

Oklahoma Law Review

Volume 71 | Number 1

Symposium: Falsehoods, Fake News, and the First Amendment

2018

Lies, Line Drawing, and (Deep) Fake News

Marc Jonathan Blitz

Follow this and additional works at: <https://digitalcommons.law.ou.edu/olr>



Part of the [Constitutional Law Commons](#), [Election Law Commons](#), and the [First Amendment Commons](#)

Recommended Citation

Marc Jonathan Blitz, *Lies, Line Drawing, and (Deep) Fake News*, 72 OKLA. L. REV. 59 (2018).

This Panel 1: Falsehoods and the First Amendment is brought to you for free and open access by University of Oklahoma College of Law Digital Commons. It has been accepted for inclusion in Oklahoma Law Review by an authorized editor of University of Oklahoma College of Law Digital Commons. For more information, please contact darinfox@ou.edu.

LIES, LINE DRAWING, AND (DEEP) FAKE NEWS

MARC JONATHAN BLITZ*

You've been bitten by someone's false beliefs. Thought contagion.

—Muse, *Thought Contagion* (Warner Music single 2018)

Don't trust a stranger's words . . .

And they hypnotized all of us. Nobody knew who they could really trust.

And they stretched us out until we split. Divided us up until there just was nothing left.

—Chad VanGalen, *Host Body*,
on LIGHT INFORMATION (Sub Pop Records 2017)

I. Introduction

Before the computer-generated virtual reality of the 1999 movie *The Matrix*, there was *The Cosmic Puppets*—a 1957 book by Philip K. Dick in which a man, after many years living elsewhere, drives back to his hometown and discovers it is not merely the town that has changed, but its entire history.¹ The house he lived in and streets he walked upon growing up have not merely been torn down and paved over—they've now never existed.² The protagonist discovers that he remains a part of the town's history: an old newspaper he obtains from the archives refers to his nine-year-old self in an obituary describing his death from scarlet fever.³ Instead of leaving the town at age nine, he has, in the town's new strange and unfamiliar reality, been dead since that time.⁴ He finally meets one other person in the town who remembers it the way he does and has also been trying to understand why the town's appearance and history has been covered with a false surface.⁵ With extraordinary mental effort, some homemade electrical technology, and stubborn refusal to accept the views of their fellow town members, they are

* Alan Joseph Bennett Professor of Law, Oklahoma City University; J.D., University of Chicago (2001); Ph.D., (Political Science) University of Chicago (2001), B.A., Harvard University (1989).

1. PHILIP K. DICK, *THE COSMIC PUPPETS* (First Mariner Books 2012) (1957).

2. *Id.* at 1–6.

3. *Id.* at 12.

4. *Id.*

5. *Id.* at 54–55.

able to slowly strip away the veneer of fake reality that has been laid over their world and see once again the real town hidden underneath.⁶

Dick's story is just one example of what is sometimes called "paranoid science fiction," a genre of science fiction where, as one encyclopedia of science fiction and fantasy describes it, "protagonists are plagued with vague intuitions of the stage-managed falsity of their perceptual experience or delusory nature of their very identities."⁷ One finds numerous other examples of paranoid science fiction in Dick's own writing, including other stories where one's environment turns out to be a carefully constructed simulation and stories (like the basis for the movie *Blade Runner*) where the people one knows might be androids virtually indistinguishable from real human beings.⁸ Other works of science fiction and fantasy also fit the genre: stories (and movies) such as *The Thing* and *Invasion of the Body Snatchers*, where people are replaced by alien look-a-likes;⁹ *Twilight Zone* episodes where a man finds his life is a screenplay in which he is a television actor;¹⁰ and *The Matrix* itself, where the day-to-day life the protagonists have grown up believing in is revealed to take place in a virtual reality simulation.¹¹ Indeed, some examples of such paranoid science fiction are much older. One arguably finds something like it in the philosopher Rene Descartes' 1641 thought experiment, wherein he imagines an omnipotent evil demon who can make him see, hear, and feel an external reality—"the sky, the air, the earth, colours, shapes, sounds and all external things"—of a kind that doesn't really exist.¹²

As the name indicates, paranoid science fiction is only fiction. In all of the stories just referenced, the "stage-managed falsity" of "perceptual

6. *Id.* at 131–36.

7. Gary Westphal, Paranoia, in 2 THE GREENWOOD ENCYCLOPEDIA OF SCIENCE FICTION AND FANTASY 585 (Greenwood Press 2005).

8. See *Blade Runner*, INTERNET MOVIE DATABASE, <https://www.imdb.com/title/tt0083658/plotsummary> (last visited July 5, 2018).

9. See *The Thing*, INTERNET MOVIE DATABASE, <https://www.imdb.com/title/tt0084787/plotsummary> (last visited July 5, 2018); *Invasion of the Body Snatchers*, INTERNET MOVIE DATABASE, <https://www.imdb.com/title/tt0049366/plotsummary> (last visited July 5, 2018).

10. See *The Twilight Zone: A World of Difference*, INTERNET MOVIE DATABASE https://www.imdb.com/title/tt0734552/?ref_=tt_urv (last visited July 5, 2018).

11. See *The Matrix*, INTERNET MOVIE DATABASE <https://www.imdb.com/title/tt0133093/plotsummary> (last visited July 5, 2018).

12. RENE DESCARTES, MEDITATIONS ON FIRST PHILOSOPHY 15 (ed., trans. John Cottingham Cambridge Univ. Press 1996).

experience” is the product of a mythical, supernatural entity or forms of alien or artificial intelligence imagined by the author. In other tales less removed from historical experience, the falseness of our historical world is the product not of an omnipotent demon or computer matrix, but rather of a control-hungry authoritarian system, like the Ministry of Truth used in Orwell’s *1984* to create “day-to-day falsification of the past” to ensure that people’s beliefs about reality serve the government’s interest.¹³

But soon, at least a modest power of reality distortion may belong not solely to evil demons, digital overlords, or ministries of propaganda, but rather to any individual with widely available computer software: new technologies for altering video and audio material will likely allow individuals to create convincingly realistic footage of events that never occurred. Indeed, “forensic specialists predict that computers will be able to generate convincing, fabricated audio and video recordings at a rapid pace in the next few years,” which, one commentator notes, “will take fake news to a whole new level.”¹⁴ Another stresses that “[a]lready available tools for audio and video manipulation . . . have begun to look like a potential fake news Manhattan Project.”¹⁵ In a recent post on *Lawfare Blog*, Bobby Chesney and Danielle Citron describe such technologically generated illusions as “deep fakes”—a phrase that has been most often used to describe a genre of artificial-intelligence-generated pornography that makes celebrities appear to engage in sexual scenes they had nothing to do with.¹⁶ Chesney and Citron describe the generation of deep fakes more broadly as “digital manipulation of sound, images, or video to impersonate someone or make it appear that a person did something—and to do so in a manner that is increasingly realistic, to the point that the unaided observer cannot detect the fake.”¹⁷ They further note that deep fakes are not likely to remain the province of governments or extraordinarily powerful corporations, but will rather “diffuse rapidly and

13. GEORGE ORWELL, 1984, at 289 (Samaira Book Publishers 2017) (1949).

14. Hilke Schellmann, *The Dangerous New Technology That Will Make Us Question Our Basic Idea of Reality*, QUARTZ (Dec. 5, 2017), <https://qz.com/1145657/the-dangerous-new-technology-that-will-make-us-question-our-basic-idea-of-reality/>.

15. Charlie Warzel, *He Predicted the 2016 Fake News Crisis. Now He’s Worried About an Information Apocalypse*, BUZZFEED NEWS (Feb. 11, 2018), https://www.buzzfeed.com/charliewarzel/the-terrifying-future-of-fake-news/fakenews?utm_term=.ofPnjaMqa#.ehdR6MXYM.

16. Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Crisis for National Security, Democracy and Privacy?*, LAWFARE (Feb. 21, 2018), <https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy>.

17. *Id.*

globally” and “end up in the hands of a vast range of actors willing to use deep fakes in harmful ways.”¹⁸

Given the havoc that would surely ensue if individuals engaged in a kind of informational Hobbesian war of all against all—with people constantly falsifying one another’s sense of what is real—it makes sense to ask whether government has the power to take measures against this sort of action. It is, after all, government that social contract theorists like Hobbes and Locke looked to help individuals escape from the insecurity of a “state of nature” where individuals were constantly vulnerable to physical attack from their neighbors. Might government likewise protect people against informational attack?

One potential barrier for government regulation of information attack, of course, is the First Amendment. Deep fakes, for example, are generally video or audio creations, and such creations have typically been considered a form of expression. So too is a painting created in the style of Rembrandt, perhaps so faithfully that even viewers educated in art history will mistake it for a Rembrandt painting. Do these types of expression receive First Amendment protection even when they are intended to deceive, and succeed in their deception? If they are protected by free speech guarantees, then how is government to protect people from the kind of havoc which writers predict may arise from deep fakes? If, on the other hand, they are *not* protected, then does this also open the door for government to restrict *other* kinds of expression—beyond doctored video- or audiotapes—such as false evidence of events that takes the form of words rather than video footage? If the First Amendment does not present a barrier to regulating deep fakes, why then would it present a barrier to government restriction of “fake news” that takes the form of a false Tweet or Facebook post? Or a factual claim about public affairs—falsely made by a skilled and persuasive speaker—to an audience ready to believe it?

My purpose in this Essay is to use these questions as a starting point for taking another look at a question that has been thoughtfully explored in recent free speech jurisprudence and scholarship. Namely, when does factually false expression qualify for First Amendment protection, and when it does not? In the past two years, this question has been raised about fake news, a vague phrase that refers to efforts to spread false information about public affairs or publicly known individuals, principally over the Internet, by means as simple as making the false claim in a Tweet, or as sophisticated as creating a fake

18. *Id.*

but seemingly real imitation of a newspaper or other established media (and perhaps fake photographic evidence to accompany it). This Essay does not offer a definitive answer to the question of when the falsity in fake news might fail to receive the First Amendment protection that normally applies to expression of all kinds (including false and inaccurate information). Nor will it offer a definitive answer to the question of whether and when deep fakes might receive First Amendment protection when they deceive with video- or audio-alteration rather than verbal expression.

Instead, this Essay aims to explore and reflect on the merits of some of the key legal frameworks that judges and scholars have offered for addressing such questions and to outline an additional proposal that is especially relevant to the problem of fabricated video and audio. It begins in Part II by introducing the key framework that the Supreme Court has offered for thinking about such issues, particularly the discussion among the Justices in *United States v. Alvarez*¹⁹ about when lying should count as protected free speech. As explained below, the key lesson of *Alvarez* for the regulation of fake news or politically relevant deep fakes is that, so long as their content is a matter of public concern—as most news of any kind *is*—then it can be proscribed by government only if it falls into a historically unprotected category of expression such as fraud or defamation.

In Parts III and IV, I will explore some possible variations of this model—each of which modestly expands the kind of false speech that government may regulate. Part III examines three possible bases, discussed in other scholars' work, for how one might do so. Lies about politics, philosophy, history, and other matters of public concern might be subject to restriction not only when they cause “legally cognizable harm,” but also (i) when they are exceptions to the general assumption that false claims about public events or political subjects are, by their very nature, matters of public concern, or (ii) when they do one or both of the following: (a) knowingly manipulate a listener with false information to serve the speaker's ends rather than his own, or (b) provide the listener with unfounded expert knowledge or other claims that an asymmetry in information makes the listener ill-equipped to critically assess. Neither of these, I will argue, seems all that likely to make a significant difference in how the *Alvarez* framework gets applied.

In Part IV, however, I will identify something that may do so, particularly given recent technological developments. Fake news may lose protection, I suggest, when it is not only a falsity, but a forgery as well. In other words, a

19. 567 U.S. 709 (2012).

distinctive type of harm may arise when the falsehood is not merely in the content of the speech that is intended to deceive, but is also in its purported source or vehicle. This is not limited to situations where video or audio make nonexistent events seem real. Imagine, for example, that a news story reaches its audience and, in doing so, provides the audience with fake content while masquerading as another publication. The story could, for example, come in the guise of an article by the *Washington Post* or the *Miami Herald*, though neither publication played any role in it. The falsity in such expression is not only in a statement contained within the writing, but in way the writing is disguised as an authoritative journalistic source.

II. Alvarez, “Low Value” Speech, and Lying

Before delving into the legal questions posed by so-called deep fakes—the “potential fake news Manhattan Project” discussed above—it is useful to briefly explore some more familiar examples of fake news that are a part of the present, rather than a dystopian near-future. On September 30, 2016—just over a month before the presidential election—the website for the Christian Times Newspaper exposed a carefully hidden plan that that was supposed to have remained a secret from voters: tens of thousands of fake ballots sat in an Ohio warehouse, ready to be used on election day to swing the election results in Ohio, overriding the voters’ wishes and assuring the election outcome desired by the secretive operation that created the fake ballots.²⁰ Accompanying the story was a photograph of the electrician who had stumbled upon the scheme, standing behind the many ballot-filled boxes.²¹ Except, in the end, it was not these hidden ballots that were fraudulent, but the news story that “revealed” them. As the *New York Times* reported four months later, the Christian Times Newspaper was not a real newspaper.²² It was a website created by a recent college graduate trying to make some money off of fake news by putting concrete evidence behind the vague fear, expressed by many Trump supporters at the time, that Hillary Clinton’s campaign was preparing to steal the election.²³ The photograph in the story was real, but it had nothing to do with ballots in Ohio: “It was a photo from

20. Scott Shane, *From Headline to Photograph: A Fake News Masterpiece*, N.Y. TIMES (Jan. 18, 2017), <https://www.nytimes.com/2017/01/18/us/fake-news-hillary-clinton-cameron-harris.html>.

21. *Id.*

22. *Id.*

23. *Id.*

The Birmingham Mail, showing a British election 3,700 miles from Columbus.”²⁴ Nothing in the photo clearly revealed that or contradicted the caption, which set it in the United States and characterized it as depicting the Trump-Clinton contest.²⁵

This was only one of a multitude of fake news stories that ran prior to the 2016 election. Enterprising young workers in Veles, Macedonia produced hundreds of articles attempting to draw interest from politically obsessed American internet users—along with the advertising revenue their “clicks” would bring.²⁶ One of these stories, for example, revealed that Hillary Clinton had said in 2013 that people “Like Donald Trump” should “Run For Office” because “They’re Honest And Can’t Be Bought.”²⁷ As the indictment obtained by Special Counsel Robert Mueller recently recounted, the Russian Internet Research Agency likewise contributed to the confusion in election reporting by borrowing real—and inventing fake—American identities, such as a Twitter account purportedly run by the Tennessee GOP, to spread misinformation.²⁸

Nor is fake news the only fake information reporters have raised concerns about. In 2013, a journalist for *Science* (trained as a biologist) revealed how easy it had been for him to get fake science accepted for publication in over a hundred open-access science journals, some run by well-known and established publishing companies.²⁹ He explained in his *Science* exposé that he had written an analysis demonstrating “the anticancer properties of a chemical . . . extracted from a lichen.”³⁰ He then submitted a variation of this made-up study to over 300 open-access journals (under a false name and false institutional affiliation, with a made-up university) to see how many would be duped by it.³¹ “More than half of the journals accepted the paper,”

24. *Id.*

25. *Id.*

26. Craig Silverman & Lawrence Alexander, *How Teens in the Balkans Are Duping Trump Supporters with Fake News*, BUZZFEED NEWS (Nov. 3, 2016), https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.hq0x0L4lL#.ptd30r4Gr.

27. *Id.*

28. Indictment, United States v. Internet Research Agency, No. 1:18-cr-00032-DLF (D.D.C. Feb. 16, 2018), 2018 WL 914777.

29. John Bohannon, *Who’s Afraid of Peer Review?*, 342 SCIENCE 60, 60–64 (2013), <http://science.sciencemag.org/content/342/6154/60.full>.

30. *Id.*

31. *Id.*

he noted, “failing to notice its fatal flaws,” which included a graph that showed results flatly at odds with those claimed by the paper itself.³²

Does First Amendment doctrine allow, and *should* it allow, government to punish such fake news, junk science, and other false statements of fact? Where such falsity is used not to test journals’ quality control, as the *Science* journalist did, but in order to deceive the public into accepting its truth, should legislatures be able to subject the speaker to civil or criminal liability? The most familiar answer in First Amendment caselaw is no. As explained below, all of the Justices in *Alvarez* found that allowing the government to punish lying on matters of public concern would require inviting it to monitor and screen out elements of public discourse in a way that is at odds with First Amendment principles. But it is useful to explore more carefully how most First Amendment analyses reach this conclusion.

Intuitively, the First Amendment status of fake news or fake science might well depend in part on what kind of fakery one believes is involved in fake news or fake science. If it is like the kind of forgery or faking one finds in the commercial sphere, then government probably has some room to regulate and to restrict it. In the commercial marketplace, government often stands ready to intervene in order to protect consumers from deception. For instance, it protects us against being sold forgeries or other fake goods. Prosecutors bring cases against those who peddle counterfeit pharmaceuticals,³³ travel scams,³⁴ fake concert tickets,³⁵ or other goods that aren’t what they purport to be. And government may generally intervene in this way even where the deception is carried out through, or consists of, expression. The sale of an artistic forgery can be a crime,³⁶ as can sale of a fake celebrity autograph.³⁷

32. *Id.*

33. See John Roth, *Counterfeit Drugs: Prosecuting the Profiteers, Protecting the Public Health*, FDA VOICE (Jul. 15, 2013), <https://blogs.fda.gov/fdavoices/index.php/2013/07/counterfeit-drugs-prosecuting-the-profiteers-protecting-the-public-health/>.

34. See *10 Accused of Scamming More than \$1 Million in New York City Travel Agency Scheme*, ABC NEWS (Dec. 12, 2014), [http://abc7ny.com/travel/10-accused-of-scamming-more-than-\\$1-million-in-nyc-travel-agency-scheme/434520/](http://abc7ny.com/travel/10-accused-of-scamming-more-than-$1-million-in-nyc-travel-agency-scheme/434520/).

35. Victor Fiorillo, *Feds Charge Philly Man in Fake Bruce Springsteen Tickets Scheme*, PHILADELPHIA (Dec. 19, 2016, 12:39 PM), <https://www.phillymag.com/news/2016/12/19/fake-bruce-springsteen-tickets/#lsoSymKGSHv6W0fX.99>.

36. See Christine E. Weller, *Lessons from Two Recent Art Forgery Cases*, 3 STETSON J. ADVOC. & L. 1 (2016) (discussing U.S. laws used to prosecute crimes related to art fraud).

37. John Meyers, *Gov. Brown Signs Law That Cracks Down on Fake Celebrity Autographs*, L.A. TIMES (Sept. 9, 2016), <http://www.latimes.com/politics/essential/la-pol->

Such governmental restriction of deception is not limited to the commercial context. It is often illegal to make false statements where government needs honest answers to questions or needs to assure there is no deception about who is authorized to take actions that only government officials may permissibly take.³⁸ State laws typically ban people from impersonating police officers or other government officials,³⁹ from email “spoofing”⁴⁰ or other use of fake online identities to cause harm to others,⁴¹ from creating or using fake IDs to enter airports or other secure areas,⁴² and from lying to building or restaurant inspectors.⁴³ By doing so, government enables us to move through public life without the kind of paralyzing paranoia that might be necessary where everyone and everything around us might be a potentially harmful impostor. Government is thus, in this sphere of our lives, very much expected to actively identify and, in a sense, quarantine and neutralize, falsehoods.

Matters are very different, however, in the “marketplace of ideas.” Here, individuals are largely on their own: government may not constitutionally exile certain ideas from the free trade in ideas as it can ban harmful goods or services from the realm of buying and selling. As the Supreme Court declared in *Gertz v. Robert Welch, Inc.*, “Under the First Amendment there is no such thing as a false idea. However pernicious an opinion may seem, we depend for its correction not on the

sac-essential-politics-updates-gov-brown-signs-law-that-cracks-down-1473451218-html story.html.

38. See Helen Klein Murillo, *The Law of Lying: Perjury, False Statements, and Obstruction*, LAWFARE (Mar. 22, 2017, 9:30 AM), <https://www.lawfareblog.com/law-lying-perjury-false-statements-and-obstruction>.

39. See *United States v. Chappell*, 691 F.3d 388 (4th Cir. 2012) (finding the First Amendment permitted “Virginia police impersonation statute, Virginia Code § 18.2–174, [which] prohibits individuals from falsely assuming or pretending to be a law enforcement officer”).

40. See L. Richard Fischer & Shannon K. Ryerson, *New York Arrests Spammer for E-Mail Spoofing*, PRIVACY & INFO. L. REP., May 2003, at 19.

41. See Victor Luckerson, *Can You Go to Jail for Impersonating Someone Online?*, TIME (Jan. 22, 2013), <http://business.time.com/2013/01/22/can-you-go-to-jail-for-impersonating-someone-online/>.

42. See, e.g., *Man Arrested with Fake ID at Sanford Airport, Authorities Say*, WESH2 (May 26, 2016), <http://www.wesh.com/article/man-arrested-with-fake-id-at-sanford-airport-authorities-say/4449756>.

43. See *Former Niagara Falls Man Sentenced for Making a False Statement*, U.S. DEP’T OF JUST. (Jan. 22, 2016), <https://www.justice.gov/usao-wdny/pr/former-niagara-falls-man-sentenced-making-false-statement>.

conscience of judges and juries but on the competition of other ideas.”⁴⁴ Thus, government cannot deal with a person’s claim of being “cheated” into adopting the wrong spiritual commitment or political ideology in the way it deals with a person’s claim that she was duped into buying a fake concert ticket, a forged piece of art, or a car that lacks the features its seller touted. She may deeply regret having lived for years with a religious commitment she now regards as erroneous. She may likewise deeply regret having furthered a political ideology and advocated for policies she now regards as damaging. And she may bitterly resent what she sees as the misleading words of a proselytizer or passionate advocate who helped bring her to the religion or political camp she now firmly rejects. But American democracy gives government neither the power to investigate and shutter the church or the political group she regrets joining, nor the power to prevent new members from entering. Government’s role here is to stand back and let individuals judge for themselves whether ideas are worthy of adherence. In the realm of ideas, wrote Justice Jackson in 1945, “every person must be his own watchman for truth, because the forefathers did not trust any government to separate the true from the false for us.”⁴⁵

Where then do we classify fake news or junk science, like the kind I described earlier, in this dichotomy? Are false articles or bogus science more like goods we purchase that aren’t what they purport to be, or are they more like spiritual or political ideas, where it is up to us, not the government, to decide which ideas we will treat as true? On the one hand, fake news or science articles aren’t generally products we rely upon to fulfill a specific, practical purpose. They are not like a set of tickets we purchase to get entry to concert or sporting event, which will entirely fail to serve their purpose if they are fake. They are rather more like the heap of other potentially mistaken or unwise claims—about politics, religion, and a host of other topics—that each of us has to sort through and evaluate in order to figure out what to believe about the world and how to build good lives in it. But on the other hand, false factual statements are unlike religious ideas and political opinions in at least one respect: they can be exposed as fake. In this respect, they are more like a forged work of art that purports to be a fifteenth century Da Vinci painting but is

44. 418 U.S. 323, 339–40 (1974).

45. *Thomas v. Collins*, 323 U.S. 516, 545 (1945) (Jackson, J., concurring) (citing *W. Va. State Bd. Of Educ. v. Barnette*, 319 U.S. 624 (1943)).

neither from that time period nor the creation of Da Vinci. The fake news article above was not from a real newspaper, and the writer did not base it on any evidence of real events. The fake science article described above wasn't based on actual results or genuine scientific analysis.

For the Justices in *United States v. Alvarez*, the likely answer is that absent legally cognizable harm, fake news and fake science fall squarely in the same category as protected speech. Like religious ideas and political opinions, they are staunchly protected against government censorship. It might seem odd to describe that 2012 case as offering a single answer of any kind to this question. The Court split three ways, with no majority opinion. The specific question that the Justices confronted was whether Congress could constitutionally subject a person, like Xavier Alvarez, to criminal liability for doing what Alvarez had done—namely, to falsely claim that he had received a Congressional Medal of Honor for battlefield heroism that he had never shown on a battlefield where he had never fought.⁴⁶ Both Justice Kennedy's plurality opinion and Justice Breyer's concurrence agreed that false statements are protected speech.⁴⁷ They also agreed that, in this case, the First Amendment barred Congress from criminalizing Alvarez's speech when it had other ways of combatting the effects of his falsehoods—by, for example, creating a public website with an authorized list of Medal of Honor winners that could show that Alvarez was not among them.⁴⁸ But they disagreed on how much protection the First Amendment provides to false statements of fact, with Justice Kennedy insisting that, absent a showing of legally cognizable harm, false statements are just as protected as any other speech⁴⁹ and Justice Breyer arguing government should generally have more leeway to regulate verifiably false statements than it has to restrict other speech content.⁵⁰

Justice Alito's dissenting opinion, by contrast, would have upheld Congress's statutory measure to combat an "epidemic of false claims about military decorations," and would have found that the government acted constitutionally when it punished Alvarez for contributing his own false claim to this epidemic.⁵¹ More generally, he argued, "false factual statements possess no intrinsic First Amendment value," and should receive no First

46. *Id.* at 713.

47. *Id.* at 724 (plurality opinion); *id.* at 730–31 (Breyer, J., concurring).

48. *Id.* at 729.

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

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

51. *Id.* at 739 (Alito, J., dissenting).

Amendment protection except where such protection is necessary to provide sufficient breathing space for true statements on related subjects.⁵²

Yet however much the Justices may have disagreed about the First Amendment status of Alvarez's false claims and similar autobiographical lies, there was much they *did* agree upon—and this agreement reflected aspects of First Amendment doctrine with a decades-long history. They agreed, first of all, that *some* types of false speech—like false speech on “philosophy, religion, history, the social sciences, the arts, and other matters of public concern”—receive the same robust protection that political or artistic expression or expression about the humanities and social sciences receives in other contexts.⁵³ Despite the falsity of these contributions to public discourse, they are *still* contributions to public discourse and, as such, an integral part of the democratic deliberation that the First Amendment strongly insulates from government management. This provides us with an important starting point for analyzing the First Amendment status of fake news and fake science. Both these types of falsity seem to deserve a place on this list. The above-cited fake news article about the 2016 presidential election and the integrity of Ohio ballots was unquestionably about a matter of public concern. It was an instance of the kind of political speech that the Court has often said is at the core of the First Amendment. Scientific debates likewise seem to be the kind of debates that even the *Alvarez* dissent would place off limits to government control.

There is, however, a second point of agreement among the Justices, and it might allow for government regulation of fake news or similar false statements in at least some circumstances: false speech that causes certain legally cognizable harms can be punished or subjected to civil liability without raising significant First Amendment concerns.⁵⁴ Government may punish or subject to liability the harm that defamation causes to reputation, that fraud causes to one's property rights, or that perjury causes to judicial truth-finding.⁵⁵ Even where a lie concerns a political topic, this doesn't mean it can never count as defamation, fraud, or perjury. A person can commit perjury by lying to hide an election law violation.⁵⁶ A newspaper reporter can

52. *Id.* at 746.

53. *Id.* at 751; *see also id.* at 731 (Breyer, J., concurring).

54. *Id.* at 718 (plurality opinion).

55. *Id.* at 718–22.

56. *See* Associated Press, *Former GOP Senate Candidate Convicted of Violating South Dakota Election Law*, FOX NEWS: POL. (May 28, 2015), <http://www.foxnews.com/politics/>

defame a public official by making knowingly or recklessly false claims about that official's alleged failures or wrongs.⁵⁷ If the author of a real newspaper article can defame the person it reports on, this can also, of course, be true of the kind of fake news article that is often constructed around an intentional lie. Consequently, if the political, historical, or scientific focus of a false account places that account safely inside the scope of the First Amendment, a legally cognizable harm caused by that account can push it out.⁵⁸

In short, the Justices in *Alvarez* agreed that government could generally restrict certain harmful false statements, such as those involving fraud, without raising First Amendment concerns.⁵⁹ By contrast, they also agreed that where false statements arise in public debate and concern matters where disagreement is an inevitable and desirable part of that debate—matters such as philosophy, history, social science, and art—then the speaker of that falsity should be just as protected as she is when she speaks a truth.⁶⁰

This two-part consensus about free speech and falsity had roots not only in prior cases, but also in familiar understandings about where government can and cannot freely regulate the claims we make about our activities. I noted above that government frequently prosecutes those who sell fake or counterfeit goods or make material misrepresentations to consumers. This would be impossible unless the First Amendment left government free to punish lies that caused legally cognizable harms. As Justice Kennedy wrote in *Alvarez*, “[w]here false claims are made to effect a fraud or secure moneys or other valuable considerations, say offers of employment, it is well established that the Government may restrict speech without affronting the First Amendment.”⁶¹ By contrast, where false speech is an inextricable part of the debates and deliberations we engage in to think through our political commitments, policy choices,

2015/05/28/former-gop-senate-candidate-convicted-violating-south-dakota-election-law.html (detailing Annette Bosworth's conviction for violation of election law).

57. *See, e.g.*, *Ball v. E.W. Scripps, Co.*, 801 S.W.2d 684, 689 (Ky. 1990) (upholding jury's finding of actual malice in newspaper's reporting about County Prosecutor).

58. *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

59. *Id.* at 729.

60. *Id.* at 751 (Alito, J., dissenting); *see also id.* at 731 (Breyer, J., concurring).

61. *Id.* at 723 (citing *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council*, 425 U.S. 748, 771 (1976)).

or religious preferences, it is in a realm which government generally has no legitimate claim to manage.⁶²

This contrast—between a government’s duty to actively patrol falsehoods in the realm of commerce and security and its obligation to refrain from doing so in the realm of ideas—has deeper roots in liberal political theory. A century before the Constitution’s enactment, John Locke wrote in *A Letter Concerning Toleration* that government is not only authorized, but also duty bound to protect “life, liberty, health, and indolence of body, and the possession of outward things such as money, lands, houses, furniture, and the like.”⁶³ Protecting our physical security and property, he argued, is the reason to submit to government in the first place. The “part of the magistrate,” he argued, “is . . . to take care that the commonwealth receive no prejudice, and that there be no injury done to any man, either in life or estate.”⁶⁴ So government must be given room to perform this role; it has to be able to protect against harm to our persons and property, including harm of the sort that occurs, in part, through deception.

What government was *not* free to do, under Locke’s framework, was to claim authority over the “care of the soul,” or the inward realm of conscience.⁶⁵ Government’s role is not, for instance, to tell us what religious doctrine is deserving of our adherence. And even where opinions are clearly “false and absurd,” government’s role, insisted Locke, “is not to provide for the truth of opinions.”⁶⁶ Instead, it is to assure “safety and security of the commonwealth.”⁶⁷ For Locke, as for American jurists who followed, it was for the individual, not the government, to decide what is true in the realm of conscience.⁶⁸ Locke was primarily interested in keeping government from interfering in our religious lives, but modern-day judges are just as emphatic that government avoid interfering in our formation of political and other opinions. As Justice Jackson emphasized in *Thomas v. Collins*, it is not

62. See, e.g., James Weinstein, *Free Speech and Domain Allocation: A Suggested Framework for Analyzing the Constitutionality of Prohibitions of Lies in Political Campaigns*, 71 OKLA. L. REV 167, 206–13 (2018).

63. JOHN LOCKE, A LETTER CONCERNING TOLERATION 10 (Hudderfeld 1796) (1689).

64. *Id.* at 40.

65. *Id.* at 27.

66. *Id.* at 48.

67. *Id.*

68. *Id.* at 33 (stating that “men cannot be forced to be saved” and that “when all is done, they must be left to their own consciences”).

only in the realm of religion where government is barred from authoritatively distinguishing true and false doctrine.⁶⁹ Rather, government is barred more generally from “assuming a guardianship of the public mind.”⁷⁰ It is barred from telling people what may and may not be a legitimate component of public discourse.

It is thus not surprising that, as the Court of the late twentieth century shaped First Amendment free speech law into a more libertarian, government-limiting doctrine than it had once been, it roughly followed this Lockean template. Indeed, the Court generally divided activity between a realm of physical and financial interactions, where government has a crucial role in securing safety and property, and a realm of ideas, where government must generally let individuals shape their own thought free from government interference.

As I have emphasized before, this Lockean model of government gives us one way to help make sense of why the Court has identified certain exceptions to what is, in current First Amendment doctrine, a “bedrock principle” that speech may not be restricted by government on the basis of the ideas it communicates:⁷¹ while government may not generally target certain speech content for restriction, it *may* do so when the content falls into a particular category the court has recognized as unprotected by free speech law or (in the case of commercial speech) less protected than other kinds of expression.⁷²

The First Amendment, for example, does not place any high judicial hurdles in government’s way when the speech it wishes to restrict consists of “true threats” that communicate a serious intent to commit unlawful violence,⁷³ defamation,⁷⁴ commercial speech,⁷⁵ or certain other

69. 323 U.S. 516, 545 (1945) (Jackson, J., concurring).

70. *Id.*

71. *Texas v. Johnson*, 491 U.S. 397, 414, (1989) (“If there is a bedrock principle underlying the First Amendment, it is that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.”); *see also* *Ashcroft v. Am. Civil Liberties Union*, 535 U.S. 564, 573 (2002) (“As a general matter, . . . government has no power to restrict expression because of its message, its ideas, its subject matter, or its content.”).

72. *See* *R.A.V. v. City of St. Paul*, 505 U.S. 377, 382 (1992).

73. *See* *Virginia v. Black*, 538 U.S. 343, 359 (2003).

74. *See* *R.A.V.*, 505 U.S. at 399, 406.

75. *Id.* at 383–86; *Ohralik v. Ohio State Bar Ass’n*, 436 U.S. 447, 456 (1978) (stating that the Supreme Court has afforded “commercial speech a limited measure of protection, commensurate with its subordinate position in the scale of First Amendment values”).

kinds of content. In at least some cases, the Court has suggested that the reason the First Amendment allows government to regulate such speech is that such speech is not only expressing ideas, but also has powerful effects on the realm of activity where government is duty bound to protect personal security and property. For example, true threats are inextricably linked to potential violence,⁷⁶ and commercial speech is “inextricably intertwined” with buying, selling, and the possible fraud that can accompany it.⁷⁷ These realms of speech, in other words, do not merely involve exchange of thought. Instead, they frequently have powerful and predictable effects in the realm of security and property that most liberal theorists since Locke have identified as the natural realm for vigorous government oversight and action.

This is not to say that First Amendment doctrine’s definition of unprotected speech categories maps perfectly onto the Lockean distinction between the outward realm of the state and individuals’ autonomy in matters of religion or other opinion. First, some categories of unprotected speech do not clearly implicate the state’s duty to protect personal security and property. Unlike commercial speech regulations that protect individuals from deceptive or coercive commercial actors, or proscriptions of true threats that protect physical security, obscenity regulations seem to protect readers’ mental welfare and culture (unless, perhaps, one can justify them as needed to combat certain physical harms that can arise from obscene speech). And certain speech torts may raise similar problems. First Amendment law allows states to impose liability for invasion of privacy, and it may not be clear why the state’s duties to protect security and property give it any more power to restrict privacy-damaging speech (such as that in “public disclosure of private facts”) than any other kind of speech.

Second, it is not only unprotected speech that can have powerful effects on security and property. Speech protected by current First Amendment doctrine often concerns those interests as well. It is not only commercial speech that may strongly impact our financial condition, but also the staunchly protected expression individuals exercise when they exhort policymakers or their fellow citizens to adopt economic policies that are protectionist, socialist, or libertarian in character, or when they argue for one understanding of Fifth Amendment takings jurisprudence

76. *See* *Virginia v. Black*, 538 U.S. 343, 359–60 (2003).

77. *Edenfield v. Fane*, 507 U.S. 761, 767 (1993).

rather than another. Speech in the political realm, after all, is not generally designed only for the purpose of providing material for abstract debates; it is meant to persuade people to take certain actions, to make a concrete difference in the world.

Still, it is at least plausible for courts and legal theorists to explain why most unprotected categories of speech content are unprotected by first emphasizing the effects they have on our physical security or financial condition. Unlike proposals for a certain economic policy, threats of violence are not meant as a contribution to deliberations in which a listener can consider and then accept or reject them. Nor, one might argue, are offers of sale or purchase, which are offered not as ideas to be critically assessed but as appeals to a listener to take some economic action in the near future. There may well be some arbitrariness and artificiality in trying to classify speech as falling on one side or the other of Locke's line between the realm of outward things—those that affect our physical security and property, for example—and the realm of inward deliberation, opinion formation, and conscience.⁷⁸ But one might argue that, under our First Amendment jurisprudence, which embraces at least some expressive or informational libertarianism, courts *have* to insulate *some* deliberative spaces from state control. It requires them to at least attempt to make some distinction between the words or arts that lie squarely in the realm of expression and belief and the words or arts that are a commercial product (like a forged artwork sold on the marketplace) or serve as an instrument of financial harm or potential violence. Some such distinction seems necessary if courts are to draw a line between realms where government can actively regulate false speech that could defraud us or damage our health, and those where the First Amendment bars government from identifying and restricting allegedly inaccurate claims about religion or politics.

It may also be the case that categories of unprotected or less-protected speech are left more vulnerable to government regulation by existing free speech doctrine not only because of what such speech does, but also because of what it does *not* do. Commentators sometimes refer to these categories of speech as “low-value” speech on the theory that they are less likely than political speech, artistic expression, or other robustly

78. See Steven J. Heyman, *Spheres of Autonomy: Reforming the Content Neutral Doctrine in First Amendment Jurisprudence*, 10 WM. & MARY BILL RTS. J. 647, 657 (2002).

protected expression to advance the core values of the First Amendment, such as promoting self-governance or enhancing individual autonomy.⁷⁹

The Court first suggested that certain types of speech content are undeserving of full First Amendment protection in 1942. In *Chaplinsky v. New Hampshire*, the Court held that free speech law does not protect use of “fighting words—those which by their very utterance inflict injury or tend to incite an immediate breach of the peace.”⁸⁰ The Court *might* have justified this conclusion by simply explaining that fighting words are inextricably linked to violent activity the state has a duty to prevent or stop (like the true threats it later made clear also lack First Amendment protection⁸¹). But it instead placed greater emphasis not on the harms that fighting words cause, but rather on value they fail to deliver. Unlike other expression, said the Court, fighting words form “no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.”⁸²

It is not clear, however, that the Court still adheres to this view that unprotected categories of speech are unprotected because they lack the First Amendment value one typically finds in expression. In *R.A.V. v. City of St. Paul*, the Court seemed to forcefully limit the view that fighting words could never form an “essential part of any exposition of ideas,” or that their value is in all respects “worthless.”⁸³ On the contrary, Justice Scalia explained, fighting words may well include a message—and while the violence-generating potential of fighting words *is not* protected, the message *is*.⁸⁴ The Court here analogized government’s restriction of low-value speech to government’s regulation of expressive conduct or of the non-speech effects of protest activity: when government restricts protesters’ burning of draft cards⁸⁵ or blocking of streets,⁸⁶ it is

79. See Genevieve Lakier, *The Invention of Low-Value Speech*, 128 HARVARD L. REV. 2166, 2221 (2015).

80. 315 U.S. 568, 572 (1942).

81. See *Virginia v. Black*, 538 U.S. 343, 359 (2003).

82. *Chaplinsky*, 315 U.S. at 572.

83. 505 U.S. 377, 385 (1992) (citing *Chaplinsky*, 315 U.S. at 572).

84. *Id.* at 386–87.

85. *United States v. O’Brien*, 391 U.S. 367, 369, 370, 376–77 (1968).

86. *Madsen v. Women’s Health Ctr., Inc.*, 512 U.S. 753, 768–770 (1994) (upholding constitutionality of a thirty-six foot buffer zone near an abortion clinic entrance and noting the state’s “strong interest in ensuring the public safety and order, in promoting the free flow of traffic on public streets and sidewalks, and in protecting the property rights of all its

allowed to counter the physical threat or disruption raised by the burning or blocking. But it is *not* allowed by First Amendment law to use its property- and safety-protecting power as a pretext for silencing anti-government messages.⁸⁷ In the same way, Justice Scalia noted in *R.A.V.*, government can protect us from the violence threatened by fighting words (or true threats), but *not* as a pretext for crushing the messages that accompany such intimidation or triggers of violence.⁸⁸ This holding appears to assume that fighting words, true threats, and other historically unprotected content *can* serve as vehicles for communicating ideas⁸⁹ and are thus fair game for government regulation not because they lack such ideas, but because—and only to the extent that—they communicate them in a way that threatens harms that government must have power to regulate in order to fulfill the duties it owes to the public.⁹⁰

In any event, the above analysis suggests that there are at least two reasons that certain kinds of speech content either fall outside of the First Amendment's scope or receive weaker shielding from it. Each of these may help explain why some kinds of lies or other false statements are unprotected. One reason is that while ideas are normally insulated from government control, the expression of ideas can take a form that causes or threatens harm to person or property. This can happen not only when the physical manner of the expression creates a risk of such harm (through non-speech conduct, like burning a flag or occupying space in a way that, for example, affects traffic safety), but also because certain kinds of words or speech *content* can itself cause risks by, for example, conveying a threat of violence or perpetuating commercial fraud. To the extent a lie or other falsity is, in part, the source of such a harm, it may well be outside of the First Amendment.

This helps explain the first of the two points of consensus previously identified in *United States v. Alvarez*: that certain kinds of false statements—like defamation, fraud, or perjury—lie outside the First Amendment because of the harm they work. This is not to say that the Court stands ready to deny First Amendment protection to a category of speech any time that it can be linked to physical or property-related harm.

citizens⁷) (citing *Operation Rescue v. Women's Health Ctr., Inc.*, 626 So. 2d 664, 672 (Fla. 1993)).

87. Geoffrey R. Stone, *Content-Neutral Restrictions*, 54 U. CHI. L. REV. 46 (1987).

88. *R.A.V.*, 505 U.S. at 385–86.

89. *Id.*

90. *Id.*

On the contrary, it has expressly refused to use such a rationale to create *new* exceptions to First Amendment protection, stating in *United States v. Stevens* that courts did not have “freewheeling authority to declare new categories of speech outside the scope of the First Amendment.”⁹¹ Still, the fact that historically unprotected categories of speech tend to produce harm of a kind government has long been duty bound to protect against helps to explain *why* they have been historically unprotected, and perhaps to better define the boundaries of categories such as true threats or commercial speech.⁹²

Second, a type of speech might be unprotected by the First Amendment when it lacks the value that justifies First Amendment protection, providing another potential basis for denying free speech protection to some kinds of lies. Lies might fall outside the First Amendment’s coverage if and when they are devoid of the kind of value that justifies free speech protection, or where their restriction poses no threat to such values. If, as some scholars suggest, the First Amendment exists to promote an individual’s autonomy⁹³ or provide a foundation for democratic deliberation and collective self-government,⁹⁴ there is a case to be made that most factual falsehoods fail to further such purposes—and are more likely to undermine them. This helps explain why the concurrence and dissent in *Alvarez* are more comfortable with giving government more leeway to restrict lies on matters *other than* “philosophy, religion, history, the social sciences, the arts, and other matters of public concern.”⁹⁵ Lies that arise in discussion of the latter topics are more likely to have First Amendment value themselves (or be inextricably intertwined with robust debate that does) than are lies unconnected to discussions which illuminate scientific or philosophical questions or forge policy in democratic discourse.

91. *United States v. Stevens*, 559 U.S. 460, 472 (2010). Courts may, however, recognize categories that have “been historically unprotected, but have not yet been specifically identified or discussed as such in our case law.” *Id.*

92. *Id.* at 470.

93. See, e.g., C. EDWIN BAKER, *HUMAN LIBERTY AND FREEDOM OF SPEECH* 50 (1989); Martin H. Redish, *The Value of Free Speech*, 130 U. PA. L. REV. 591, 620 (1982).

94. See Alexander Meiklejohn, *The First Amendment Is an Absolute*, 1961 SUP. CT. REV. 245, 255; see also James Weinstein, *Participatory Democracy as the Central Value of American Free Speech Doctrine*, 97 VA. L. REV. 491, 513 (2011); Robert Post, *Participatory Democracy and Free Speech*, 97 VA. L. REV. 477, 482 (2011).

95. *United States v. Alvarez*, 567 U.S. 709, 751 (2012) (Alito, J., dissenting); see also *id.* at 731 (Breyer, J., concurring).

These two rationales for denying First Amendment protection to certain kinds of speech, including certain kinds of false statements, are not mutually exclusive. One might argue that there are times they essentially merge into one another: the harm generated by an unprotected category of speech may itself consist of the fact that it undermines a certain free-speech value where we rely on constitutional protections to further it. Thus, Steven Heyman and Christina Wells have each offered distinctive arguments that many low-value categories of speech are low value because they undermine the capacities for individual autonomy that free speech rights are supposed to protect and enhance.⁹⁶ But it is helpful to recognize that excluding certain speech from the First Amendment's scope on the basis of the nature of the harm it causes is conceptually distinct from doing so on the basis that it lacks First Amendment value. These two possible bases for treating certain speech as unprotected sometimes operate separately. As noted earlier, the Court in *R.A.V. v. City of St. Paul* seemed to argue that fighting words can be proscribed because of the harm they cause—not because they lack expressive value, but in spite of that value. And unlike the dissent in *Alvarez*, which argues that lies about one's personal life can typically be regulated (at least in part) because they “possess no intrinsic First Amendment value,”⁹⁷ the plurality opinion finds factual falsity no less valuable or less worthy of First Amendment protection than arguments that are “unreasoned” or “uninformed,” and excludes such lies from the scope of free speech only when they generate “legally cognizable harm.”⁹⁸

The consequence, in any case, is that whatever rationale is adopted for identifying unprotected or less-protected speech categories and explaining why they are outside, or at the outer edges of, the First Amendment, the Justices in *Alvarez* all appeared to agree at least that the realm of false statements fits with this double level of more and less protected speech. Lies that have consequences for financial wellbeing, physical safety, or the wielding of government's coercive force (as the result of a trial, for example) can be targeted or penalized by government. This is true, at the very least because of the harm they raise, and also, perhaps, because they lack the value that inheres in mistaken or misleading political,

96. See Heyman, *supra* note 78; Christina E. Wells, *Reinvigorating Autonomy: Freedom and Responsibility in the Supreme Court's First Amendment Jurisprudence*, 32 HARV. C.R.-C.L. L. REV. 159, 188–95 (1997).

97. *Alvarez*, 567 U.S. at 746 (Alito, J., dissenting).

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

philosophical, or historical claims. Other false statements occur in the realm of vigorous religious, political, or other public debates largely insulated from government control. They may lead us to adopt spiritual, philosophical, or political commitments we come to regret—but these are not the kind of harms government has any role in correcting. And they occur in debates and intellectual explorations whose value would be very much at risk if government were allowed to interfere.

The focus of the disagreement between the Justices in *Alvarez* was in how to assess both the risk of harm and the value at stake in false speech—like Xavier Alvarez’s false claim to have won a Medal of Honor—that, at first glance at least, fits into *neither* of the above two categories. Alvarez’s claim was not a claim about philosophy, politics, religion, science, or some other discipline for how individuals should understand their place in the world. Neither was it a claim regarding how Americans should understand their country and the policies that would best serve it. It was simply an assertion that Alvarez had earned a badge of heroism he hadn’t earned.⁹⁹ Thus, it didn’t automatically receive the robust First Amendment protection that all Justices would offer to claims made about matters of public concern or other areas of knowledge where there is robust disagreement and debate.

Nor did Alvarez’s boast cause a kind of harm that had traditionally subjected speech to common law liability or criminal punishment. Alvarez didn’t perjure himself, make a defamatory claim about someone else, defraud others, or otherwise cause a recognized form of legally cognizable harm. So his fabrication wasn’t automatically excluded from the First Amendment scope as all of these false statements are.

For Justice Kennedy and the plurality, this lack of legally cognizable harm was decisive—if Alvarez’s lie wasn’t the type that is excluded from free speech coverage, then Alvarez must still be inside of (and shielded by) First Amendment’s protection for expressive freedom.¹⁰⁰ For Justice Alito and the dissent, by contrast, the lack of First Amendment value in Alvarez’s mundane lie about a fact concerning his own past was decisive in a different way. Utter falsity, in the dissenters’ opinion, has no value, and where government restriction of it raises no risk for free and vigorous

99. It is still possible to claim that such a false claim about one’s own history might be a statement about a matter of “public concern” and I consider why this might be the case below. See *infra* notes 126–130 and accompanying text.

100. *Alvarez*, 567 U.S. at 719.

public debate, there is no reason for the First Amendment to prevent such restriction.¹⁰¹

For the concurrence, by contrast, the challenge presented by government measures against personal lies was a more complicated one. Instead of finding that personal lies were maximally protected (except where they caused legally cognizable harm) or entirely unprotected (except where they were interwoven with political debate), the concurrence found such lies existed in a challenging First Amendment middle ground. In this middle ground, judges could not know if government restriction was permissible or impermissible under the First Amendment until they took a closer, context-sensitive look at the both the potential harms raised by the falsity in question and the First Amendment benefits that might be lost by letting government attack it.¹⁰²

In any case, government restriction of fake news is unlikely to occasion similar disagreement. The kinds of reports that typically get labelled as fake news tend to focus on the very subjects that all the Justices found to be strongly protected by the First Amendment: they tend to concern political life, historical events, or findings generated by science.

One might wonder why the Justices in *Alvarez* should grant the same First Amendment protection to verifiably false statements concerning philosophy, history, science, or similar topics as is given to claims that are true, might be true, or are opinions (that thus cannot be classified as true or false). Consider a public debate about the wisdom of the continuing U.S. presence in Afghanistan to combat the Taliban and those they support in the Al Qaeda. It is probably clear to most Americans why the First Amendment staunchly protects a person's right to call for withdrawing all American troops from Afghanistan or to state, more generally, that our presence in Afghanistan "accomplished nothing of value," even if government officials and many members of the public believe such a statement is entirely wrong and that, for example, important counterterrorism objectives require a continued U.S. military presence there. But why, one might wonder, should First Amendment protection also extend to false factual claims that a person makes about American operations or what led to them? Why, for example, would the First Amendment protect a person in giving a false account of an alleged high-casualty American operation that never took place? Or in falsely

101. *Id.* at 739, 746 (Alito, J., dissenting).

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

insisting that the terrorist attacks Al Qaeda conducted on September 11, 2001, and commanded while being sheltered by the Taliban, actually had nothing to do with Al Qaeda and were ordered by the Bush administration itself?

Other scholars have provided cogent explanations as to why such false statements of fact might receive First Amendment protection, some of which have been endorsed by the Supreme Court in past First Amendment cases.¹⁰³ In a careful and systematic analysis of this question, Helen Norton identifies three principal reasons that government measures taken to restrict lies, or possibly other false statements of fact, may violate the First Amendment. One is that some lies have value. Lies can protect privacy; they can “trigger confrontation and rebuttal” and by doing so “lead to increased public awareness and understanding of the truth.”¹⁰⁴ In other cases, lies may be an essential part of strategies for uncovering the truth. Investigative journalists and police informants, for instance, might lie about their identities and motivations because wrongdoers would be unlikely reveal the harms they inflict on the public to people they knew to be serving the public interest. This is one of the reasons that Justice Breyer, in his *Alvarez* concurrence, finds that lies and other false statements sometimes receive First Amendment protection.¹⁰⁵

A second reason for First Amendment protection of false facts, notes Norton, is not that lies have intrinsic value but rather that, in a world where false speech could be subjected to harsh punishment, individuals may avoid even true or otherwise valuable speech out of fear that they might include falsity.¹⁰⁶ As the Supreme Court noted in *New York Times v. Sullivan*, not only might the fear of slipping into falsity chill other speech; the fear of being accused of falsity (and having to defend against it in court) even when one hasn’t said or written anything false may achieve the same end.¹⁰⁷ Moreover, it is at times unclear whether a statement constitutes an opinion or idea or instead constitutes a fact. In *Gertz*, the Court stated “there is no such thing as

103. See, e.g., Alan K. Chen & Justin Marceau, *High Value Lies, Ugly Truths, and the First Amendment*, 68 VAND. L. REV. 1435 (2015); Helen Norton, *Lies and the Constitution*, 2012 SUP. CT. REV. 161, 200; Frederick Schauer, *Facts and the First Amendment*, 57 UCLA L REV 897 (2010); Jonathan D. Varat, *Deception and the First Amendment: A Central, Complex, and Somewhat Curious Relationship*, 53 UCLA L REV 1107 (2006).

104. Norton, *supra* note 103, at 165–66.

105. *Alvarez*, 567 U.S. at 733 (Breyer, J., concurring).

106. Norton, *supra* note 103, at 169–79.

107. 376 U.S. 254, 270, 300 (1964).

a false idea” (or opinion, for which it used idea as a synonym).¹⁰⁸ But there can be, it said, a “false statement[] of fact.”¹⁰⁹ One might thus propose a First Amendment regime that bars government from punishing opinions it identifies as unworthy but allows it to punish facts it can demonstrate to be false in a way that would presumably count as objective. It may not be clear, however, when a statement can be demonstrated to be false in this way.

For example, the claim that “serious crime is rising in major American cities,” may be regarded as factually false if it is taken to mean—as many listeners would understand it—that the number of recorded murders and other violent crimes is higher this month or year than in previous time frames. In that case, one could presumably compare the statement to a trustworthy analysis of some trustworthy source of data. But the statement could also be understood as a vague claim about certain unspecified “serious” crimes becoming more of a problem in unspecified “American cities” (and “rising” could refer not to numbers but to increased effects, or fear of, crime). Other claims that make a statement, but leave out context, may mislead some audiences even though the speaker may assume that a critical listener should be aware that statements cannot always include context. The consequence of this blurriness in the boundary between a verifiably false fact and an unverifiable expression of one’s belief or perception about the world is that individuals may end up on the wrong side of this poorly marked line. And, as a consequence of crossing such a boundary, an individual might lose First Amendment protection for their speech, or at least create a situation where an opponent can bring a suit (or a prosecutor to bring charges) by presenting a claim as a false factual statement even where it may not actually be. A First Amendment guarantee against such legal trouble or punishment gives people freedom to engage in robust debate without worrying that they will be sued or penalized for it.

Third and finally, Norton writes, the First Amendment may bar government from targeting lies in public discourse, not in order to protect the lying, but in order to stave off the detrimental consequences of letting government exercise coercive authority over the exchange of ideas.¹¹⁰ It was this last reason for protecting falsity that won the explicit support of all of the Justices in *Alvarez*. As Justice Kennedy emphasized, having a government truth commission to identify and punish such falsities would be reminiscent

108. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 339–40 (1974).

109. *Id.* at 340.

110. Norton, *supra* note 103, at 170–72.

of the Orwellian “Oceania’s Ministry of Truth.”¹¹¹ Justice Alito’s dissent likewise emphasized that, in debates about matters of public concern, “it is perilous to permit the state to be the arbiter of truth,” even where the arbitrating is limited to identifying flagrantly false statements about history, philosophy, or other topics.¹¹² Inviting the government to exercise control over public debate is to invite it to wield coercive power in a realm where the public must be as free as possible to form its own opinions and consider all proposals (even those the vast majority of the public regards as preposterous), where individuals must be left free to shape their own ideas, and where society must be left free to forge its own collective preferences.

Jonathan Varat similarly explains that government cannot be given anything close to a blank check to regulate false or misleading information because public discourse and private deliberations alike are full of such information.¹¹³ To give government free reign—or anything resembling it—to restrict falsity would thus effectively give government free reign to restrict much of our communication. Giving government unlimited power to drain our conversations of deceptive content would give it power undermine “our rights to personal and political self-rule.”¹¹⁴

In public debate or personal conversations about a person’s values or understanding of the world, the remedy for falsehood thus comes not from government, in these Justices’ view, but from the critical faculties of the discussion’s participants and other sources they can recruit to evaluate the veracity and quality of claims addressed to them. As Justice Kennedy declared in his *Alvarez* opinion, “[T]he remedy for speech that is false is speech that is true.”¹¹⁵ The First Amendment’s solution to the challenges raised by non-factual or controversial claims is also its solution for verifiably false claims. Just as the permissible First Amendment response to unreasoned speech is not government suppression but rather rational speech, so the permissible response to a “straight out lie” is not government restriction but rather “the simple truth.”¹¹⁶ Rather than having to recruit government coercion to defend against falsity, as individuals do in the commercial realm or in certain security contexts, they can engage in a kind of self-defense by using their own powers of thought and

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

112. *Id.* at 752 (Alito, J., dissenting).

113. Varat, *supra* note 103, at 1108–1109.

114. *Id.*

115. *Id.* at 727 (plurality opinion).

116. *Id.*

expression. Justice Alito did not agree that this stance applied to Alvarez's flagrant lie about his personal history.¹¹⁷ But with respect to matters of public concern, he seemed to take the same approach that Justice Kennedy took to the broader category of speech lacking legally cognizable harm: it is for the marketplace of ideas to identify and correct, not government intervention.¹¹⁸

The question this raises is whether there are times that (1) such self-defense will be insufficient to combat a falsehood in expression; (2) the First Amendment should permit government to offer speakers and listeners some type of aid in this effort, of a kind government is normally barred from providing in individuals' efforts to identify falsehood and resist being convinced by it; or (3) both. These two questions might each receive separate answers. It might at first seem to follow that, if speakers and listeners are unable to identify and avoid being duped by false statements, then some form of government control is needed, and the First Amendment should allow room for it. But it also may not be true that the solution to ineffectual self-defense against falsehood is recruiting government defense. Even if fake news or false speech raises a genuine problem, government intervention may be an ineffective solution—or one that, even if it succeeds, would bring even worse distortion into public debate or personal reflection than the one it was designed to combat. As the Court itself emphasized in *Gertz*, even where we cannot trust the marketplace of ideas to separate truth and falsehood, there is good reason to distrust government with that task even more strongly.¹¹⁹ If we wish to carefully elaborate—and perhaps rethink—the common ground the Justices laid out in *Alvarez* as we confront modern techniques and technologies for spreading fake news, we thus have to offer separate analysis of (1) the ways in which such fake news or similar misinformation might be able to overwhelm our ability to analyze it and counter the deception it aims to generate; and (2) what role, if any, government might play in addressing this problem, and what alternatives exist to government restriction where the latter would be intolerable given our First Amendment commitments. In Part III, I briefly review how some scholars have suggested the First Amendment may allow regulation of fake news.

117. *Id.* at 739 (Alito, J., dissenting).

118. *Id.*

119. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 340–41 (1974).

III. Is There Deceptive Harm Without Legally Cognizable Harm?

Would the framework I have described above need reshaping, or elaboration, to provide government with power to counter fake news, fake video or audio, or other forms of false expression that have raised concerns among journalists, legal commentators, and other observers? There is reason to think so: even Justice Alito's dissenting opinion in *Alvarez*, the opinion most comfortable with utilizing government power to regulate false statements, did not wish to leave the government with power to restrict falsehoods that arise in robust debates about "philosophy, religion, history, the social sciences, the arts, and other matters of public concern."¹²⁰ When fake news sites report voter fraud or other crimes by public figures that never happened, peddle conspiracy theories about the American government orchestrating the September 11 attacks, or claim the Sandy Hook killings were staged, the false information is not only flatly wrong, it is a flatly wrong claim about politics, history, crime, or war. There is little doubt that the questions they address are questions about matters of public concern—however insane and untenable their answers to these questions may be. Thus, they are not merely factually false statements, they are factually false statements that fall into a category of false statement that even Justice Alito assumed—and that all of the Court's Justices agreed—are virtually off limits to government restriction.

Might the First Amendment nonetheless leave government with room to restrict or combat at least some of this falsehood? My main purpose in this Part is to explore some possible answers to this question. The most obvious possible avenue for justifying such regulation is to show that the two categories of lies I have thus far treated as mutually exclusive—lies that count as defamation or otherwise generate legally cognizable harms, and lies about matters of public concern, such as political life, history, and religion—can sometimes come packaged together in the same statement. A lie might be about public affairs while also defaming someone (such as a candidate it falsely describes as having committed a crime).

In such a case, courts are likely to find that government may restrict or penalize the false statement in spite of its connection to matters of public concern in order to prevent the harm it causes. This, after all, is true of other circumstances where the government regulates political speech that

120. *Alvarez*, 567 U.S. at 751 (Alito, J., dissenting).

has a non-speech component, such as speech that threatens violence. A protestor who is passionately expressing anger against Wall Street might still be arrested if he expresses himself by blocking traffic at an intersection or throwing a brick through the window of a bank. Likewise, someone who incites website viewers to kill or injure doctors who perform abortions is still subject to punishment for incitement even if he can show his incitement is motivated by and expresses his strong protest against abortion.¹²¹ The First Amendment will not insulate traffic disruption, vandalism of property, or incitement to violence from government restriction just because such actions are interwoven with a political message.

Similarly, imagine that a blogger maliciously—and falsely—states that a political figure is guilty of bribery. Further imagine that he does so because he wants that political figure, who has called for faster trials and harsher punishments, to understand what it is like to be falsely accused of a crime. The political dimension of the blogger’s false accusation won’t insulate him from liability for defamation. It will, to be sure, make it harder for a plaintiff to prevail. Under existing free speech doctrine, public figures can only prevail in a defamation case if they can show that the defendant had actual malice in making a false claim—that is, that the defendant knew the claim to be false or did so with reckless disregard of its falsity.¹²² But assuming the public figure can meet this hurdle, he can show the legally cognizable harm necessary to attack this falsehood free from heightened First Amendment scrutiny. Thus, where fake news is not merely fake, but is also harmful in a way that the law has viewed as the basis of a common law claim, the First Amendment allows it to be attacked by both the government and private litigants.

One might argue that where lies or other falsehoods serve a particularly valuable First Amendment purpose, they should be shielded from government punishment or civil liability even when they would otherwise count as fraud. Justin Marceau and Alan Chen describe circumstances where investigative journalists have to lie to uncover information about lawbreaking or behavior that merits exposure and public

121. *Planned Parenthood of Columbia/Willamette, Inc. v. Am. Coal. of Life Activists*, 290 F.3d 1058, 1079–80, 1085 (9th Cir. 2002) (en banc) (finding that circulation of “wanted poster” and web site displaying faces and personal information of doctors providing abortions was unprotected by First Amendment).

122. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 281 (1964).

condemnation.¹²³ Just as police informants might perpetuate fraud that would be illegal if they were not police informants, perhaps investigative journalists should similarly be shielded from liability that would otherwise apply when they lie in order to gain access to a facility that is trying to keep secrets from the public. In other words, there may be First Amendment benefits to information gathering that can succeed only with dissimulation; perhaps these benefits transform what would otherwise be low-value misrepresentation (amounting to fraud) into a crucial component of high value expression.

This might also be true in other circumstances. David Han has argued that autobiographical lies are valuable instruments of self-definition. A person exercises autonomy in part by exercising control over how she presents herself to the world.¹²⁴ On Han's account, such autobiographical lies clearly merit constitutional protection when their intended effect is "purely psychological"—that is, where its aim is to shape others' beliefs about the liar rather than to shape others' actions in ways resulting in "material harm."¹²⁵ It is at least conceivable, however, that, in some cases, a lie used as an essential part of such self-presentation—particularly if made to counter a harsh representation of the person by others in a community—might merit free speech protection even where it fits the definition of fraud. Just as investigative journalists may need to engage in deception that is normally impermissible to play a crucial role in promoting self-government and raising public awareness of abuses of power, so individual autonomy may likewise require deception of a kind that is normally illegal or unethical.

For example, imagine a person who revises pages in his diary and then refers to such a forgery in arranging a paid speaking engagement. It is likely that person should face liability for using false information to obtain payment he may not have obtained if he had been honest about his personal history. But, if Han is correct that the power to mold one's biography, even with tall tales, is one variant of a power of self-definition central to First Amendment purposes, then a court might pause before imposing liability on use of such a power. It might use a framework like Justice Breyer's balancing test in *Alvarez* to ask whether the harm done

123. Chen & Marceau, *supra* note 103, at 1438.

124. David S. Han, *Autobiographical Lies and the First Amendment's Protection of Self-Defining Speech*, 87 N.Y.U. L. REV. 70, 117–18 (2012).

125. *Id.*

by a person's creative forgery of his own past personal identity really justifies punishing it in spite of its First Amendment benefits.

The major concern of this Essay, however, is not whether the misrepresentations that all of the Justices in *Alvarez* regard as low value might become high-value First Amendment activity in some circumstances, and thus merit the kind of strict or heightened scrutiny normally denied to them. It is rather the opposite inquiry: Are there times when falsehoods that appear to be on matters of public concern might be treated as akin to low-value speech and, thus, lose the strong First Amendment protection they normally receive, making them potential targets of government regulation? Might fake news or junk science fall outside of the First Amendment entirely despite the fact that it is presented and sometimes perceived as news or as science? Might a fake video depicting a police shooting or battlefield bombing that never occurred be denied First Amendment protection—even where there is no legal argument presented that it contains defamation, or constitutes incitement? If not, might it at least fall into a place with weaker First Amendment protection?

One possible response to these questions is to criticize my assumption that all or most examples of fake news would necessarily count as a part of the categories of falsehoods about “philosophy, religion, history, the social sciences, the arts, and other matters of public concern.” The boundaries of such bodies of knowledge are unclear. Still, if the relevant inquiry is whether the speech is on a “matter of public concern,” the Supreme Court has interpreted this category very broadly. In *Snyder v. Phelps*, a case that asked about the public-concern test in the context of intentional infliction of emotional distress, the Supreme Court said that, although “the boundaries of the public concern test are not well-defined,” it generally encompasses speech that can “be fairly considered as relating to any matter of political, social, or other concern to the community,” or that “is a subject of legitimate news interest; that is, a subject of general interest and of value and concern to the public.”¹²⁶

There is little question, then, that most fake news would fit this description if it weren't fake. Consider false claims of voter fraud designed to nullify the votes of certain groups of voters. Where real voter fraud or voter suppression occurs, this is unquestionably a matter of

126. *Snyder v. Phelps*, 562 U.S. 443, 452–53 (2011) (quotations marks omitted) (quoting *City of San Diego v. Roe*, 543 U.S. 77, 83–84 (2004)).

public concern. So although fake reports of voter fraud—such as those circulated by the Russian operated Twitter account “Tennessee GOP”¹²⁷—are false, they are also about an issue of public concern.

But perhaps, one might argue, certain discrete statements that individuals make in the course of making a statement of political import should *not* automatically be treated as speech on matters of public concern and, thus, insulated from government control. Imagine, for example, that rather than simply stating that he was a Medal of Honor winner and describing his battlefield heroics, Xavier Alvarez made this statement as a prelude to another statement that “having experienced the horrors of war,” he “knows better than most people why the United States’ missile attacks in the Middle East are reckless in risking another war.” In this circumstance, the statement made criminal by the Stolen Valor Act (that Alvarez is a Medal of Honor winner) is simply an autobiographical lie used to make a subsequent political claim more persuasive. It seems unlikely, however, that Justice Alito would thus classify it as a being a statement about a matter of public concern or as insulated from government restriction on the ground that state restriction of such claims would make it too easy for the “state to use its power for political ends.”¹²⁸ And some of the falsity in fake news is similar: when Russian agents established the Twitter account “Tennessee GOP,” this falsely portrayed them as an American, Tennessee-based Republican party organization, but it did so as a prelude to their use of the Twitter handle to make political statements about the election. Others within the U.S. might similarly lie about their own pasts in order to frame the speech they then make about political matters. Thus, the position of the *Alvarez* dissent might allow the state to restrict at least *this part* of fake news and junk science. And Justice Breyer’s concurrence could similarly subject it to a balancing analysis of a sort that government might win.

But this seems problematic, especially given that “boundaries of the public concern test are not well-defined.”¹²⁹ Why should a verifiably false biographical claim made in political discourse or historical argument be any more vulnerable to government suppression than another verifiably

127. See Jake Lowary, *Fake Tennessee GOP Twitter Account Highlighted in Indictment of 13 Russians*, TENNESSEAN (Feb. 16, 2018), <https://www.tennessean.com/story/news/2018/02/16/fake-tennessee-gop-twitter-account-highlighted-indictment-13-russians/345835002/>.

128. *United States v. Alvarez*, 567 U.S. 709, 752 (2012) (Alito, J., dissenting).

129. *Snyder v. Phelps*, 562 U.S. 443, 452 (2011) (quotation marks omitted) (quoting *City of San Diego v. Roe*, 543 U.S. 77, 83 (2004)).

false claim made in such debates? Why, for example, is it any less troublesome to punish Alvarez's lie about past military experience he never had than it would be to punish an intentionally false claim of his about battlefield events that never really occurred? There may be answers one can offer to such questions. But without clearer answers than those provided by the *Alvarez* dissent, such considerations provide reason to adhere to Justice Kennedy's assumption that *all* false statements should be treated like falsehoods in public discourse unless they raise the kind of harms recognized to justify state intervention.

It is worth considering two other varieties of arguments that even such falsehoods on matters of public concern should be subject to some government restrictions or liability in spite of the risks that arise when the First Amendment allows "the state to be the arbiter of truth."¹³⁰ Each variety comes with its own line-drawing challenges. First, some scholars have argued that where people lie to manipulate a listener, such manipulation through false speech should be excluded from First Amendment's protection. Such manipulation, of course, might be accomplished through speech about politics, history, and science. It might also be accomplished by speech that is not necessarily defamatory and does not defraud anybody of anything of value in the way that is typically necessary for false speech to count as fraud under state law. Such manipulation through falsehood may occur, for example, where a fake news piece is designed to manipulate someone into supporting or denouncing a candidate on the basis of false information.

The second argument concerns situations when a falsehood comes from an expert or other speaker with an exclusive claim to certain knowledge—one in whom listeners will predictably and justifiably place their trust even when the speaker conveys her knowledge in public discourse rather than in a fiduciary relationship. Freedom of speech, Robert Post has argued, covers communication that a listener can "autonomously query."¹³¹ But certain kinds of expert knowledge don't fit this description—certain knowledge is produced by experts whom the listener is poorly positioned to question. One might argue that even where such an asymmetry occurs in public discourse, there are situations where government might have at least limited power to assure it is not abused.

130. *Alvarez*, 567 U.S. at 752 (Alito, J., dissenting).

131. Robert Post, *Recuperating First Amendment Doctrine*, 47 STAN. L. REV. 1249, 1254 (1995).

A. Manipulation Through Falsehoods

One circumstance in which courts might allow government to restrict falsehoods—even when such falsehoods deal with political issues or other matters of public concern—is where doing so is necessary to prevent manipulative lying or manipulation through other false statements of fact.

About twenty years before *United States v. Alvarez* was decided, David Strauss argued that the First Amendment should not protect a speaker's use of lies to manipulate a listener such that the listener serves the speaker's ends rather than her own.¹³² In fact, says Strauss, such manipulative lying not only lacks First Amendment value, it violates what he takes to be the core principle underlying First Amendment doctrine: the "persuasion principle."¹³³ The persuasion principle, as Strauss explains, stems from one of the most valuable functions of speech, and the one that the First Amendment is largely designed to protect—namely, the function speech performs when it "persuades," inducing the listener to "action through a process that a rational person would value."¹³⁴ When a person gets a chance to hear and consider acting on ideas, even ideas the state views as dangerous, she is exercising her autonomy vis-à-vis her right to decide upon, and act according to, her own ends. But when the listener is deceived by false speech, she is doing something that autonomous and rational people never want to do: build their actions upon confidence in false factual premises. Moreover, where these false factual premises are fed to her by a speaker who uses them to steer her in ways she would never move herself, her autonomy is not only left unsupported, it is undermined, making the listener a tool of the speaker rather than an agent forming and acting on her own rational decisions.

As a consequence, Strauss concludes that while the First Amendment's support for an individual's intellectual autonomy normally requires it to restrain the state from interfering in the processes by which individuals try to persuade each other, the same is *not* true when those proposals are manipulative lies.¹³⁵ Far from enhancing autonomy, such "[l]ying is the

132. David A. Strauss, *Persuasion, Autonomy, and Freedom of Expression*, 91 COLUM. L. REV. 334, 355 (1991).

133. *Id.* at 335.

134. *Id.*

135. *Id.* at 354.

clearest case of the coercion-like, autonomy-invading manipulation that the persuasion principle is intended to prevent.”¹³⁶ This is, thus, “a core area in which the harm of private manipulation seems great enough to justify government restrictions on speech.”¹³⁷

The consequence of such a framework depends heavily on what makes a lie count as manipulative. On one, fairly broad understanding of what constitutes manipulation, Strauss’s exclusion of manipulative lies from the First Amendment’s scope in some ways could provide a striking contrast with the positions taken by the Justices in *United States v. Alvarez*. This is true when manipulation might occur even where it leads merely to the adoption of certain beliefs or attitudes held by a listener, without some concrete action that the speaker desires the listener to take. It might also be true where the lying *is* designed to produce a specific action by the listener, but one without any legally cognizable harm.

In either case, Strauss’s approach would be different than Justice Kennedy’s argument in the plurality opinion, where lies and other false statements of fact lose First Amendment protection only when they cause legally cognizable harm.¹³⁸ After all, it is hard to argue that such legally cognizable harm results from all lies that are manipulative in the broad sense described above. If Alvarez tells a person that he has engaged in heroic activity on the battlefield, but does so in an attempt to win admiration rather than to defraud his listener of money, such a lie is manipulative because it aims to get the listener to react in ways he would not react but for the lie. But it is unlikely its manipulation would count as the kind of legally cognizable harm that would justify a lawsuit or criminal penalty.

Manipulative lying of this kind would also likely be provided First Amendment protection by all the Justices in *Alvarez* where it not only lacks legally cognizable harm but also deals with issues of politics, history, philosophy, or other matters of public concern. This is the kind of manipulative lying that one might find in articles, tweets, or Facebook posts that spread fake news. Consider, for example, a tweet that makes a false claim about immigrant crime rates in order to motivate its readers to oppose legislative approval of the Deferred Action for Childhood Arrival Act (DACA). If the author of the tweet realizes the information she is

136. *Id.* at 366.

137. *Id.*

138. *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

sharing is false and shares it anyway in order to generate political opposition to DACA, she is manipulating her readers with false information. So long as the claim does not defame particular individuals or organizations, however, this lie is likely to be the kind of lie that the Justices in *Alvarez* were unwilling to exclude from First Amendment coverage and to which they would have provided the same robust First Amendment protection that covers accurate claims about politics or political opinions.

But Strauss's approach is not inevitably in conflict with the *Alvarez* opinions. As noted above, how much room Strauss's framework provides for government restriction of false statements on political matters, or other matters of public concern, would depend on how one defines what counts as manipulative.

One important question is what exactly is denied First Amendment protection by a proposal denying it to "false statements of fact by private speakers."¹³⁹ It seems clear that Strauss means to exclude *lying*—a situation where the false statement of fact is known to be false by the speaker. The speaker "inject[s] her own false information into the thought processes of the listener for the purpose of making those processes produce the outcome that the speaker desires."¹⁴⁰

But lying is not the only way a listener's thought processes can become infected by false information. A listener can also receive false information from a speaker who genuinely believes that the false information is true. This kind of deception occurs, for example, when a conspiracy theorist spreads false information, such as a story that a school shooting was staged by the federal government to justify greater gun control, not in order to intentionally deceive someone but to convince the listener of something the speaker believes. On the one hand, the outcome is still in tension with the persuasion principle. A "rational person," Strauss argues, "never wants to act on the basis of false information," and by urging such a person to act on the basis of false information, the conspiracy theorist is urging someone to act irrationally even though the conspiracy theorist is subject to the same delusion.¹⁴¹ On the other hand, from the conspiracy's theorist own vantage point, she is, unlike a liar, doing something the First Amendment staunchly protects: trying to

139. Strauss, *supra* note 132, at 366.

140. *Id.*

141. *Id.* at 335.

convince someone else to adopt her own conclusions by presenting what she regards as valid evidence for such conclusions. And it seems quite likely that much of the fake news spread on Twitter is believed to be true by those who disseminate it.

Second, one might question not only whether the manipulative use of falsehood loses First Amendment protection even when it is deception other than lying, but also what counts as manipulative for purposes of applying this principle.¹⁴² Consider, again, why lying is a perfect example of the “coercion-like, autonomy-invading manipulation that the persuasion principle is intended to prevent.”¹⁴³ It is because a “speaker tells a lie in order to influence the listener’s behavior, . . . making it serve the speaker’s ends instead of the listener’s.”¹⁴⁴ What kind of goals count as influencing behavior, or making someone serve the speaker’s ends? In cases where such manipulation might amount to fraud, it is because the victim of the fraud has been tricked into paying money or sacrificing something else of value. It seems likely that manipulation has also occurred where someone lies to a voter in order to get that voter to cast her vote for the opposing candidate. But in many false statements of fact online, the goal is vaguer. The false statement is not intended to generate a specific action, but rather to generate anger or disdain in the listener towards a certain political or social group. In other words, the false statement is not made in order to get them to take a specified action, but rather to encourage a general opposition to certain political views or leaders. Consider some of the information spread by the Russian Internet Research Agency, whose members have now been indicted by a grand jury for various violations of United States law.¹⁴⁵ This group spread information on Twitter making false claims, for example, about voter fraud by Hillary Clinton.¹⁴⁶ But their goal was not simply to motivate readers to vote against Clinton, but instead to “sow discord in the U.S. political system.”¹⁴⁷ In some cases, liars may wish to spread information just because they want others to share their ideology, perhaps to advance it in practical

142. This is a question that has also been discussed in other analyses of lying’s First Amendment status. *See, e.g.*, Han, *supra* note 124, at 115–19 (exploring how one can draw a distinction between lies with “purely psychological effects” and thus that result in harm).

143. Strauss, *supra* note 132, at 366.

144. *Id.*

145. Indictment, *supra* note 28.

146. *Id.* at 19.

147. *Id.* at 4.

ways in some future day or perhaps, in the near term, simply to make the ideology less of a fringe belief.

First Amendment principles provide good reason to err on the side of adopting narrower, more restrictive answers to the above questions. First, it is dangerous to give government significant power to punish, or subject to liability, those who spread falsehoods that they believe to be true. Again, if we passionately defend statements about politics that we believe to be true, we are engaging in—what from our perspective—is persuasive speech of the kind that lies at the core of the First Amendment. Where someone’s reputation is at stake, the law *will* sometimes make us liable for false statements we believe to be true. Individuals can sometimes be liable for defamation even when their reporting of false information is negligent or reckless rather than intentional lying. But applying the same standards of negligence or recklessness to general public statements would likely chill great swathes of public discourse. Second, if manipulation is defined broadly enough to include *all* effects a speaker desires to cause in a listener’s mind or disposition to take further unspecified action, then virtually all lying would become manipulative lying.

Such considerations weigh in favor of defining “manipulative” quite narrowly. And Strauss seems to lean this way: in calling for the exclusion of false statements from the First Amendment’s scope, he adds the caveat that, although letting the government restrict “false statements of fact by private speakers . . . will do more good than harm,” doing so is likely to be safe for First Amendment freedom only if “the category of false statements of fact is . . . defined very narrowly.”¹⁴⁸

The upshot of this analysis is that Strauss’s arguments for excluding manipulative lying from the scope of First Amendment protection aren’t likely to leave much more speech unprotected than the Justices in *Alvarez* already did when they concluded that lying causing legally cognizable harm is outside the scope of the First Amendment. This is true because if manipulative lies are defined narrowly to cover only those lies where a speaker uses falsity to cause the listener to act in a specific way, many such lies will, in any event, count as fraud or some other already recognized legally cognizable harm.

Such narrowness in defining manipulative lies is also important for another reason. Even when lying is manipulative in that it is aimed at making a listener embrace certain beliefs she would otherwise reject, and perhaps

148. Strauss, *supra* note 132, at 366.

then act in accord with those beliefs, some of the lies scholars and judges identify as valuable may have this character—at least, if the category of manipulative lies is defined too broadly. As David Han argues, autobiographical lying is often used by individuals to engage in self-definition.¹⁴⁹ While listeners may not want to be deceived by such lies from others, they might strongly prefer a world where such autobiographical lying is legal, and thus constitutionally protected, so they remain free to engage it themselves. Remaining vulnerable to such deception, in other words, may be a price they are willing to pay if that is the only way they can retain for themselves the continued freedom to present themselves in the way they wish to be seen. They may likewise want to live with the risk of being deceived by an investigative journalist if doing so is the price they have to pay for such journalists to expose dangers they and their fellow citizens need to know about.

B. Expert Truths

The premise I explored in the last subsection was that it is not at odds with the First Amendment for government to assume that all individuals have a duty to avoid manipulating others with false information. But it is also possible to conceive a narrower truth-telling duty, one that generally binds not all individuals, but only those who have expert or specialized knowledge or, for some other reason, have to be trusted by listeners. As Robert Post writes, much of our knowledge of the world comes not from direct perception, but rather from what we learn from scientific experts, historians, or other experts whose methods for producing that knowledge are quite different from the chaotic debates one finds in public discourse. As Post writes, “[E]xpert knowledge requires exactly what normal First Amendment doctrine prohibits.”¹⁵⁰ Where normal public discourse requires that we decide for ourselves what opinions to embrace, without government or any other authority ordering us to favor one opinion over another, science, history, and journalism can produce content “we have reason to trust” only if the right kinds of experts can “distinguish meritorious from specious” claims.¹⁵¹

It is not always the case, however, that affirmations by experts or others with privileged knowledge (such as an eyewitness) are required, by law, to be accurate. That depends on the context in which they occur. In certain

149. Han, *supra* note 124, at 115.

150. ROBERT C. POST, DEMOCRACY, EXPERTISE, AND ACADEMIC FREEDOM: A FIRST AMENDMENT JURISPRUDENCE FOR THE MODERN STATE 9 (2012).

151. *Id.* at 13.

circumstances, experts have a legal duty to provide statements consistent with the standards of their discipline. This occurs, for example, when an expert is hired to provide medical or legal services. A doctor who has agreed to treat a patient has to answer that patient's medical inquiry with an answer that constitutes competent practice of medicine. And a lawyer has to provide her client with competent legal advice. Similarly, expert witnesses at trial are bound to provide honest answers about their knowledge and may not be called upon to do so unless their expertise is genuine and is of the kind that will be helpful to the trier of fact in addressing particular questions.

But in the realm of public discourse, even such experts are no longer bound to voice disciplinary wisdom. Whereas a doctor's advice is supposed to accord with standards of her profession in her treatment of a patient, she is not similarly constrained when she writes a book or a newspaper column. In "speech to the general public," writes Post, a doctor (or other expert) has free speech rights they do not have in the conversations they have with their patients.¹⁵² The doctor might even receive protection if she falsely claims, for example, that evidence shows vaccination increases the risk of autism when it does not. Of course, she may suffer the opprobrium of colleagues and others in the medical community if she makes false or misleading medical claims in public discourse. She may find it difficult to publish an article about vaccination in any journal highly regarded within the medical community. But, thanks to the First Amendment, she cannot suffer a penalty at the hands of the state or federal government.¹⁵³

How does this relate to fake news or junk science? It is useful to focus on the latter example first. One might argue that perhaps—even in public discourse, on matters of public concern—the government ought to be able to impose limits on when experts can invoke their reputation to promote, as confirmed facts, claims that are roundly rejected by everyone else in their field. This is a possibility Jane Bambauer has recently explored in her scholarship. Where an expert's false statement is likely to be relied upon by a listener in ways that cause harm, she argues, the statement's falsity should be fair game for government restriction even if it occurs in public discourse.¹⁵⁴

Rather than generating rules for "false statements of fact," she instead considers how different First Amendment rules might apply to three categories of knowledge she calls "accepted knowledge," "contested

152. *Id.*

153. *Id.*

154. Jane Bambauer, *Snake Oil*, 92 WASH. L. REV. 73, 76 (2018).

knowledge,” and “anti-knowledge.”¹⁵⁵ The first and third of these categories, respectively, are refinements of what many others call true and false knowledge. Accepted knowledge consists of knowledge that is “supported by enough observations and credible evidence to clear the high bar established by the relevant experts” and others who are able to apply whatever epistemological standards determine what observation and evidence are sufficient.¹⁵⁶ Anti-knowledge includes “statements that are in direct conflict with the statements contained in accepted knowledge,” such as “sets of claims that have been proven, based on prevailing scientific standards, to be incorrect.”¹⁵⁷ In between these two categories is contested knowledge, consisting of “claims that may have some evidence in support, and perhaps some evidence in conflict, but not enough of either sort to conclusively place the statement into the accepted knowledge or anti-knowledge buckets.”¹⁵⁸

While government can, of course, already restrict anti-knowledge where it comes in the form of bad medical advice given by a doctor to a patient, Bambauer argues that “government should also be permitted regulate anti-knowledge within the public discourse so long as the claim is likely to cause the listeners to take action that puts themselves or others in serious risk of harm, and the speaker has a sufficiently culpable mental state.”¹⁵⁹ This argument is focused on scientific knowledge in public discourse, and she illustrates it by questioning the Ninth Circuit’s conclusion in *Winter v. G.P. Putnam & Sons* that the publishers of a reference book called “The Encyclopedia of Mushrooms” could not be liable, under product liability law, for identifying dangerous mushrooms as safe to eat.¹⁶⁰

It is not clear how much application such an analysis might have to journalism, to claims about politics or history, or to areas of science where a person’s health, safety, or property are not at stake. Junk science and bad medicine can lead individuals to place themselves in danger (for example, by eating poisonous mushrooms they fail to realize are poisonous). Erroneous directions on a chart or map can likewise cause trouble by leading someone

155. *Id.* at 111–12.

156. *Id.* at 85.

157. *Id.*

158. *Id.* at 86.

159. *Id.* at 133.

160. *Id.* at 89–90 (citing *Winter v. G.P. Putnam’s Sons*, 938 F.2d 1033, 1037–38 (9th Cir. 1991)) (finding that the content of books were not suitable subjects for product liability claims).

into dangerous territory.¹⁶¹ But it seems less likely that equally concrete harms will flow from falsehoods about politics or history. Individuals may regret a vote they cast on the basis of inaccurate political information or a protest they attended to express outrage against an event that never occurred. But such regrets are similar to the abandoned spiritual or ideological commitments which, as I noted earlier, individuals can't treat as harms sufficient to create legal liability for those who lured them into a religious or political movement. A journalist's flawed reporting might conceivably cause more harm—if, for example, it includes inaccurate descriptions of dangerous events (such as an impending tornado). But the question that needs answering here is why the First Amendment should allow any greater restriction of such speech in public discourse than it already does when it permits people to sue for defamation or intentional infliction of emotional distress, or allows the state to punish incitement, true threats, commercial speech, and low-value speech.

The same question could arise about another variant of this argument: it is also possible, one might claim, that First Amendment law could treat certain false claims even by non-experts as subject to duties to avoid harmful false statements about facts to which a speaker has exclusive access. A comparison with trial evidence may be helpful here. It is not only the expert witness who brings to such a trial knowledge to which he has special claim. Rather, it is also the lay witnesses who have directly perceived events or other aspects of the world that no one else might have seen or heard.¹⁶² Just as the expert witness must testify honestly regarding the subject of his expertise, the lay witness is under an obligation to testify honestly about his experiences. Public discourse, of course, is not subject to the same constraint. If I write a blog post about events I have seen or heard or describe them in response to inquiries by a reporter, I am not under the same obligation to tell the truth that I am when I am on the witness stand. One might argue, however, that where a lie about my personalized knowledge will predictably cause certain harms to individuals who rely upon this knowledge to make certain decisions, then such a lie should perhaps be more vulnerable to government restriction than a lie I tell about some other subject (which no one has any need to rely upon). But again, one may ask, why doesn't existing free speech doctrine on low-value categories of speech—or other speech outside the boundaries of the First Amendment—already provide all the guidance one needs for addressing

161. *Id.* at 91.

162. FED. R. EVID. 701.

when dishonest reports of personalized knowledge may be punished or subjected to civil liability? Why not hold, as Justice Kennedy argued in *Alvarez*, that a false report of an event may be punished when one lies to a police officer or other government official; when one commits perjury, or defrauds or defames someone; or when one otherwise causes a harm the legal system has recognized as a basis for a lawsuit or prosecution instead of granting the state the potentially dangerous power to punish any falsehood that it can characterize as a source of harm? Answering such a question once again depends on discussing line drawing and describing more specifically how the line between speech where reliance interests are and aren't present in public discourse would differ from the line drawn in Kennedy's plurality opinion between lies that cause legally cognizable harm and those that do not.

IV. Free Speech and the Future of Forgery

When someone critically assesses an argument, like the one I am making in this Essay, they can do so without seeing the world as I see it or asking me to provide facts in my sole possession. Someone might read this piece, consider its claims and the reasons it offers them, and decide that the reasons don't convince her. They might do the same with any of the opinions in *United States v. Alvarez*. One does not have to learn about previously unknown facts, or rely heavily on social science or other data, to come to the conclusion that they disagree with some of the reasoning of Justice Kennedy's plurality opinion, Justice Breyer's concurrence, or Justice Alito's dissent. Of course, certain data may well be helpful in judging certain claims within such arguments. It might be good to have data, for example, on the psychological processes that occur as individuals try to evaluate arguments. But such data isn't essential for a reader to make some judgments about the persuasiveness of any of these opinions, or a law review essay about them. This is thus one situation where individuals at least have the potential to act as their own watchmen for truth. In any event, this is not a role government can constitutionally wrest from individuals through use of its coercive power: we cannot trust government to mandate which arguments we should accept and which we should reject. And even though some questions require an authoritative answer from a court or legislature so that law and policy can take a certain form, that only means that we have to live under that decision, not that we have to agree with it or consider it correct.

A person is typically less self-reliant, however, when assessing a factual claim about the external world. There aren't many such claims I am in a position to test against my own personal experience. Rather, I rely

on a complex array of social practices and practical realities to help me expand my factual knowledge beyond the very narrow sample of reality I can perceive by myself. For example, if I wanted to know, while in Oklahoma, the weather conditions in New York City, I would have to call a friend there and ask them, watch a weather report on TV that provides that information, or check a weather-oriented website or smartphone app. Similarly, almost the entirety of my knowledge of the facts central to the nation's political life comes second- or third-hand as well; indeed, I don't interact with the key players in Congress or the Executive Department regularly and have met only a few of them. Thus, my knowledge of what President Trump or members of the Senate or House of Representatives are doing each week has to come from other reports whose veracity I can trust—this time not from distant friends, but from the journalists who, as the Supreme Court once said, “act as the ‘eyes and ears’ of the public.”¹⁶³

At times, the inaccessibility of a factual claim results not only from the great distance or other barriers between me and the fact or event I can't directly observe, but also because it is the kind of fact the world conceals from our senses entirely until experts use scientific experimentation and analysis to extract it. We only know about the role DNA plays in cell function and reproduction or about the properties of electrons, for example, because scientists have been able to learn such information from a systematic study of the world.¹⁶⁴

The inaccessibility of facts to any one individual raises another potential challenge for the First Amendment model that the Supreme Court developed during the mid-to-late twentieth century and that served as crucial background for *United States v. Alvarez*. That is, we can act as our own watchmen for truth, when it comes to factual claims, but not because we are capable of perceiving most of the factual realm directly. We clearly aren't. Rather, we play this role *only* because we can rely on and trust in a complex set of social institutions and practical technologies to reveal aspects of the world to us. Americans and members of certain other Western societies can learn about and monitor their countries' political life only because they can rely, to some extent, on journalists, historians, and, at times, fellow citizens to share and aggregate information about such events. We can understand the hidden

163. *Houchins v. KQED, Inc.*, 438 U.S. 1, 8 (1978).

164. POST, *supra* note 150, at 1, 7 (“In fact I have learned about the carcinogenic properties of cigarettes by studying the conclusions of those whom I have reason to trust.”).

characteristics of the world that scientists uncover only if the scientific community accurately and honestly uses its disciplinary study to uncover this realm and vet the work of other scientists to assure it is reliable.

In a sense, these social practices extend our perception. Video and audio technology are perhaps the most obvious examples of such extended perception: they allow us to see and hear events that we cannot witness personally. But we also rely on other individuals' perception—like that of journalists or the individuals interviewed by journalists—and treat it as a basis for knowledge about the world.

One starting point for elaborating and perhaps revising the consensus understandings in *United States v. Alvarez*¹⁶⁵ is to make it clear that the First Amendment should, if possible, leave the government with room to protect us not only from false speech that causes us material harm but also from falsity that inserts itself into, or convincingly disguises itself as, at least some of the channels of indirect knowledge that we crucially rely upon to deliver perceptual experiences beyond those which we can create for ourselves.¹⁶⁶

And modern technology has, in some ways, made us far more vulnerable to such attacks on the knowledge ecosystem for this purpose. As noted earlier, the rise of computer technology has made certain kinds of forgeries and imitations easier. This is, in part, perhaps because we have altered some of the aspects of the knowledge ecosystem in ways that make it easier to fake. For example, we increasingly communicate by email or text messaging in circumstances where we would have previously used a phone call. Where we once might have had a good basis to know it was really our employer, colleague, or friend contacting us on the phone with new, important information about significant developments in the world, we may now find that we too quickly accept as real an email purporting to be from that employer, colleague, or friend but that really is from someone who is imitating them to manipulate us. Thus, cybersecurity specialists now face the challenge of helping people avoid falling for “spear-phishing” emails, wherein attackers “disguise

165. For instance, that false statements on matters of public concern are fully protected unless they cause legally cognizable harm. *See* 567 U.S. 709, 719–20 (2012).

166. *See* Jonathan Adler, *Epistemological Problems of Testimony*, STANFORD ENCYCLOPEDIA OF PHIL. Winter 2017, <https://plato.stanford.edu/archives/win2017/entries/testimony-episprob/> (noting that every person has no choice but to rely on reports, the accuracy of which one cannot “check on . . . for oneself”; rather, we must “depen[d] on the reliability and sincerity of others, often strangers or mere acquaintances”).

themselves as a trustworthy friend or entity to acquire sensitive information,” using information about the intended victim such as “their friends, hometown, employer, locations they frequent, and what they have recently bought online.”¹⁶⁷ Some such attacks also disguise themselves as internal company emails, using information gathered from surveillance of companies.

Our increasing reliance on internet-based communication may likewise make us vulnerable not only to fake personal communications but also to fake news. Now that newspapers are on websites, the code or design of which can be easily copied by digital means, creating fake versions of established newspapers is far simpler than it was when newspaper production relied on possessing and using a powerful printing press. And fake newspaper creators who are unwilling to risk subjecting themselves to trademark infringement suit from the *New York Times*, the *Miami Herald*, *Time Magazine*, or another publication might instead create an entirely fake publication with the appearance and feel of a real newspaper or magazine. The creator of the Christian Times Newspaper, the fake publication that ran a made-up story about election fraud, attempted to do this. Legitimate news sites themselves have arguably contributed to this problem. When a news network’s home page includes links not only to the journalism it produces itself but also to commercially sponsored “stories” that nonetheless have an appearance very similar to those of the news network, this risks confusing readers about what content is produced according to the normal practices of journalists and what is offered by commercial entities who are doing so with an eye to encourage particular types of consumer behavior.

As noted before, developments in audio- and video-editing technology raise an even more significant threat. Almost twenty years ago, in 1998, science fiction writer and technologist David Brin warned, “One of the scariest predictions now circulating is that we are about to leave the era of photographic proof. . . . We are fast reaching the point where expertly controlled computers can adjust an image, pixel by microscopic pixel, and not leave a clue behind.”¹⁶⁸ Now, many articles are reporting such a

167. Nena Giandomenico, *What Is Spear-Phishing? Defining and Differentiating Spear-Phishing from Phishing*, DIG. GUARDIAN BLOG (June 27, 2016), <https://digitalguardian.com/blog/what-is-spear-phishing-defining-and-differentiating-spear-phishing-and-phishing>.

168. DAVID BRIN, *THE TRANSPARENT SOCIETY: WILL TECHNOLOGY FORCE US TO CHOOSE BETWEEN PRIVACY AND FREEDOM?* 28 (1998).

similar technological transformation is occurring not just in the realm of still pictures but also in the realm of audio-visual recording. National Public Radio (NPR), *The Verge*, and other media outlets, for example, recently ran stories on a Canadian company called Lyrebird that, according to NPR, has “come up with a way to recreate anyone’s voice and get it to say almost anything.”¹⁶⁹ It uses computer algorithms to capture the distinctive features of someone’s voice from a voice sample as short as a minute in length. Programs can then produce a voice eerily similar to that of Donald Trump, Barack Obama, or any other voice they analyze and instruct it to say whatever content is fed to it. The voices produced are still noticeably artificial, but this technology will only improve.¹⁷⁰ And, as *The Verge* reports, Lyrebird’s audio simulation technology can reportedly “infuse the speech it creates with emotion, letting [users of the software] make voices” that they simulate “sound angry, sympathetic, or stressed out.”¹⁷¹ Further, Lyrebird is not the only company creating such technology. Google has also produced technology for simulating distinctive voices, and Adobe’s Project VoCo “can edit human speech like Photoshop tweaks digital images.”¹⁷²

Other researchers at the University of Washington and Stanford have generated tools for manipulating video of speakers. Using this technology, a video of Donald Trump or Barack Obama can be altered so that it shows their lips forming words they never said, to go with the fabricated audio of words they never voiced.¹⁷³ As *The Guardian* notes, this “new breed of video and audio manipulation tools, made possible by advances in artificial intelligence and computer graphics, . . . will allow for the creation of realistic looking footage of public figures appearing to say [] anything.”¹⁷⁴ The article describes this as “the future of fake news” and noted it means we will have to question not only “everything we

169. *New Software Can Mimic Anyone’s Voice*, NAT’L PUB. RADIO (May 5, 2017), <http://www.npr.org/2017/05/05/527013820/>.

170. See James Vincent, *Lyrebird Claims It Can Recreate Any Voice Using Just One Minute of Sample Audio*, VERGE (Apr. 24, 2017, 12:04 PM), <https://www.theverge.com/2017/4/24/15406882>, (“The results certainly aren’t indistinguishable from human speech, but they’re impressive all the same, and will no doubt improve over time.”).

171. *Id.*

172. *Id.*

173. Olivia Solon, *The Future of Fake News: Don’t Believe Everything You Read, See or Hear*, GUARDIAN (Jul. 26, 2017), <https://www.theguardian.com/technology/2017/jul/26/fake-news-obama-video-trump-face2face-doctored-content>.

174. *Id.*

read, but soon we'll have to question everything we see and hear as well."¹⁷⁵ As with audio simulation, the video simulation is not quite yet good enough to fool a careful observer.¹⁷⁶

One widely discussed version of this fabricated reality is the deep fake, a video scene that, as discussed above, shows someone doing something they didn't actually do, thanks to sophisticated computer technology that imposes their face on the body of the person actually in the video footage. The computer technology involves a kind of artificial intelligence called "deep learning" (hence the name "deep fake").¹⁷⁷ To date, the technology's most noticed use has been to create pornography films featuring celebrities who neither appeared in them nor consented to have their image use for such purposes.

But like other forms of fake video and audio, deep fakes can be used for other malicious purposes. As discussed in the introduction, a recent *Lawfare* blog post uses the term deep fake to describe alteration of an image, video, or audio source to make it appear that someone did something that they did not actually do—and with such technological sophistication that the "unaided observer cannot detect the fake."¹⁷⁸ Such fabrication of reality, the blog post notes, can create havoc for the way individuals understand the world and thus unbalance their foundation for action.¹⁷⁹ Moreover, such fabrications might not only gull individuals into taking actions with negative personal consequences, but also may lead to harms that tear "the very fabric of democracy"—for example, by generating fake evidence of race-based violence or of war crimes.¹⁸⁰ Interestingly, the authors stress the threat raised by deep fakes by revising the same metaphor that Justice Oliver Wendell Holmes used to explain, in 1919, why First

175. *Id.*

176. Greg Kumparak, *This System Instantly Edits Videos to Make It Look Like You're Saying Something You're Not*, TECHCRUNCH (Mar. 18, 2016), <https://techcrunch.com/2016/03/18/this-system-instantly-edits-videos-to-make-it-look-like-youre-saying-something-youre-not> ("It's not pixel-perfect yet—even in the relatively low-res clips we're shown, there's an uncanny valley effect of something being not quite right.")

177. James Vincent, *Why We Need A Better Definition of 'Deepfake'*, VERGE (May 22, 2018, 2:53 PM), <https://www.theverge.com/2018/5/22/17380306/deepfake-definition-ai-manipulation-fake-news>.

178. Bobby Chesney and Danielle Citron, *Deep Fakes: A Looming Crisis for National Security, Democracy and Privacy?*, LAWFARE (Feb. 21, 2018), <https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy>.

179. *Id.*

180. *Id.*

Amendment speech rights must have limits. Holmes “warned a century ago of the danger of shouting fire in a crowded theater”; with the use of deep fakes, “now those false cries might go viral, fueled by the persuasive power of hyper-realistic evidence in conjunction with the distribution powers of social media.”¹⁸¹

As many writers have noted, more primitive equivalents of such tools are already being used by those who wish to commit fraud. Already, many criminals are trying to use fake identities to commit computer crime or other crime that requires deception.¹⁸² As the fact-checking website Snopes.com describes the scheme, “a scam artist gleans just enough information about a family (e.g., names, ages, addresses, phone numbers) to be able to impersonate one of them during a brief phone call to another family member.”¹⁸³ The scam artist then calls a member of the family—usually a grandparent—claiming to be the grandchild and facing significant distress and in need of money.¹⁸⁴ Such scams have worked even without the technology I have described above. Such a scam could, of course, be far more convincing if the impersonator could not only use the grandchild’s name and other information, but also a carbon copy of her voice. This technology also gives criminals methods of creating even more damaging versions of the spear-phishing emails described above. Instead of an email from your employer asking you to provide certain sensitive information to a certain email address or wire money to a bank account, you might receive a phone call in which such an instruction comes from a simulation of your employer’s voice.

There are, to be sure, already laws that allow government to aggressively pursue and seek punishment for criminals who use such methods. The Computer Fraud and Abuse Act already criminalizes using some of these methods to gain unauthorized access to computer information (or access that exceeds what is authorized).¹⁸⁵ Identity theft and wire fraud laws might also apply. Many states have laws defining and imposing punishment on “criminal impersonation” or “false

181. *Id.*

182. *See Family Member Impersonation Scam*, SNOPE (Aug. 3, 2014), <https://www.snopes.com/fact-check/the-family-way/>.

183. *Id.*

184. *Id.*

185. *See* 18 U.S.C. § 1030(a)(2) (2012).

personation.”¹⁸⁶ Oklahoma’s statutes, for example, make it a crime to “falsely personate another” and perform certain actions in that “assumed character.”¹⁸⁷ One violates this law, when, while impersonating another person, one “[s]ubscribes, verifies, publishes, acknowledges or proves, in the name of another person, any written instrument, with intent that the same may be delivered or used as true.”¹⁸⁸ One likewise violates the law when one performs actions that would, if done by the victim of the impersonation, make the latter “liable to any suit or prosecution, or to pay any sum of money, or to incur any charge, forfeiture or penalty” or that brings “any benefit . . . to the party personating, or to any other person.”¹⁸⁹

It is unlikely any of these legal restrictions on impersonation or misuse of others’ identity would face First Amendment difficulties. As noted above, even the opinion in *Alvarez* that extended the strongest free speech protection to false statements of fact, that of Justice Kennedy, emphasized that “[w]here false claims are made to effect a fraud or secure moneys or other valuable considerations, say offers of employment, it is well established that the Government may restrict speech without affronting the First Amendment.”¹⁹⁰ The crime of false personation seems to fit squarely within this category of entirely unprotected lying. So too do all of the spear-phishing attacks I have described, as well as the analogues of such attacks that might be carried out with fabricated video- or audio-recordings.

The more difficult question is whether, and when, the law may *also* criminalize, or subject to civil liability, the use of such techniques to create fake news, to disguise content as expert knowledge when it is not, or to otherwise inject false content into public discourse. Should government have greater power to prevent such manipulation of communicative media or sources than it has, under the *Alvarez* framework, to punish false statements?

The proposal I wish to briefly consider here is that judges and scholars should be open to answering yes—at least in some circumstances. Even if the Justices were right to assume in *Alvarez* that the false *content* of a

186. See, e.g., CAL. PENAL CODE § 529(a) (West 2011); 21 OKLA. STAT. § 1531(4) (2011); N.Y. PENAL LAW § 190.25 (McKinney 2008).

187. 21 OKLA. STAT. § 1531(4).

188. *Id.*

189. *Id.*

190. *United States v. Alvarez*, 567 U.S. 709, 723 (2012).

statement about politics, history, or other matters of public concern should not by itself normally suffice to eliminate First Amendment protection, the rule might differ for methods that distort the medium that carries it. This is because, at least right now, we do not have the same repertoire of tools for dealing with falsified mediums or sources that we have for uncovering false claims. When confronted by a false statement from an individual or organization on Twitter or in a Facebook post, we can conceivably respond with skepticism—and then withhold acceptance of the claim until we see it confirmed by some more reliable source, such as multiple reports from professional journalists or video footage showing that the event described on Twitter or Facebook. By contrast, if this reliable journalistic or video check is itself rendered unreliable by widespread and easy-to-implement falsification methods, and particularly methods which (to quote Lawfare once again) make it impossible for an “unaided observer” to “detect the fake,” it is hard to see what basis of knowledge we will have to fall back on.

There is, of course, a counterargument: we can learn to bring the same skepticism to audio and video evidence that many people already show to verbal reports. Indeed, as some who make this counterargument point out, such increased skepticism has already arisen for photos. In response to concerns about malicious uses of Lyrebird’s technology, for instance, one representative of Lyrebird stated that just as people “are now aware that photos can be faked,” they will, in the future, regard “audio recording” as “less and less reliable.”¹⁹¹ If an unaided observer cannot detect that a deep fake video is fake, she might perhaps withhold judgment about its accuracy until she can obtain the aid needed to meet that challenge—such as the aid of an expert in forensic video analysis, or in fact-checking suspicious visual evidence.

Such a response, however, at the very least requires elaboration. It is true, of course, that in earlier times, when we lacked any technology for creating video and audio evidence, we had to find some way to form beliefs about faraway events without it. But it is not clear how well modern life can function in a situation where photographic evidence, audio-recording, and video-recording can be used not only to fabricate reality, but also to fabricate the other sources of indirect knowledge that we might use to test it against. If the absence of video and audio evidence

191. See Vincent, *supra* note 170.

leaves us to rely only on verbal reports, the problem is that these verbal reports can already very easily be designed to endorse false facts.

Such an environment could leave us not with a free market of ideas, but rather, as noted in the introduction, in a kind of Hobbesian informational “war of all against all”¹⁹² where any and every source of factual information might be a fake. This brings us at least a little bit closer to the kind of nightmarish dystopia described in this Essay’s introduction: the kind common in the genre of paranoid science fiction, where all of our perceptual experience might be a delusion.¹⁹³

To be sure, a world of forged photos, videos, and newspapers is not quite as unsettling as a world where human beings are themselves fake, or where the objects they see can and see and touch are illusions. But the informational anarchy and paranoia that could characterize such a world might at least present a serious (and possibly insuperable challenge) to those elements of individual decision making or collective self-rule that require us to trust some sources of external information.¹⁹⁴

Consequently, scholars and jurists should at least explore adding the following addendum to the framework from *United States v. Alvarez*: where false statements do not merely state false facts, but are also given in a form that 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. It should not, under this approach, always have to demonstrate that additional harm flows from the forgery of a video or faked news story from a major

192. THOMAS HOBBS, *LEVIATHAN* (Clarendon Press 1909) (1651).

193. 2 GARY WESTPHAL, *THE GREENWOOD ENCYCLOPEDIA OF SCIENCE FICTION AND FANTASY* 585 (2005).

194. There are, of course, conspiracy theorists in the United States who already believe that much of what is reported in the mainstream media consists of the kind of stage-managed falsity I have described above and that the stage manager is the government. They already believe that many of their fellow citizens are being manipulated and that many experts cannot be trusted. This belief has provided a willing audience for claims that events such as the San Bernardino and Las Vegas shootings and school shooting in Parkland, Florida, are government operations designed to manipulate audiences. While it is not the purpose of this Essay to discuss or analyze their views in depth, one thing that may make them more open to placing credence in tweets, Facebook posts, or comments in YouTube videos by individuals with such views is that they have already rejected the sources of knowledge—mainstream journalism or video and audio records—of events that, for others, continue to provide a reliable source of knowledge about the world, and continue to provide the basis for a consensus about much of the factual content that lies beyond each individual’s direct experience.

publication. The trick such an alteration plays on our senses and our ability to serve as watchmen should, in some cases, *by itself* be enough to act against the forgery.

This is already how the First Amendment applies where individuals imitate government officials—for example, in impersonating a police officer. Justice Kennedy acknowledged in *Alvarez* that laws may “prohibit impersonating a Government officer” and punish someone who engages in such impersonation even without a showing of “‘actual financial or property loss’ resulting from the deception.”¹⁹⁵ He argues that the First Amendment can allow such punishment of falsity in the absence of a showing of financial or other concrete harm because such impersonation threatens another kind of harm: undermining the dignity and good repute of government. But fabrication of video- or audio-recordings and of other social practices that allow us to identify reliable records of outside events can likewise do harm even where it does not cause financial or physical harm. It can undermine our ability to generate, and then draw upon, reliable sources of knowledge of the world beyond our direct perceptions.

This approach is necessarily tentative and comes with three major caveats. First, if it is to be an addendum to the *Alvarez* consensus among the Justices rather than a rejection of it, it cannot empower government to restrict the content of our communications in public discourse. Consistent with that consensus, individuals should still be able make whatever claims—even false ones—they wish to make about matters of public concern so long as they avoid doing so in ways that defame, defraud, or otherwise cause legally cognizable harm to others. What they should *not* have free rein to do under the cover of the First Amendment, under the approach considered here, is disguise the source of their claims in authoritative clothing by using technology such as video- or audio-fabrication or digital forgery to give it an appearance of reality that the expressive content alone cannot create. In other words, while the First Amendment gives someone the right in a discussion of public affairs to provide any answer they like—even a false one—to the question, “What should I believe?” and even to the follow-up question, “Why should I believe that?,” it doesn’t give them the right to answer the latter question by creating an illusion rather than an explanation.

195. *United States v. Alvarez*, 567 U.S. 709, 721 (2012).

This means that courts must somehow be able to mark a line between the expressive content itself and the features of that content's medium or source—and to assure that the expressive content remains free from government censorship in the absence of legally cognizable harm. But drawing such a line may be quite challenging. The analysis above has been built in part on the intuition that there is an important distinction in the way we perceive different kinds of evidence about events in the external world and, more specifically, that we tend to view video or audio footage as a form of indirectly perceiving an event for ourselves. What we see or hear in a recording is often treated by people not as a form of evidence that may be subject to distortion or manipulation, but as a window into reality. By contrast, we are more likely to view verbal reports as possibly mistaken or dishonest. But one might object to giving this intuitive difference constitutional significance. After all, in response to the question “why should I believe” a certain proposition, one might answer with verbal explanation rather than pointing to a purported video or audio recording. If, for instance, someone is asked to prove that a riot occurred in a certain city, he might respond not by showing video or audio of the riot, but rather by answering that he was there and providing a description of the experience that seems rich in detail. The proposal I am making here—that courts might bar individuals from creating their own indicia of reliability (by, for example, creating video evidence)—requires accepting First Amendment doctrine that allows restriction of one method of making the case for a belief (through altered video) but not another (by a verbal performance intended to give the false impression that certain memories and experiences actually occurred).

Drawing a line between source or medium and content is likely to be even more difficult where one turns from recording to other indicia of reliability. How, for example, should the First Amendment treat a fake, web-based newspaper that adopts the look and feel of a real newspaper but does so without using any real newspaper's trademark or trade dress? On the one hand, a website's or print publication's use of certain designs or names can effectively manipulate readers into assuming that the information is from a professional news organization that follows journalistic techniques for gathering and verifying information. On the other hand, the fake newspaper's design choices intuitively seem to be as much a part of that newspaper's expressive content as are the words or artistry on a book cover. Moreover, real journalists could conceivably find other, more foolproof ways of authenticating themselves to readers—for example, with professional certification standards. And this

authentication method, then, may be something that the government might prevent someone from forging.

Helen Norton has already analyzed a very similar challenge, exploring how the government can protect citizens from the confusion they might face “when private speakers seek the government’s perceived imprimatur to manipulate onlookers’ common—indeed, sometimes automatic—reliance on an idea’s source as the measure of its value.”¹⁹⁶ Her focus is on taking such concern into account and looking at a speech’s “source cues” to determine whether speech is private or government speech for purposes of the First Amendment.¹⁹⁷ But it also stresses the importance—to listeners and viewers—of being able to accurately identify speech that legitimately comes from government.¹⁹⁸

My major point in this subsection is a related one: our experience of the world—and our capacity to make confident judgments about it—requires that we have *some* sources of evidence that, as Norton puts it, allow us to place “common” and perhaps “automatic” “reliance on an idea’s source as the measure of its value.”¹⁹⁹ Moreover, where we do treat evidence that way, then First Amendment ground rules cannot insist that we be our own “watchmen for truth,” with all the uncertainty that entails: courts cannot demand that each of us show the kind of skeptical attitude and hesitation to accept a claim that is the hallmark of being a watchman with regard to evidence that we must be able to accept automatically and without anxiety about its truth. Courts already generally read the First Amendment as allowing government to help safeguard individuals’ reliance in certain contexts such as commercial or professional interactions. It may have some room to do so even outside of those social spheres where individuals need to take certain evidence on faith.

A second caveat is that even if courts can meet the difficult challenge of drawing a line between falsification of a speech source or medium and false speech content, they could not simply assume that the First Amendment offers no protection on the non-content side of this line. The video-altering technology that allows individuals to undermine each other’s grasp of what is real will likely have other, more benevolent uses. It might, for example, provide moviemakers with yet another tool to

196. Helen Norton, *The Measure of Government Speech: Identifying the Expression’s Source*, 88 B.U. L. REV. 587, 597 (2008).

197. *Id.* at 597–618.

198. *Id.* at 615.

199. *Id.*

create the special effects that can make narrative films feel real to an audience. Thus, the same technology that might lose First Amendment protection when it fabricates news might merit robust First Amendment protection when it is, like other tools of modern filmmaking, a means of telling a story. We thus appear to face, in addressing this technology, some of the same concerns that have pushed scholars like David Strauss to define the category of unprotected false statements “very narrowly.”²⁰⁰ Giving the government too much power to control how we use image-altering technology risks empowering it not only to prevent thorough deception, but also to restrict how we tell stories or otherwise express ourselves with technology.

A third caveat is that the proposal considered here is far from a complete antidote for all of the dangers that flow from what Eugene Volokh calls “cheap speech.”²⁰¹ In a 1995 article outlining the effects that the then-nascent internet might have on First Amendment activity, Volokh noted that while the internet would democratize speech and have other positive effects, “when speakers can communicate to the public directly, it’s possible their speech will be less trustworthy: they might not be willing to hire fact checkers, or might not be influenced enough by professional journalistic norms, or might not care enough about their long-term reputation for accuracy.”²⁰² More recently, Tim Wu has linked the rise of cheap speech to the listeners’ ability to sort quality speech from falsehoods, thus also giving rise to forms of private censorship by speakers themselves. Wu emphasizes that whereas “it was once hard to speak, it is now hard to be heard,” because the flood of information available on the internet makes it hard to capture listeners’ attention.²⁰³ Richard Hasen similarly argues that the internet-driven shift to cheap speech has caused a “collapse of traditional media” and “a rise in false news stories (‘fake news’) spread via social media.”²⁰⁴ Even if government can preserve the indicia of reliability of certain speech, that is no guarantee that listeners will notice the speech or do what they have to

200. Strauss, *supra* note 132, at 366.

201. Eugene Volokh, *Cheap Speech and What It Will Do*, 104 YALE L.J. 1805, 1807 (1995).

202. *Id.* at 1838.

203. Tim Wu, *Is the First Amendment Obsolete?* 7 (Columbia Pub. Law Research, Paper No. 14-573, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3096337.

204. Richard L. Hasen, *Cheap Speech and What It Has Done (to American Democracy)*, 18 FIRST AMEND. L. REV. 204–05 (2018).

do to locate it. Nor can any law assure that they will continue to use such indicia of reliability. Even if the First Amendment gives government leeway to stop individuals from exploiting and ultimately undermining the trust individuals place in certain media or sources of speech, it cannot force individuals to place more trust in professional journalists than in the wild speculations of conspiracy theorists.

V. Conclusion

The proposal I have just explored is that, by giving government more leeway to protect the marks of reliability in certain speech, the First Amendment can allow government to aid in preserving the value of social practices, like the practice of journalism or of scientific disciplines, that are themselves independent of government. What the government should be allowed by the First Amendment to do on this approach, in other words, is not take a leading role in sorting truth from falsehood. That is precisely what Supreme Court has said the Constitution cannot trust government to do. Rather, instead of replacing such social practices with its own truth-sorting mechanism, government's role should be to preserve those social practices and technological possibilities that have already evolved and function free of government direction or interference. As I described above, such sorting is already done, though less dangerously, by certain social practices that individuals inevitably rely upon to form reliable knowledge about the world that lies beyond their direct perception—that is, it is already done by a knowledge ecosystem. If, even with the help of government defense against attacks of the sort made by digital impostors and forgery artists, this constellation of social practices and technology still finds itself collapsing, the solution will lie not in First Amendment law (or constitutional law of any sort), but in other measures that might entail the emergence of new social practices and technologies.

At a high level of generality, what the First Amendment may have to allow room for is similar to the role played by the Digital Millennium Copyright Act (DMCA) in the realm of copyright law.²⁰⁵ Congress did not invent the copyright protection technologies that various rights-holders have generated to protect movies, for example, against the enhanced threat the computers have created to copyright (given the ease with which one can copy and disseminate a digital movie). What it did

205. See Digital Millennium Copyright Act, Pub. L. No. 105-304, sec. 103, § 1201, 112 Stat. 2860, 2863 (1998).

instead is let private parties create such technology themselves, and then legally shield the technology from those who would circumvent it.²⁰⁶ Although commentators have understandably objected to the high, often insuperable, barriers placed by the DMCA in the way of individuals who wish to make legal and “fair use” of movies or music,²⁰⁷ the general template is a common one and one that might have application in First Amendment law: when law establishes and protect rights, it does so in conjunction with other practices that make such rights possible, such as social norms or physical constraints. In fact, legal protection of rights is often most effective when it adds its protection to freedom-supporting features of the natural or social world or to technologies invented by private parties. Here, too, this may be a model that post-*Alvarez* First Amendment doctrine on false statements should allow room for. Conceding that it is not for law alone to save people from being deceived by fake news, judges might at least interpret the First Amendment to let government support the social practices and practical realities that do so (or may evolve to do so in the future).

On the other hand, such a project also carries with it the same kinds of risks that have led the Court to hesitate before excluding false statements from the First Amendment’s scope. Just as allowing government to censor damaging lies might simultaneously empower it to exercise a more far-reaching power over public discourse, allowing it to restrict the way individuals use video-altering or other source-imitating technology may give it more power than it should have to control the way individuals express themselves (artistically or otherwise) with emerging technologies. The question raised by deep fakes and similar technology, then, is whether First Amendment law can leave government with room to protect the social foundations that allow individuals to serve as their own “watchmen for truth” without simultaneously inviting officials to control and restrict how they play that role.

206. See Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 519 (1999).

207. *Id.*