simultaneously overbroad and underinclusive, and is not the least restrictive means of achieving any stated goal.”); *Tomei v. Finley*, 512 F. Supp. 695, 698 (N.D. Ill. 1981) (holding that the use of “REP” in the election context was not protected by the First Amendment because it falsely suggested that the defendants were Republicans). For arguments that such restrictions should be invalidated, see Marshall, *supra* note 151, at 300–22; Geoffrey R. Stone, *The Rules of Evidence and the Rules of Public Debate*, U. CHI. LEGAL F. 127, 136–37 (1993); and James Weinstein, *Free Speech and Domain Allocation: A Suggested Framework for Analyzing the Constitutionality of Prohibitions of Lies*, 71 OKLA. L. REV. 167, 206–13 (2018).

154. Hannah Arendt’s words remain relevant: “While probably no former time tolerated so many diverse opinions on religious or philosophical matters, factual truth, if it happens to oppose a given group’s profit or pleasure, is greeted today with greater hostility than ever before.” See ARENDT, *supra* note 1, at 231.

155. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279–80 (1964).

ence to the question of truth or falsity, the Constitution permits the government to impose liability. If *Sullivan* can be extended, then some version of the Dirty Tricks Act stands on firm ground. On the other hand, *Alvarez* seems to make it clear that intentional falsehoods are generally protected unless they cause demonstrable harm that cannot be avoided through speech-protective alternatives. In addition, the plurality in *Alvarez* gave great weight to time-honored categories of cases in which the government has forbidden false statements.¹⁵⁶ It loathed adding new categories, apparently on the theory that longevity created a kind of legitimacy.¹⁵⁷ Under this approach, because it goes beyond libel, the Dirty Tricks Act cannot claim longevity. But the basic point is that under *Alvarez*, intentional falsehoods enjoy a great deal of protection because they can be countered with more speech, rather than enforced silence.

For that reason, *Sullivan* and *Alvarez* are in evident, serious, and apparently unnoticed tension: under the reasoning of *Alvarez*, *Sullivan* would seem to have done far too little to protect speech; counterspeech might be used to respond to a libelous statement, even if the actual malice standard is met. Shouldn't the *Sullivan* Court have recognized that point, at least if *Alvarez* is correct? If someone says that a presidential candidate committed a violent crime and knows that the statement is false, counterspeech — and not a damage action — might be the constitutionally required response. And indeed, that does seem to be the logic of *Alvarez*, which emphasizes that in the face of an intentional falsehood about receipt of the Congressional Medal of Honor, the constitutionally mandated remedy is counterspeech.¹⁵⁸ And yet *Alvarez* does not question *Sullivan*.

One way to square the two decisions is to return to the *Alvarez* plurality opinion and emphasize history, not abstract theory. The idea of libel is time honored, which means that the Court will allow libel actions, so long as they are suitably constrained.¹⁵⁹ History does not honor a general exception for false statements of fact, even if they count as lies.¹⁶⁰ Another way to square the two opinions is to insist that libelous statements are indeed difficult to meet with counterspeech, and that the harm is so likely, and often so serious, that a damage remedy is legitimate. With a false claim about having received the Congressional

156. *United States v. Alvarez*, 567 U.S. 709, 717–18 (2012) (plurality opinion).

157. In this respect, there is an unmistakable Burkean strand to the plurality opinion in *Alvarez*, which can be read to say: Time has sanctified certain content-based restrictions, but we will allow no others. *Id.* On Burkeanism in constitutional law, see generally CASS R. SUNSTEIN, *A CONSTITUTION OF MANY MINDS* 35–59 (2009). By contrast, Justice Breyer's concurrence is more pragmatic, focused on policy and principle, rather than (strictly) historical provenance.

158. *Alvarez*, 567 U.S. at 726 (plurality opinion).

159. *Id.* at 718–19; see also *id.* at 737 (Breyer, J., concurring).

160. *Id.* at 719 (plurality opinion).

Medal of Honor, we have a form of pathetically narcissist lying, but nothing as grave as libel.

Where does this leave the Dirty Tricks Act? If it goes beyond libel, there is a strong argument that it is constitutionally invalid under *Alvarez*, except insofar as it reaches statements that cause serious harm and that cannot be sufficiently reduced through counterspeech. The meaning of the exception, signaled by *Alvarez*, is less than entirely clear. But if someone intentionally spreads a falsehood, and if the falsehood has an adverse effect on the democratic process, is it really right to insist that the Constitution forbids the government from imposing some kind of sanction? Is that so clear?

To be sure, we have good reason to fear the potential bias of any enforcing entity, particularly if it is in the hands of a political actor (such as the President) with clear political incentives. For that reason, there is a strong argument in favor of ensuring that any Dirty Tricks Act is enforced by an independent agency, such as or akin to the Federal Election Commission. The proposals sketched above — including a modest monetary penalty and a right to some kind of retraction — might well be used to accommodate the relevant interests.

Of course, social media providers have much more room to maneuver. Unconstrained by the First Amendment, they can respond to dirty tricks as they see fit. In view of the evident risks posed by false statements about political candidates, Facebook, Twitter, and YouTube should be doing more than they are now doing to take down such statements, even if they are made in political advertisements.¹⁶¹ It is not nearly enough to say that counterspeech is the right remedy.¹⁶²

IV. DEEPPFAKES AND DOCTORED VIDEOS

Libel is an old issue, but something like libel, or perhaps a modern version of the thing itself, comes from deepfakes: products of techniques, based on artificial intelligence or machine learning, for creating

161. For various views, see Danielle Abril, *Google and Twitter Changed Their Rules on Political Ads. Why Won't Facebook?*, FORTUNE (Nov. 22, 2019, 7:56 PM), <https://fortune.com/2019/11/22/facebook-google-twitter-political-ads-policy> [<https://perma.cc/PS42-3V26>]; Shelly Banjo, *Facebook, Twitter and the Digital Disinformation Mess*, WASH. POST (Oct. 2, 2019, 3:25 AM), https://www.washingtonpost.com/business/facebook-twitter-and-the-digital-disinformation-mess/2019/10/01/53334c08-e4b4-11e9-b0a6-3d03721b85ef_story.html [<https://perma.cc/2C3Z-EQ4W>]; and Cecilia Kang & Mike Isaac, *Defiant Zuckerberg Says Facebook Won't Police Political Speech*, N.Y. TIMES (Oct. 21, 2019), <https://nyti.ms/2qn01Fa> [<https://perma.cc/9K3J-CJ8N>].

162. See, e.g., Robert D. Richards & Clay Calvert, *Counterspeech 2000: A New Look at the Old Remedy for "Bad" Speech*, 2000 B.Y.U. L. REV. 553, 554–55; Kathleen M. Sullivan, *Resurrecting Free Speech*, 63 FORDHAM L. REV. 971, 974 (1995).

apparently real images or videos, in which people might be shown doing or saying something that they never did or said.¹⁶³ If an image of a person can be found, the technology is now available to make it look as if that person embraces racism, despises his country, engages in shoplifting, or dances wildly to Rolling Stones songs — anything at all.¹⁶⁴ Deepfake pornography is now pervasive.¹⁶⁵ Deepfakes could easily be used to discredit candidates for public office.

Let us understand “doctored videos” as products of techniques by which a real video is altered to make it seem as if people are doing or saying something other than what they did, or differently from how they did it.¹⁶⁶ A doctored video might show people supporting a cause that they abhor, committing a crime, showing disloyalty to their country, acting inappropriately when they did nothing of the kind, or being inebriated or otherwise impaired.¹⁶⁷ In some cases, doctored videos are quite credible, in the sense that viewers do not believe that they have been doctored. They might turn out to be libelous.¹⁶⁸ I will be speaking here of deepfakes, but the analysis applies to doctored videos as well.

A. *The Horizon*

For orientation — and for a sense of what is coming — consider the following cases:

- (1) Jane Jones is a high school student. A deepfake shows her in a romantic situation with a teacher.
- (2) Philip Cross is a candidate for public office. A deepfake shows him endorsing Hitler and the Holocaust.

163. Marc Jonathan Blitz, *Lies, Line Drawing, and (Deep) Fake News*, 71 OKLA. L. REV. 59, 61 (2018); Chesney & Citron, *supra* note 15, at 1757–58.

164. See Chesney & Citron, *supra* note 15, at 1757; Egor Zakharov et al., *Few-Shot Adversarial Learning of Realistic Neural Talking Head Models*, ARXIV, Sept. 25, 2019, at 2, <https://arxiv.org/pdf/1905.08233.pdf> [<https://perma.cc/5GJY-ZZT5>].

165. Danielle Keats Citron, *Sexual Privacy*, 128 YALE L.J. 1870, 1921–22 (2019).

166. See, e.g., BBC News, *Fake Obama Created Using AI Video Tool — BBC News*, YOUTUBE (July 19, 2017), <https://www.youtube.com/watch?v=AmUC4m6w1wo> [<https://perma.cc/89DT-Y4C6>]; Fortune Magazine, *What Is a Deepfake? Video Examples with Nicolas Cage, Jennifer Lawrence, Obama Show Troubling Trend*, YOUTUBE (Mar. 11, 2019), <https://www.youtube.com/watch?v=-yQxsIWO2ic> [<https://perma.cc/WW9F-JZ9F>]; Bernhard Warner, *Deepfake Video of Mark Zuckerberg Goes Viral on Eve of House A.I. Hearing*, FORTUNE (June 12, 2019, 12:31 PM), <https://fortune.com/2019/06/12/deepfake-mark-zuckerberg> [<https://perma.cc/52PM-2DGE>].

167. Cf. Blitz, *supra* note 163, at 61–62.

168. Corp. Training Unlimited, Inc. v. Nat’l Broad. Co., 868 F. Supp. 501, 507 (E.D.N.Y. 1994) (explaining that when scrutinizing “audio and video portions of a television program,” courts should be cognizant of “the possibility that a transcript which appears relatively mild on its face may actually be . . . highly toxic,” and that “a clever amalgamation of half-truths and opinion-like statements, adorned with orchestrated images and dramatic audio accompaniment, can be devastating when packaged in the powerful television medium”).

- (3) John Simons is eighty years old. A deepfake shows him taking an “energy pill” and then competing fiercely and well in a pickup basketball game.
- (4) A deepfake shows the Beatles playing Taylor Swift songs.
- (5) A deepfake shows the Mona Lisa speaking like a teenager.
- (6) A deepfake shows a Labrador Retriever dancing like Michael Jackson.
- (7) A deepfake shows the Attorney General looking drunk and disoriented.

Alvarez might well be taken to suggest that deepfakes cannot be regulated simply because they are deepfakes. They are essentially falsehoods, and to that extent, they are protected unless they cause harm.¹⁶⁹ Do they cause any? For the subjects of deepfakes, harm takes the form of reputational injury. A deepfake can be favorable or innocuous, but it can also be akin to or a form of libel;¹⁷⁰ it can hold people up to ridicule or contempt in their communities.¹⁷¹ If it is libelous — if its propositional content amounts to a libel — it can be regulated under the standards of *Sullivan* and *Gertz*.¹⁷² Apart from injuring the person who is depicted, it can injure the community by, for example, discrediting a candidate for public office and thus distorting the democratic process.

But deepfakes, as such, need not be libelous. They might be positive; they might make people look impressive or wonderful. For example, they might show a member of Congress playing tennis or golf at a professional level. And if people do not believe that a deepfake is real, there should be no harm. If a rock group from the 1960s (two members of whom are dead) is seen playing songs by a contemporary musician, we are dealing with something like whimsy, humor, or satire — all of which should be protected by the First Amendment.¹⁷³ The risk of harm arises if and when people think that what they are seeing actually happened. If people are falsely shown in a romantic situation or endorsing a political position that they abhor, they may in fact be harmed. In this

169. In *Alvarez*, Justice Breyer noted that “few statutes, if any, simply prohibit without limitation the telling of a lie, even a lie about one particular matter. Instead, in virtually all these instances limitations of context, requirements of proof of injury, and the like, narrow the statute to a subset of lies where specific harm is more likely to occur.” *United States v. Alvarez*, 567 U.S. 709, 736 (2012) (Breyer, J., concurring). He continued by explaining that “the limitations help to make certain that the statute does not allow its threat of liability or criminal punishment to roam at large, discouraging or forbidding the telling of the lie in contexts where harm is unlikely or the need for the prohibition is small.” *Id.*

170. See Citron, *supra* note 165, at 1933 n.421.

171. Cf. *Corp. Training*, 868 F. Supp. at 505–06.

172. See *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974); *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279–80 (1964).

173. *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 54 (1988).

light, consider the following proposition: *The government can regulate or ban deepfakes, consistent with the First Amendment, if (1) it is not reasonably obvious or explicitly and prominently disclosed that they are deepfakes, and (2) they would create serious personal embarrassment or reputational harm.* Note that (2) is meant to include but go beyond the existing law of libel.¹⁷⁴

Here is a possible response inspired by the prevailing opinion in *Alvarez*¹⁷⁵: the best response to a deepfake is disclosure, not censorship. A social media provider might conclude that platform users should be notified that doctored videos are doctored videos; it might reject the view that doctored videos should be taken down.¹⁷⁶ In its view, users should be able to see the videos so long as they receive the information they need to put the videos into perspective. As the plurality put it in *Alvarez*: “The remedy for speech that is false is speech that is true. This is the ordinary course in a free society.”¹⁷⁷

As a matter of constitutional law, something similar might be said about what governments may do. Perhaps they should be required to choose the approach that is maximally protective of speech. In *Alvarez*, for example, the plurality explained: “The Government has not shown, and cannot show, why counterspeech would not suffice to achieve its interest. The facts of this case indicate that the dynamics of free speech, of counterspeech, of refutation, can overcome the lie.”¹⁷⁸ It added: “There is, however, at least one less speech-restrictive means by which the Government could likely protect the integrity of the military awards system. A Government-created database could list Congressional Medal of Honor winners. Were a database accessible through the Internet, it would be easy to verify and expose false claims.”¹⁷⁹

But it would be difficult to defend the view that this is a sufficient response to deepfakes. If someone were portrayed as doing something that she never did or as endorsing a position that she despised, an accessible database would be most unlikely to undo the damage. Under *Alvarez*, there should be no constitutional barrier to allowing controls on deepfakes, at least on a sufficient showing of harm — and the same analysis applies to doctored videos.¹⁸⁰ Those controls might take the

174. Even if deepfakes did not threaten to harm their subjects, they could nonetheless cause harm. Consider, for example, a deepfake that portrays a political candidate as doing something heroic when he did nothing of the kind. The discussion of a Dirty Tricks Act, *supra* Section III.C, bears on the question whether deepfakes that benefit political candidates might be regulable.

175. *United States v. Alvarez*, 567 U.S. 709, 727 (2012) (plurality opinion).

176. For Facebook’s current position, see *infra* Section IV.B; and *Bullying and Harassment*, *supra* note 150.

177. *Alvarez*, 567 U.S. at 727 (plurality opinion).

178. *Id.* at 726.

179. *Id.* at 729.

180. For a valuable discussion, see generally Blitz, *supra* note 163. Professor Blitz noted that “where false statements do not merely state false facts, but are also given in a form that

form of a regulatory approach, operating perhaps via an independent commission, or (more interestingly) a tort-like approach, operating through a civil cause of action, building on libel law, and creating a kind of property right in one's person.

One could respond that, in an important sense, deepfakes and doctored videos are nothing new. They are equivalent to false statements of fact. If a deepfake depicts the Attorney General looking drunk and disoriented, perhaps it is not so different, or any different, from this apparently credible statement: "I saw the Attorney General looking drunk and disoriented." The propositional content of deepfakes is not regulated or banned (unless there is an independent ground — for example, they might be defamatory or obscene). Why should deepfakes themselves be treated differently?

A possible answer is that deepfakes (and doctored videos) have a unique kind of authenticity; they are more credible than merely verbal representations. In a sense, they are self-authenticating. The human mind does not easily dismiss them, and if it does, there is some part of it that remains convinced. To borrow a distinction from behavioral science,¹⁸¹ System 1 and System 2 do not distinguish between deepfakes and reality, and once System 2 is properly informed, System 1 is likely to remain under the influence; it hears a kind of mental echo. For these reasons, it is plausible to say that deepfakes (and doctored videos) are properly the subject of regulatory attention even if statements that embody their propositional content are not. And as we have seen, some false statements of fact, adversely affecting the political process, may be regulable even if they are not defamatory.¹⁸²

It would also be possible to object to an admittedly fuzzy component of the test that I have proposed, which focuses on whether it is reasonably obvious that a deepfake is not real. What exactly does that mean? There is no crisp and clean answer to this question. As in other domains of the law, courts must do their best to decide how reasonable people would regard the deepfake in question. More specific standards would undoubtedly emerge over time. It is important to say that most cases are easy. Some deepfakes (and some doctored videos) are extremely difficult to tell from the real thing. Others are clearly a product of technology.

carries with it indicia for reliability (such as a falsified newspaper or video or audio tape), the government should have greater power to regulate than it typically has to regulate false words." *Id.* at 110. For a careful analysis of the constitutional issues, see Chesney & Citron, *supra* note 15, at 1790–93 (finding *Alvarez* a significant obstacle to regulation).

181. See KAHNEMAN, *supra* note 46, at 20–21.

182. See *supra* notes 60–62 and accompanying text.

B. Facebook's Incomplete Solution

These conclusions have strong implications for social media platforms such as Facebook, YouTube, and Twitter. If the government has power to regulate certain deepfakes and doctored videos, it follows that such platforms ought to exercise their broad authority to do exactly that. Restrictions of the kind outlined here, or some suitable variation, could protect, at once, democratic processes and individual lives. In 2019, Facebook responded to considerations of this general kind with an announcement that it is banning deepfakes.¹⁸³ That is an important step in the right direction, but is it enough?

As Facebook pointed out in its announcement, media can be manipulated for benign reasons like making video sharper and audio clearer.¹⁸⁴ Some forms of manipulation are clearly meant as jokes, satires, parodies, or political statements — as, for example, when a rock star or a politician is depicted as a giant or as having superpowers. That is not a basis for regulation. Under its policy, Facebook says that it will remove “misleading manipulative media” only if two conditions are met:

- (1) “It has been edited or synthesized — beyond adjustments for clarity or quality — in ways that aren’t apparent to an average person and would likely mislead someone into thinking that a subject of the video said words that they did not actually say.”¹⁸⁵
- (2) “It is the product of artificial intelligence or machine learning that merges, replaces or superimposes content onto a video, making it appear to be authentic.”¹⁸⁶

In a sense, the first condition is close to what I have suggested here. The average person must, by hypothesis, be reasonable. And if the two conditions are taken together, they suggest a broader prohibition than I have suggested. There is no requirement of serious personal embarrassment or harm (and perhaps there should be no such requirement insofar as we are dealing with a private actor, not a government). The two conditions are meant to be tailored to Facebook’s concern: use of new or emerging technologies to mislead the average person into thinking that someone said something that they never said.

183. Monika Bickert, *Enforcing Against Manipulated Media*, FACEBOOK (Jan. 6, 2020), <https://about.fb.com/news/2020/01/enforcing-against-manipulated-media> [<https://perma.cc/A4SV-878X>].

184. *Id.*

185. *Id.*

186. *Id.*

Facebook's announcement also makes it clear that even if a video is not removed under the new policy, other safeguards might be triggered.¹⁸⁷ If, for example, a video contains graphic violence or nudity, it will be taken down. And if it is determined to be false by independent third-party fact-checkers, those who see it or share it will see a warning informing them that it is false.¹⁸⁸ Its distribution will also be greatly reduced in Facebook's News Feed.¹⁸⁹

While the resulting approach should be seen as significant progress over what preceded it, two problems remain. The first is that, even if a deepfake is involved, the policy does not apply if the deepfake depicts deeds rather than words. Suppose that artificial intelligence is used to show a political candidate working with terrorists, beating up a small child, or using heroin. Nothing in the new policy would address those depictions. That is a serious gap. Deepfakes should be banned if they present actions that would otherwise fall within the prohibition, even if no one is shown to have said anything.

The second problem is that the prohibition is limited to products of artificial intelligence and machine learning. But why? Suppose that videos are altered in other ways — for example, by slowing them down so as to make someone appear drunk or drugged, as in the case of an infamous doctored video of Nancy Pelosi.¹⁹⁰ Or suppose that a series of videos, directed against a candidate for governor, are produced not with artificial intelligence or machine learning, but nonetheless in such a way as to run afoul of the first condition; that is, they have been edited or synthesized so as to make the average person think that the candidate said words that she did not actually say. What matters is not the particular technology used to deceive people, but whether unacceptable deception has occurred — at least insofar as we are dealing with private actors.¹⁹¹ For governments, some showing of harm should be required.

187. *Id.*; see also Guy Rosen et al., *Helping to Protect the 2020 US Elections*, FACEBOOK (Oct. 21, 2019), <https://about.fb.com/news/2019/10/update-on-election-integrity-efforts> [<https://perma.cc/Q6DV-YA4Q>] (announcing additional safeguards to reduce the impact of viral misinformation on U.S. elections).

188. Bickert, *supra* note 183.

189. *Id.*

190. See Cass R. Sunstein, *Act Now to Head Off Looming "Deepfakes" Disasters*, BLOOMBERG (May 29, 2019, 1:01 PM), <https://www.bloomberg.com/opinion/articles/2019-05-29/pelosi-video-russia-threat-and-deepfake-threat> [<https://perma.cc/89PJ-C8MX>].

191. The considerations raised here suggest grave doubts about Facebook's policy in favor of allowing falsehoods in campaign advertisements, though that issue deserves extended discussion. For a very brief treatment, see Cass R. Sunstein, *Facebook Can Fight Lies in Political Ads*, BLOOMBERG (Oct. 9, 2019, 8:00 AM), <https://www.bloomberg.com/opinion/articles/2019-10-09/facebook-can-fight-lies-in-political-ads> [<https://perma.cc/2VA7-QXUL>].

V. CONCLUSION

In politics, truth matters. In the end, truth might matter more than anything else.

But what, exactly, can governments do to restrict the dissemination of falsehoods in systems committed to freedom of speech? On numerous occasions, the Supreme Court has said that false statements are valueless and that the First Amendment does not protect them. At the same time, the Court has long held that the First Amendment imposes serious restrictions on libel actions and that many false statements are indeed protected by that amendment. In *Alvarez*, the Court made this point clear, suggesting that false statements are protected unless the government can produce a powerful justification for regulating or banning them. The Court has explained this conclusion largely by pointing to the risk that regulating falsehoods might deter truth.

We have seen an assortment of arguments for protecting false statements of fact. First, the government's own judgments about what is true and what is false might not be trustworthy. Second, regulation of false statements might and probably will chill truthful statements. Third, people can learn from false statements; engagement with false statements can deepen their understanding. Fourth, it is important for people to know what other people think, even if what they think is not true. Fifth, banning false statements can simply drive them underground and increase their power and allure; counterspeech can be far more effective than prohibition.

The problem is that, taken individually or as a whole, these arguments do not support the broad conclusion that a system of free expression must give strong protection to false statements. Some falsehoods and some lies make general propositions of this kind seem inadequate or even silly. Some involve time-honored prohibitions: consider perjury, fraud, and false commercial advertising (“this product prevents cancer”). These generally involve serious dangers, but not all of them always do. For example, a pathetic and doomed effort to defraud people counts as fraud, even if it is pathetic and doomed. Nonetheless, a legal system committed to freedom of expression could do much worse than to adopt a presumption to the effect that false statements of fact are protected by the First Amendment unless they threaten to produce serious harm. But the presumption is no more than that, and in *Alvarez*, the plurality was myopic in focusing on time-honored categories of cases in which false statements of fact may be regulated.¹⁹² There are other cases — involving falsity but not libel — in which the presumption should also be overcome.

192. See *United States v. Alvarez*, 567 U.S. 709, 718–20 (2012) (plurality opinion).

These points leave a degree of vagueness. But they offer a distinctive perspective on both old and new problems. They suggest that current constitutional law does not strike the right balance, and that public officials, actors, musicians, and athletes should be able to do far more than they are now permitted to do to control libel and other falsehoods. The same is true for ordinary citizens subject to damaging falsehoods. In addition, public officials have considerable power to regulate deep-fakes and doctored videos.

Private institutions, including social media providers, should be acting far more aggressively to control libel and other falsehoods. They should be doing far more than they are now doing to prevent the spread of doctored videos and reduce the coming spread of deepfakes. These are specific conclusions, but they bear on some of the largest and most general questions in all of American public law. Hannah Arendt put it this way: “What is at stake here is this common and factual reality itself, and this is indeed a political problem of the first order.”¹⁹³

193. See ARENDT, *supra* note 1, at 232.