COMMENT

THE DEEPFAKE DILEMMA: RECONCILING PRIVACY AND FIRST AMENDMENT PROTECTIONS

Shannon Reid[*]

INTRODUCTION

Deepfakes are realistic videos created using artificial intelligence software to replace the face of one person with the face of another.[1] The technology used to produce these fake videos or digital representations is becoming increasingly sophisticated and available to the masses.[2] Since their creation by an anonymous Reddit user in late 2017,[3] deepfakes have challenged the effectiveness of U.S. law at punishing those who publish deepfakes of others without their consent. While deepfakes are often utilized as comedic or

---

* Shannon Reid is a 2020 graduate of the University of Pennsylvania Law School. She focuses her practice on labor and employment law and has legal research experience in the areas of privacy and government investigations.

1 *See* Sierra Lyda, *"Deepfakes" Technology and Pornography Laws*, N.C. J.L. & TECH. (Feb. 14, 2018), htttp://ncjolt.org/deepfakes-technology-pornography-laws/ (describing the open-source "machine learning" technology used to create deepfakes); *see also* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1757 (2019) (defining deep fakes as "the full range of hyper-realistic digital falsification of images, video, and audio").

2 *See* Grace Shao & Evelyn Cheng, *The Chinese Face-Swapping App That Went Viral Is Taking the Danger of 'Deepfake' to the Masses*, CNBC (Jan. 17, 2020, 2:50 AM), https://www.cnbc.com/2019/09/04/chinese-face-swapping-app-zao-takes-dangers-of-deepfake-to-the-masses.html (discussing how "[a] face-swapping app that surged to the top of China's domestic download rankings has raised concerns about how fabricated but realistic-looking videos may be breaking into the mainstream").

3 *See* Samantha Cole, *AI-Assisted Fake Porn is Here and We're All Fucked*, VICE (Dec. 11, 2017, 2:18 PM), https://www.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn (discussing deepfake technology with the creator); *see also* Meredith Somers, *Deepfakes, Explained*, MIT SLOAN (July 21, 2020), https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained#:~:text=The%20term%20%E2%80%9Cdeepfake%E2%80%9D%20was%20first,open%20source%20face%2Dswapping%20technology ("The term 'deepfake' was first coined in late 2017 by a Reddit user of the same name[.]").

satirical tools,[4] around the world, deepfakes are also being used to humiliate and harass individuals. The latter, more harmful use leads to detrimental consequences for those targeted.

Rana Ayyub, an investigative journalist in India, became a target of this practice when a deepfake sex video showing her face on another woman's body was circulated on the Internet in April 2018.[5] It was spread via Facebook, Twitter, and WhatsApp, and it was sometimes sent with rape threats and her home address.[6] Ayyub stated that she "endured online harassment for years[,]" but she found the deepfake "uniquely visceral, invasive and cruel."[7] Ayyub threw up when she saw the video, cried for days afterward, and was ultimately rushed to the hospital, "overwhelmed with anxiety."[8] In reflecting upon the physical, mental, and emotional harm the video caused her, she stated that the video "is a lot more intimidating than a physical threat. [It] has a lasting impact on your mind. And there's nothing that could prevent it from happening to me again."[9]

Unfortunately, Rana Ayyub is not alone. Celebrities, such as Scarlett Johansson, are often targets of deepfake creators.[10] The technology is being used to "take the face of one real person (like a celebrity) and splice it onto the body of another (like a porn star), creating videos that lack the consent of multiple parties."[11] In addition, political figures like Barack Obama are depicted in incredibly realistic deepfakes speaking on important issues that

---

4 *See* Casey Newton, *Facebook's Deepfakes Ban Has Some Obvious Workarounds*, VERGE (Jan. 8, 2020, 6:00 AM), https://www.theverge.com/interface/2020/1/8/21054906/facebook-deepfakes-ban-loopholes-parody-satire-cheap-fakes (discussing how deepfakes created for parody and satire are easily spread on the social media network despite a general ban against deepfakes to protect those that did not consent to the use of their image in the creations).

5 Drew Harwell, *Fake-Porn Videos are Being Weaponized to Harass and Humiliate Women: 'Everybody Is a Potential Target,'* WASH. POST (Dec. 30, 2018, 10:00 AM), https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/ (describing the plight of Ayyub after she was featured in a deepfake without her consent).

6 *Id.*

7 *Id.*

8 *Id.*

9 *Id.*

10 *See id.* (discussing the plight of celebrities in battling fake online videos).

11 David Greene, *We Don't Need New Laws for Faked Videos, We Already Have Them*, ELEC. FRONTIER FOUND. (Feb. 13, 2018), https://www.eff.org/deeplinks/2018/02/we-dont-need-new-laws-faked-videos-we-already-have-them.

could easily fool the unsuspecting viewer.[12] If some of the most famous and resourceful in American society are being taken advantage of without recourse, harm to the average U.S. citizen could be exceptional.

Moreover, the proliferation of deepfakes that spread disinformation on matters of public interest could cripple public discourse, and, as a result, undermine democracy.[13] There are several factors that increase the threat of deepfake disinformation in the political context. These factors include the tendency of humans, as social creatures, to be attracted to the shocking material that is often included in deepfakes, which drives larger audiences and facilitates dissemination.[14] There is also the fear that hysteria over fake videos could lead people to deny legitimate video evidence or overwhelm people to the point of "reality apathy[,]" in which one rejects all video evidence as unreliable and maintains their previous position or affiliation.[15] Thus, in addition to the private harms resulting from deepfakes, the deepfake technology could lead to serious public harms as well.[16]

For these reasons, a comprehensive reevaluation of U.S. law is needed to identify opportunities to strengthen the privacy protections available to those victimized by this rapidly advancing technology. Currently, the First Amendment is a significant challenge to the U.S. government's ability to regulate deepfakes because of First Amendment restrictions on limiting free speech.[17] Relatedly, deepfake creators often have a First Amendment defense in civil claims against them. This Comment suggests that federal courts could better balance First Amendment interests and privacy protections by

---

[12] *See* Edward Lee, *Can the U.S. Government Prohibit Deepfake Videos Intended to Deceive Voters?*, FREE INTERNET PROJECT (Feb. 15, 2019), https://thefreeinternetproject.org/blog/can-us-government-prohibit-deepfake-videos-intended-deceive-voters (discussing Jordan Peele's deepfake of Barack Obama and the danger of deepfakes).

[13] *See* Amy Mitchell, Jeffrey Gottfried, Galen Stocking, Mason Walker, & Sophia Fedeli, *Many Americans Say Made-Up News Is a Critical Problem That Needs To Be Fixed*, PEW RSCH. CTR. (June 5, 2019), https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/ (reporting that "[m]any Americans say the creation and spread of made-up news and information is causing significant harm to the nation" and makes it difficult to discern the basic facts of current events).

[14] *See* Drew Harwell, *Top AI researchers race to detect 'deepfake' videos: 'We are outgunned,'* WASH. POST (June 12, 2019, 4:44 PM), https://www.washingtonpost.com/technology/2019/06/12/top-ai-researchers-race-detect-deepfake-videos-we-are-outgunned/ (discussing how the "reward structure of the modern Web" can erode confidence in information published online).

[15] *Id.*

[16] *See* Chesney & Citron, *supra* note 1, at 1768–84 (detailing harms including, but not limited to, exploitation, sabotage, and the undermining of public safety, diplomacy, and democratic discourse).

[17] *See infra* Part II.

recognizing a constitutional right to defend one's personality and reputation via the Fourteenth Amendment right of autonomy. It begins by giving an overview of deepfake technology and U.S. law's shortcomings in deterring the use of nonconsensual deepfake videos. It then discusses the historical link between U.S. common law privacy protections and the more legally authoritative right to personality under German constitutional law. It concludes that the two privacy regimes share historical origins in protecting human dignity, which includes reputation and autonomy; and therefore, U.S. courts could reasonably infer a constitutional privacy right that would rival the First Amendment in legal actions against deepfake creators.

I. DEEPFAKE TECHNOLOGY

Deepfakes are "manipulated videos, or other digital representations produced by sophisticated artificial intelligence, that yield fabricated images and sounds that appear to be real."[18] Today, artificial intelligence or "AI" refers to "machines that respond to stimulation consistent with traditional responses from humans, given the human capacity for contemplation, judgment, and intention."[19] Deepfake technology utilizes "deep learning," a subset of artificial intelligence that involves the "arrangements of algorithms that can learn and make intelligent decisions on their own."[20] "A deep-learning system can produce a persuasive counterfeit by studying photographs and videos of a target person from multiple angles, and then mimicking their person's behavior and speech patterns."[21] Once a "preliminary fake" has been produced, a method known as generative adversarial networks (GANs), detects flaws in the forgery and improves them, making the fake more

---

[18] Grace Shao, *What 'Deepfakes' Are and How They May Be Dangerous*, CNBC (Jan. 17, 2020, 2:47 AM), https://www.cnbc.com/2019/10/14/what-is-deepfake-and-how-it-might-be-dangerous.html.

[19] Darrell M. West, *What Is Artificial Intelligence?*, BROOKINGS (Oct. 4, 2018), https://www.brookings.edu/research/what-is-artificial-intelligence/ (citing Shukla Shubhendu S. & Jaiswal Vijay, *Applicability of Artificial Intelligence in Different Fields of Life*, 1 INT'L J. SCI. ENG'G & RSCH. 28, 28 (2013)).

[20] Shao, *supra* note 18.

[21] *Id.*

believable.[22] Individuals can initiate and manage these deep learning processes through software generally known as a "neural network."[23]

While deep-learning is a significant part of how all deepfakes are created, there are several ways in which deepfake creators can engage with the software to manipulate images and video of the individuals they are targeting. While some older versions of deepfake software can create convincing deepfakes by analyzing only a few still photos of an individual's face, newer software can change what someone appears to be saying on video just by editing the text in the audio-video transcript.[24] Thus, as deepfake technology advances, it becomes easier for its users to manipulate digital media.[25]

Once a deepfake is created, it can be disseminated easily on the Internet and is virtually impossible to retract. Major social media and video platforms, such as Facebook, Twitter, and YouTube, enable deepfake creators to upload fake videos and images for rapid dissemination to millions of other media users around the world.[26] The U.S. government has sponsored "more than a dozen academic and corporate groups" that are conducting research on how to detect deepfakes, but researchers say they remain "vastly overwhelmed by a technology that they fear could herald a damaging new wave of disinformation campaigns[.]"[27] A major barrier to the development of effective state-sponsored deepfake detection and removal tools is an asymmetry of resources and time. According to Hany Farid, a computer-science professor and digital-forensics expert at the University of California at Berkeley, "[w]e are outgunned . . . [t]he number of people working on the video-synthesis side, as opposed to the detector side, is 100 to 1."[28] Because of this resource difficulty, top political officials are hoping that social networks and video sites

---

[22] *Id.* (citation omitted).

[23] *What Is a Deepfake?*, ECONOMIST: THE ECONOMIST EXPLAINS (Aug. 7, 2019), https://www.economist.com/the-economist-explains/2019/08/07/what-is-a-deepfake.

[24] *See* Harwell, *supra* note 14 (citing Ohad Fried et al., *Text-based Editing of Talking-head Video*, *in* 38 ACM TRANSACTIONS ON GRAPHICS 1, 2 (2019)).

[25] For a more technical explanation of deepfake technology, see Alan Zucconi, *Understanding the Technology Behind DeepFakes*, (Mar. 14, 2018), https://www.alanzucconi.com/2018/03/14/understanding-the-technology-behind-deepfakes/ [https://perma.cc/L7UE-2WWF].

[26] *See* Harwell, *supra* note 14 (discussing the government's push to get major media platforms, such as Facebook, Twitter, and YouTube, to regulate the presence of deepfakes on their sites to mitigate the spread of malicious deepfakes online).

[27] *Id.*

[28] *Id.*

will find and remove the worst fakes.[29] Major tech companies, however, have differing policies on takedowns, and some do not require that uploaded videos be true.[30] In addition, a government approach that necessitates sharing deepfake detection software with media outlets and other parties could potentially enable deepfake creators to "examine the code and find workarounds."[31] For these reasons, the rapidly advancing nature of deepfakes makes offensive use of the technology extremely difficult to combat through physical or technical means.

II. THE LACK OF SUFFICIENT PROTECTIONS AGAINST DEEPFAKES

Unfortunately, U.S. federal and state law fails to provide a sufficient remedy for those targeted in deepfake productions. Privacy laws fall significantly short of targeting the technologies and behaviors posing the greatest threats.[32] And while several federal criminal and intellectual property statutes appear to address the challenges presented by deepfakes, like U.S. privacy law, they are narrowly applied by the courts or vulnerable to defenses that nearly extinguish legal remedies for potential plaintiffs.

A. Tort Law

While civil privacy violations are often remedied under state tort law, the First Amendment is a significant barrier to redress, especially if a deepfake involves a public figure or public matter. In *Pavesich v. New England*, the Supreme Court of Georgia held that:

---

[29] *See id.* (explaining how technical and personnel difficulties leave government officials largely unprepared to handle a large deepfake disinformation campaign or otherwise prevent malicious deepfakes from spreading on the Internet).

[30] *See* Allyson Chiu, *Facebook Wouldn't Delete an Altered Video of Nancy Pelosi. What About One of Mark Zuckerberg?*, WASH. POST (June 12, 2019, 6:23 AM), https://www.washingtonpost.com/nation/2019/06/12/mark-zuckerberg-deepfake-facebook-instagram-nancy-pelosi/ (discussing Facebook's takedown policy).

[31] Harwell, *supra* note 14.

[32] *See* Steven Chabinsky & F. Paul Pittman, *USA: Data Protection Laws and Regulations 2020*, ICLG.COM (June 6, 2020), https://iclg.com/practice-areas/data-protection-laws-and-regulations/usa (detailing an exhaustive list of federal privacy protections provided by, for example, the Health Insurance Portability and Accountability Act ("HIPAA"), Fair Credit and Reporting Act ("FCRA"), Gramm-Leach-Bliley Act ("GLBA"), and Family Educational Rights and Privacy Act ("FERPA"), but failing to document explicit federal protection for privacy invasions involving deepfakes or artificial intelligence generally).

> [t]he right of privacy, or the right of the individual to be let alone, is a personal right . . . . It is the complement of the right to the immunity of one's person. The individual has always been entitled to be protected in the exclusive use and enjoyment of that which is his own.[33]

The court found that "a violation of the right of privacy is a direct invasion of a legal right of the individual. It is a tort, and it is not necessary that special damages should have accrued from its violation in order to entitle the aggrieved party to recover."[34]

In 1960, William Lloyd Prosser drew upon all post-*Pavesich* state law cases and defined four distinct invasion of privacy torts: intrusion upon seclusion, publicity given to private life, false light publicity, and appropriation for advantage ("The Four Prosser Torts").[35] The first two of the Prosser Torts, "intrusion upon seclusion" and "publicity given to private life," are the least applicable to harms suffered by deepfake targets.[36] The remaining two torts, "appropriation of name or likeness" and "false light publicity," on their face, appear applicable to the privacy threat presented by deepfakes. The former is committed when "[o]ne . . . appropriates to his own use or benefit the name or likeness of another[,]"[37] while the latter occurs when "[o]ne . . . gives publicity to a matter concerning another that places the other before the public in a false light[.]"[38] A person is subject to false light liability if: "(a) the false light in which the other was placed would be highly offensive to a reasonable person, and (b) the actor had knowledge of or acted in reckless disregard as to the falsity of the publicized matter and the false light in which the other would

---

[33] Pavesich v. New England Life Ins. Co., 50 S.E. 68, 78 (Ga. 1905) (quoting Roberson v. Rochester Folding Box Co., 64 N.E. 442, 449 (N.Y. 1902) (Gray, J., dissenting)).

[34] *Id.* at 73 (citation omitted). The court was reviewing a Georgia statute and held that the publication of plaintiff's picture without consent was an invasion of that individual's right to privacy. *Id.*

[35] RESTATEMENT (SECOND) OF TORTS, § 652B-E (AM. L. INST. 1977). The Second Restatement was published after Prosser's death in 1972, but was based on his work in 1960. William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383 (1960).

[36] An "intrusion upon seclusion" committed when "[o]ne who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns . . . , if the intrusion would be highly offensive to a reasonable person." RESTATEMENT (SECOND) OF TORTS, *supra* note 35, at § 652B. "Publicity given to private life" involves "[o]ne who gives publicity to a matter concerning the private life of another . . . , if the matter publicized is of a kind that (a) would be highly offensive to a reasonable person, and (b) is not of legitimate concern to the public." *Id.* § 652D.

[37] *Id.* § 652C.

[38] *Id.* § 652E.

be placed."[39] Before bringing a claim for any of these torts, a plaintiff must show injury in fact that is particularized (specific to them) and concrete (an actual harm or risk of harm); the injury must also be traceable to the defendant's conduct and redressable.[40]

The problem with these privacy tort protections, however, is that they are vulnerable to a First Amendment defense that deepfakes are protected speech, rendering them potentially useless against the unique privacy harms presented by deepfakes.[41] Under the First Amendment, "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances."[42] In *New York Times Co. v. Sullivan*, the Supreme Court recognized the right to free speech, specifically the public's interest in freedom of expression in the area of public matters and public figures, as a defense in tort actions:

> "Although this is a civil lawsuit between private parties, the Alabama courts have applied a state rule of law which petitioners claim to impose invalid restrictions on their constitutional rights of speech and press. It matters not that that law has been applied in a civil action and that it is common law only . . . . The test is not the form in which state power has been applied but, whatever the form, whether such power has in fact been exercised."[43]

After *Sullivan*, the First Amendment became a go-to defense against tort actions that penalized free speech.[44] In addition to those who make tortious statements, parties that disseminate false information, such as news outlets,

---

[39] *Id.*

[40] *See* Spokeo, Inc. v. Robins, 136 S. Ct. 1540, 1548 (2016) (further defining standing requirements for privacy torts as requiring that they be plead with particularity and describe concrete harms).

[41] *See* George C. Christie, *The Uneasy and Often Unhelpful Interaction of Tort Law and Constitutional Law in First Amendment Litigation*, 98 MARQ. L. REV. 1003, 1003–04 (2015) (discussing "increasing tensions between the First Amendment and the common law torts of intentional infliction of emotional distress, defamation, and privacy"); *see also* Anita L. Allen, *Privacy Torts: Unreliable Remedies for LGBT Plaintiffs*, 98 CALIF. L. REV. 1711, 1711–14 (2010) (arguing that Prosser's Four Privacy Torts are weak in practice and do not effectively remedy harms for a significant portion of the U.S. citizenry, most specifically LGBTQ individuals).

[42] U.S. CONST. amend I.

[43] N. Y. Times Co. v. Sullivan, 376 U.S. 254, 265 (1964) (citations omitted).

[44] *See* James M. Beck, *How the First Amendment Affects Tort Law*, LEXOLOGY (Dec. 14, 2012), https://www.lexology.com/library/detail.aspx?g=1b0cef82-2385-42a0-9631-8a3846724833 (explaining how the First Amendment and other federal law often poses a viable defense to any tort claim that undermines free speech).

can raise a First Amendment defense to defeat liability.[45] Courts consider the act of awarding money damages for a tort claim involving free speech or expression a "direct regulation on speech."[46] In *United States v. Alvarez*, the Supreme Court explained that

> 'the First Amendment means that [the] government has no power to restrict expression because of its message, its ideas, its subject matter, or its content." As a result, the Constitution "demands that content-based restrictions on speech be presumed invalid . . . and that the Government bear the burden of showing their constitutionality.'"[47]

The defense extends to other torts, such as defamation (libel and slander) and intentional infliction of emotional distress, which could otherwise offer plaintiffs some redress.[48] For example, the *Sullivan* case involved libel or written defamation.[49] By extending First Amendment protection to libel, the *Sullivan* court overturned nearly 200 years of precedent holding that defamation is not protected under the First Amendment.[50] Moreover, in *Hustler Magazine, Inc. v. Falwell*, the Supreme Court employed the *Sullivan* standard to defeat liability in an infliction of emotional distress claim.[51] In *Hustler Magazine*, an advertisement "parody" described a well-known minister as losing his virginity "during a drunken incestuous rendezvous with his mother in an outhouse."[52] The Court concluded that additional proof of

---

[45] *See* Phila. Newspapers, Inc. v. Hepps, 475 U.S. 767, 776–78 (1986) (discussing how the First Amendment abrogates the common law presumption that defamatory public speech regarding private persons is false).

[46] *See, e.g.*, *In re* Orthopedic Bone Screw Prods. Liab. Litig., 193 F.3d 781, 792 (3d Cir. 1999) (discussing the negative impact on free speech if protected expression could lead to civil liability and damages).

[47] United States v. Alvarez, 567 U.S. 709, 716–17 (2012) (alteration in original) (quoting Ashcroft v. Am. Civ. Liberties Union, 542 U.S. 656, 573, 660 (2004)).

[48] A statement is defamatory if it "tends so to harm the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him." RESTATEMENT (SECOND) OF TORTS, *supra* note 35, at § 559. Infliction of emotional distress claims require a showing that the defendant's conduct was "extreme" or "outrageous." *Id.* § 46. For a more comprehensive review of cases recognizing the First Amendment defense in various forms of tort litigation, see *supra* note 44.

[49] *See* N.Y. Times Co. v. Sullivan, 376 U.S. 254, 256–57 (1964) (explaining the elements of libel via case-specific facts).

[50] *See id.* at 264–65 (holding that the First Amendment can be a viable defense to a libel claim involving matters of public concern).

[51] Hustler Mag., Inc. v. Falwell, 485 U.S. 46, 50–52 (1988).

[52] *Id.* at 48.

falsity with actual malice was necessary "to give adequate 'breathing space' to the freedoms protected by the First Amendment" when the speech involves public figures or matters of public concern.[53]

The actual malice standard can make it incredibly difficult for a plaintiff in a tort case involving a public figure or public matter, even if the case is between two private parties.[54] A plaintiff can establish "actual malice" by proving that the defendant knew or acted with reckless disregard as to the falsity of a statement.[55] But this assumes that a plaintiff can identity the deepfake creator. Absent an identifiable defendant, a plaintiff may only be able to pursue a third-party disseminator of the deepfake (i.e. a news agency or social media platform). A third party that disseminates defamatory information, however, may not have actual knowledge or reckless disregard for a statement's falsity, which may become increasingly harder to prove as deepfakes become more convincing.[56] In addition, courts have yet to thoroughly address what could constitute a "matter of public concern" or a "public figure" in the age of social media and deepfakes. For example, someone who a decade ago may have been considered a private figure may, in 2020, have thousands of followers and subscribers (both nationally and internationally) on social media or YouTube, respectively. Would they be a public figure? Would information about them be a matter of public concern? Right now, a public figure can be someone who merely thrusts themselves into the public eye and invites comment or criticism.[57] If and how this definition evolves upon the advent of

---

[53] *Id.* at 53, 56.

[54] *See N.Y. Times Co.*, 376 U.S. at 279–80 (establishing the First Amendment defense to liability and "actual malice" standard regarding a public figure and a private newspaper).

[55] *See Hustler Mag., Inc.*, 485 U.S. at 53, 56. The "actual malice" standard (requiring proof that the defendant knew or acted with reckless disregard as to the falsity of a statement) may be easily proven against a deepfake creator that intentionally alters the original or true depictions of a private party. *Id. But see infra* note 56 (explaining how third parties that contribute to the spread of the defamatory information are unlikely to be liable as deepfakes become more convincing).

[56] This is especially likely considering that online media platforms with high content volume cannot remove defamatory information with the discernment required under the First Amendment, even when they are notified about a statement's falsity. *See, e.g.*, Zeran v. Am. Online, Inc., 129 F.3d 327, 333 (4th Cir. 1997) (referencing First Amendment concerns in broadly interpreting a federal statutory provision that immunizes online service providers from tort liability for information posted on their websites by third parties).

[57] *See* Gertz v. Robert Welch, Inc., 418 U.S. 323, 345, 351 (1974) (holding that:

> [T]hose classed as public figures have thrust themselves to the forefront of particular public controversies in order to influence the resolution of the issues involved. An individual may achieve such pervasive fame or notoriety that he becomes a public figure for all purposes

more convincing deepfake technologies could severely impact tort remedies for deepfake targets.[58]

Moreover, a plaintiff's ability to engage in counter speech could undermine their tort claim.[59] While this expectation is imposed against governments more than often than private parties,[60] it sets a discouraging precedent for individuals defamed in deepfake videos. It overestimates the power of counter speech, especially considering the realistic and pervasive nature of deepfake technology. American comedian and director, Jordan Peele, created a deepfake of former U.S. President, Barack Obama, to demonstrate how easily deepfakes can be believed and disseminated on the Internet.[61] Global social media platforms like Twitter, Instagram, YouTube, and Facebook serve as ultra-efficient channels for deepfake creators to share their deepfakes with millions of individuals worldwide.[62] For these reasons, courts' consideration of counter speech, even regarding public figures and governments, significantly undermines tort remedies. This is especially true when considering that, even if an individual is successful in proving that a video

---

> and in all contexts. More commonly, an individual voluntarily injects himself or is drawn into a particular public controversy and thereby becomes a public figure for a limited range of issues.).

58 *See* Ellyn M. Angelotti, *Twibel Law: What Defamation and Its Remedies Look Like in the Age of Twitter*, 13 J. HIGH TECH. L. 430, 432 (2013) (discussing unanswered questions on how the public figure and actual malice doctrines apply in the age of social media, specifically Twitter); *see also* Deven R. Desai, *Speech, Citizenry, and the Market: A Corporate Public Figure Doctrine*, 98 MINN. L. REV. 455, 456–59 (2013) (explaining the public figure doctrine as it relates to corporations and advocating for a specific corporate public figure doctrine).

59 *See* United States v. Alvarez, 567 U.S. 709, 726 (2012) (striking down a federal statute against false speech because, in part, the government failed to show why counter speech or refutation could not "overcome the lie"); *see also N.Y. Times Co*, 376 U.S. at 278–79 (holding that false speech enjoys constitutional protection insofar as its prohibition would chill truthful speech).

60 *See Alvarez*, 567 U.S. at 726–28 (discussing the government's ability to disprove whether an individual has won a military medal of honor).

61 Kaylee Fagan, *A Viral Video That Appears to Show Obama Calling Trump a 'dips--' Shows a Disturbing New Trend Called 'Deepfakes,'* BUS. INSIDER (Apr. 17, 2018, 4:48 PM) https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4.

62 *See, e.g.*, *How is Fake News Spread? Bots, People like You, Trolls, and Microtargeting*, CTR. FOR INFO. TECH. & SOC'Y, https://www.cits.ucsb.edu/fake-news/spread (last visited Oct. 31, 2020) (explaining how "bots," microtargeting, trolls, and people in general are used to spread fake content on social media).

is fake (i.e. the case of Indian journalist, Rana Ayyub), the reputational, emotional, and economic harm may already be done.[63]

For these reasons, the First Amendment is widely invoked to protect speech, authors who create it, and entities that disseminate it, from tort liability for injuries allegedly caused by such ideas.[64] Those same protections, however, can nearly extinguish tort remedies that would otherwise offer victims of injurious deepfakes at least some redress for their harms.

B. Copyright Law

Copyright law also fails to provide a remedy for those victimized by deepfakes because deepfake creators are likely to succeed in arguing a First Amendment fair use defense to copyright protection. Article I, Section 8, Clause 8 of the Constitution secures for a limited time "to [a]uthors and [i]nventors the exclusive [r]ight to their respective [w]ritings and [d]iscoveries[.]"[65] If a plaintiff captured photos or video that was later transposed into a deepfake, they would have copyright ownership of the original media.[66]

A deepfake, however, is likely a transformative work or parody under Section 107 of the Copyright Act, rending a plaintiff unlikely to defeat a fair use defense. Fair use of a copyrighted work is allowed for purposes such as criticism, comment, news reporting, teaching, scholarship, or research.[67] In determining whether use in any case is fair, the courts must consider the following factors:

> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> (2) the nature of the copyrighted work;
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

---

63 *See* Harwell, *supra* note 5 (describing emotional and physical distress caused by a pornographic deepfake).

64 *See* Beck, *supra* note 44 (arguing that the First Amendment defense is widely employed in tort law).

65 U.S. CONST. art. I, § 8, cl. 8. (commonly referred to as the "Copyright Clause").

66 Douglas Harris, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 107 n.57 (2019).

67 17 U.S.C. § 107 (2018).

> (4) the effect of the use upon the potential market for or value of the copyrighted work.[68]

The U.S. Supreme Court has made clear that the fourth factor is the most important factor in the fair use analysis.[69] The fair use analysis involves a "sensitive balancing of interests" between factors, requiring that the factors be weighted together.[70] Thus, the first and fourth factors are often combined to evaluate the commerciality of a creator's work and whether it interferes with the market value of the original.[71] But under factor three, "the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use."[72] Therefore, even if a deepfake creator is selling a nonconsensual deepfake containing copyrighted images, the creator would still be likely to succeed in a fair use defense if the deepfake is transformative under the third factor.

A new work that merely copies the essential physical or nonphysical qualities of an original work, copies the heart of the work and is not transformative.[73] Parodies are transformative by humorously criticizing a former work and providing some social benefit.[74] While deepfakes concerning public figures or matters of public concern are likely to be considered parodies, pornographic deepfakes are not for lack of commentary on an underlying work or expression.[75] A new work can still be transformative, however, even if it is not a parody. Transforming an original work into something new with different purpose or character can satisfy the third element.[76] Unfortunately for plaintiffs, deepfakes are the epitome of using a

---

[68] *Id.*

[69] *See* Harper & Row, Publishers. v. Nation Enters., 471 U.S. 539, 566 (1985) (stating that "the Act focuses on 'the effect of the use upon the potential market for or value of the copyrighted work[,]' the last factor, is undoubtedly the single most important element of fair use" (footnote omitted)).

[70] Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 449 n.32, 455 n.40 (1984).

[71] *Harper & Row*, 471 U.S. 539 at 602-03 (1985).

[72] Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994).

[73] *See* Chi. Bd. of Educ. v. Substance, Inc., 354 F.3d 624, 627–30 (7th Cir. 2003) (holding that a schoolteacher's publishing of six full-length, copyrighted standardized exams to critique them used too much of each exam).

[74] *See Campbell*, 510 U.S. at 580 (defining parody and distinguishing it from mere mockery).

[75] *Compare* Harris, *supra* note 66, at 108–09 (detailing why "personal" pornographic deepfakes are likely unprotected by parody law), *with* Jessica Ice, Note, *Defamatory Political Deepfakes and the First Amendment*, 70 CASE W. RSRV. L. REV. 417, 435 (2019) (discussing the vitality of parody as a defense to deepfake creators using public figures or commenting on public matters).

[76] *See Campbell*, 510 U.S. at 587–88.

previous copyrighted work (i.e. a picture or video) and transforming it into something wildly different from the purpose or character of the original work.[77] Therefore, copyright law is not a promising source of redress for those victimized by nonconsensual deepfake creations.

C. Right of Publicity Law

In addition to the challenges First Amendment protections pose to individuals seeking redress under state tort law, right of publicity laws require showings that make it difficult for plaintiffs to recover for the nonconsensual use of their image or likeness. The right of publicity protects the economic value of a person's likeness or image; it is very broad and generally extends past death.[78] Right of publicity laws are rooted in state-level privacy laws, but they are often considered an intellectual property protection because of their focus on commerciality.[79] The tension between classifying the right of publicity as a property or privacy protection makes it difficult for lawmakers to draft strong right of publicity protections.[80] And, like the privacy torts discussed above, they are only recognized on the state, not federal, level.[81]

While a legitimate right of publicity claim can defeat a fair use defense,[82] requiring a plaintiff to prove the economic value of their image to secure a remedy is a significant barrier to non-famous plaintiffs. For example, in *White v. Samsung*, Vanna White sued Samsung for an advertisement depicting a robot resembling the "Wheel of Fortune" host.[83] White was ultimately unsuccessful on her right to privacy claim under California law, but she

---

[77] *See* Harwell, *supra* note 5 (noting that deepfakes "are effectively new creations, meaning they could be protected as free speech").

[78] J. THOMAS MCCARTHY & ROGER E. SCHECHTER, THE RIGHTS OF PUBLICITY AND PRIVACY, § 1:2 (2d ed. 2020).

[79] *See id.* (footnote omitted) ("The right of publicity grew historically out of the state law right of privacy. Today, all states recognize some aspect of the right of privacy, either at common law or by statute.").

[80] Joshua L. Simmons & Miranda D. Means, *Split Personality: Constructing a Coherent Right of Publicity Statute*, A.B.A. (2018), https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2017-18/may-june/split-personality/.

[81] *Id.*

[82] Many states require the plaintiff to prove that a defendant's alleged use is not transformative of the plaintiff's likeness or image. *See, e.g.*, *In re* NCAA Student-Athlete Name & Likeness Licensing Litig., 724 F.3d 1268, 1276 (9th Cir. 2013) (holding that a video game developer's use of the likenesses of college athletes in its video games was not transformative and thus not protected by the First Amendment).

[83] White v. Samsung Elecs. Am., Inc., 971 F.2d 1395, 1396 (9th Cir. 1992).

succeeded after appeal on her right of publicity claim.[84] In reversing the lower court's finding on White's right of publicity claim, the court reasoned that "[a]lthough the defendants in these cases avoided the most obvious means of appropriating the plaintiff's identities [sic], each of their actions directly implicated the commercial interests which the right of publicity is designed to protect."[85] The court also explicitly rejected the defendant's First Amendment parody defense, arguing that the defendant's spoof of White was only "tangentially related" to the ad's primary purpose of persuading viewers to buy Samsung products and holding that "[t]he difference between a 'parody' and a 'knock-off' is the difference between fun and profit."[86]

The average U.S. citizen, however, may not have the social leverage to argue that their likeness is being appropriated without consent or compensation. As the Court in *White* stated,

> [t]he right of publicity has developed to protect the commercial interest of celebrities in their identities. The theory of the right is that a celebrity's identity can be valuable in the promotion of products, and the celebrity has an interest that may be protected from the unauthorized commercial exploitation of that identity.[87]

This was also a concern in *In re NCAA Student-Athlete Name & Likeness Licensing Litigation*, in which the court held that a video game developer's use of the likenesses of college athletes in its video games was not protected by the First Amendment.[88] And, therefore, a former college football player's right of publicity claims against the developer were not barred by California law.[89] While *In re NCAA* was a step in the right direction toward protecting uber-famous individuals, the decision was still rooted in economically-focused intellectual property law, which fails to capture the more nuanced privacy harms done the average citizen. For these reasons, the right of publicity also falls short of comprehensively addressing nonconsensual deepfakes.

---

[84] *See id.* at 1400–02 (overturning the lower court's ruling regarding White's right of publicity and Lanham Act claims).

[85] *Id.* at 1398.

[86] *Id.* at 1401.

[87] *Id.* at 1398 (alteration in original) (quoting *Carson v. Here's Johnny Portable Toilets, Inc.*, 698 F.2d 831, 835 (6th Cir. 1983)).

[88] *NCAA Student-Athlete Name & Likeness Licensing Litig.*, 724 F.3d at 1284.

[89] *See id.* ("Under California's transformative use defense, [defendant's] use of the likenesses of college athletes like Samuel Keller in its video games is not, as a matter of law, protected by the First Amendment.").

D. Criminal Revenge Porn and Nonconsensual Pornography Law

"Revenge porn" or nonconsensual pornography statutes were created almost as a direct response to the proliferation of individuals' personal photos and videos on the Internet without their consent.[90] There are two types of revenge porn laws. The first is a "narrower" type that "prohibits distributing sexually explicit images of another without the other person's consent . . . with the intent to harm."[91] This type of offense usually arises when an individual is seeking revenge against a former intimate partner by "sharing sexually explicit images that the individual obtained during the period of their intimacy," even if the image was originally taken with the subject's consent.[92] The second type of revenge porn is broader and prohibits "distributing sexually explicit images of another without the other person's consent, even [if] the distribution is not done with the intent to harm."[93] This broader criminal offense recognizes a right to control images of oneself in intimate settings.[94]

Currently, there is no remedy under federal law for adult victims of nonconsensual pornography.[95] Federal invasion of privacy laws do not apply to deepfakes because they require the unlawful intercepting, obtaining, or accessing of a person's electronic, telephone, or Internet communications.[96] And while there are federal protections for obscenity,[97] the Supreme Court has issued conflicting decisions on how to define "obscene," which calls into question the doctrine's integrity and applicability to Internet crimes.[98]

---

[90] *See State Revenge Porn Policy*, ELEC. PRIV. INFO. CTR., https://epic.org/state-policy/revenge-porn/ (last visited Oct. 30, 2020) (articulating states' interest in creating revenge porn and nonconsensual pornography statutes).

[91] ORIN S. KERR, COMPUTER CRIME LAW 245–47 (4th ed. 2018).

[92] *Id.*

[93] *Id.*

[94] *Id.*

[95] *See State Revenge Porn Policy*, *supra* note 90 ("Currently, federal law does not provide a remedy to victims of nonconsensual pornography . . . .").

[96] KERR, *supra* note 91, at 248 (detailing the elements of the Wiretap Act, 18 U.S.C. § 2511; Stored Communications Act, 18 U.S.C. § 2701; and Pen Register Statute, 18 U.S.C. § 3121).

[97] KERR, *supra* note 91, at 264, 275, 279 (explaining the obscenity standard established in *Miller v. California*, 413 U.S. 15 (1973) and preexisting federal obscenity statutes, 18 U.S.C. §§ 1460–70).

[98] KERR, *supra* note 91, at 273 (discussing how the Supreme Court's decision in *Lawrence v. Texas*, 539 U.S. 558 (2003), undermined the obscenity doctrine).

As a result, 46 states, the District of Columbia, and Guam have enacted revenge porn laws.[99] These laws, however, align more with the narrower type of revenge porn that requires the person distributing the images to have an intent to harm the subject of the images or video.[100] These laws can also require the subject to have had a reasonable expectation of privacy in the images for the distributor to be criminally liable.[101] Additionally, several of these laws require the media at issue to include the "intimate" areas or parts of the targeted individual, which usually means the "unclothed genital areas."[102]

All of these requirements are barriers to redress for those targeted in nonconsensual deepfake pornography. First, the limited application of revenge porn laws is particularly concerning regarding pornographic deepfakes because a deepfake creator may not *intend* to hurt the subject of their depictions. Deepfake creators that share their productions among friends or online without harmful intent would not be criminally liable in states with intent-to-harm requirements.[103] Second, a deepfake producer could easily utilize photos and/or video from a targeted individual's social media accounts to create highly offensive deepfakes. A defendant deepfake producer could argue successfully that someone does not have a reasonable expectation of privacy in media they publish online themselves. And finally, given that deepfakes are usually created by superimposing images of people's faces onto other images, a deepfake will rarely involve images or video of the targeted individual's actual intimate areas. Thus, revenge porn and nonconsensual pornography criminal statutes fail as an effective remedy for both public and private individuals depicted in nonconsensual deepfakes.

---

[99] *46 States + DC + One Territory Now Have Revenge Porn Laws*, CYBER CIV. RTS. INITIATIVE, https://www.cybercivilrights.org/revenge-porn-laws/ (last visited December 4, 2020).

[100] At least twenty-five state jurisdictions contain a culpability requirement that the individual must intend to cause harm to the other individual by disseminating the sexually explicit photograph or video. Harris, *supra* note 66, at 121; *see, e.g.*, KY. REV. STAT. ANN. § 531.120 (West 2018) (requiring intent to cause harm); 11 R.I. GEN. LAWS § 11-64-3 (2020) (requires "knowledge or reckless disregard for the likelihood that the depicted person will suffer harm").

[101] At least sixteen state jurisdictions have a reasonable expectation of privacy requirements. *See 46 States + DC + One Territory Now Have Revenge Porn Laws*, *supra* note 99 (providing an overview of all the state statutes that have revenge porn laws); *see also* Harris, *supra* note 66, at 121–22 (describing the expectation of privacy language used in different state statutes); *see also* MINN. STAT. § 617.261 (2019) (requiring the subject to have a reasonable expectation of privacy).

[102] Harris, *supra* note 66, at 122.

[103] *See* Harris, *supra* note 66, at 121 ("The Producers that share deepfakes amongst friends or post online without any harmful intent are not criminally liable.).

E. Criminal Harassment and Threat Law

Federal law generally prohibits threats and harassment.[104] Laws against threats and harassment that are most likely to apply to nonconsensual deepfake creators include cyberstalking and cyberbullying laws. Threats and harassment, however, are just as limited as their counterparts in revenge porn law.

1. Cyberstalking

18 U.S.C. § 2261A (more commonly known as the federal "Cyberstalking Statute") requires the defendant to have engaged in qualified stalking behavior or know that their behavior would cause serious harm to the individuals depicted in their deepfakes.[105] Like the limitations imposed by the intent requirement of other U.S. laws, it does not address the harm done to individuals based on the mere dissemination of a deepfake containing their image because of its focus on the creator's knowledge or intent.

For example, Selena Gomez, a famous American singer and performer, was depicted in a pornographic deepfake that quickly went viral in mid-2019.[106] A quick Google search returns tens of results to websites hosting one or more deep fake creations featuring Gomez that are incredibly arduous and difficult to remove, especially for the average individual. Under the Cyberstalking Statute's current language, a creator could easily establish that they could not have foreseen death or serious bodily injury to Gomez or her loved ones when they created the fake. Similarly, in the case of Rana Ayyub,[107] that deepfake creator would likely succeed in arguing that Ayyub's physical suffering and

---

[104] *See, e.g.*, 18 U.S.C. § 875 (2018) (prohibiting extortion or any threat of injury to the property or reputation of another, including threats of kidnapping, in interstate communication).

[105] *See* 18 U.S.C. § 2261A (2018) (prohibiting the use of "any interactive computer service or electronic communication system of interstate commerce" to engage in activities "with the intent to kill, injure, harass, intimidate . . . another person" and behavior that the actor knows would place another person "in reasonable fear of the death of or serious bodily injury to a person . . . .").

[106] *See* Charlotte Walsh, *What is a deepfake? This video technology is spooking some politicians*, USA TODAY, (May 6, 2020, 8:14 PM), https://www.usatoday.com/story/news/politics/2019/03/15/what-deepfake-video-technology-spooking-some-politicians/3109263002/ (using the Selena Gomez deepfake to foreshadow how the technology could be used against politicians).

[107] *See* Harwell, *supra* note 5 (discussing the physical and emotional harm suffered by an Indian journalist, Rana Ayyub, after she was featured in a deepfake without her consent).

hospital visits were entirely unforeseeable when they made the deepfake.[108] Therefore, cyberstalking laws do not provide a remedy for individuals depicted in a majority of nonconsensual deepfakes, in which physical harm to the targeted individual is not intended.

2. Cyberbullying

Virtually every state has "cyberbullying" or electronic harassment laws that were created to protect children from physical and electronic bullying.[109] What constitutes the most accurate definition of "cyberbullying" is a debatable topic.[110] A commonly used definition, however, is "speech that is 'defamatory, constitutes bullying, harassment, or discrimination, discloses personal information, or contains offensive, vulgar, or derogatory comments.'"[111] Most definitions encompass forms of bullying that use technology.[112]

These laws are often struck down as too broad when expanded to include advancing technologies.[113] For example, in *People v. Marquan*, the New York Court of Appeals found that a local law prohibiting cyberbullying against "any minor or person" or children in the county was overbroad under the Free Speech Clause of the First Amendment.[114] Furthermore, most anti-cyberbullying and harassment laws, even if they do not have intent or knowledge requirements, do not capture the "intense" form of harassment deepfakes can generate.[115]

---

[108] *Id.*

[109] *See* Sameer Hinduja & Justin W. Patchin, *State Bullying Laws*, CYBERBULLYING RSCH. CTR., https://cyberbullying.org/Bullying-and-Cyberbullying-Laws.pdf (last updated Nov. 2018) (categorizing cyberbullying and harassment laws across the U.S.).

[110] *See* SHAHEEN SHARIFF, CONFRONTING CYBER-BULLYING: WHAT SCHOOLS NEED TO KNOW TO CONTROL MISCONDUCT AND AVOID LEGAL CONSEQUENCES 39–40 (2009) (discussing the overly simplistic nature of some "cyberbullying" definitions and the resulting inability of cyberbullying measures to address cyberbullying harms).

[111] *Id.* at 41.

[112] *Id.*

[113] *See* Sameer Hinduja & Justin Patchin, *State Bullying Laws*, CYBERBULLYING RESEARCH CTR. (2008), https://cyberbullying.org/Bullying-and-Cyberbullying-Laws.pdf (analyzing the clash between cyberbullying laws and the vagueness and overbreadth doctrines).

[114] People v. Marquan M.,19 N.E.3d 480, 488 (N.Y. 2014).

[115] *See* Owen Bowcott, *Criminal law not keeping pace with digital world – report*, GUARDIAN (Oct. 31, 2018, 8:01 AM), https://www.theguardian.com/society/2018/nov/01/criminal-law-not-keeping-pace-with-digital-world-report ("At present the criminal law does not treat ['pile on'] abuse as an intense form of harassment. Future reforms could consider whether the conduct associated with 'pile on"

For these reasons, U.S. law is severely lacking in legal recourse for those victimized by nonconsensual deepfake productions. Emerging laws, such as the Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act (the DEEP FAKES Accountability Act), set out to address the specific harms presented by nonconsensual deepfakes.[116] Until those measures are passed and survive First Amendment scrutiny, however, most targeted individuals will remain vulnerable. Analyzing U.S. privacy law's longstanding connectedness to protecting human dignity and reputation, however, could serve as grounds upon which courts could shape stronger constitutional privacy protections against nonconsensual deepfakes. In turn, those victimized by and seeking to regulate deepfakes could have a better chance at defeating First Amendment challenges.

III. INFERRING STRONGER CONSTITUTIONAL PRIVACY PROTECTIONS

U.S. privacy law's root in preserving human or personal dignity provides some justification for courts to interpret the Fourteenth Amendment's right to autonomy as protecting the unique harms caused by deepfakes. While the right to autonomy is most often applied to one's physical control over their own body, connecting the privacy protection it to its broader origins in safeguarding personal dignity could justify a constitutional privacy protection that rivals the First Amendment. Furthermore, while the Supreme Court traditionally avoids balancing competing constitutional interests,[117] the unique

---

harassment' such as coordinating and inciting this behaviour [sic], could be more effectively targeted.").

[116] *See* Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019, H.R. 3230, 116th Cong. (2019) (illustrating that this bill was introduced by Rep. Yvette D. Clarke (D-NY-9) and referred to the House Committee on the Judiciary and the Committees on Energy and Commerce, and Homeland Security in June 2019); *see also* Devin Coldewey, *DEEPFAKES Accountability Act would impose unenforceable rules — but it's a start*, TECH CRUNCH (June 13, 2019, 3:25 PM), https://techcrunch.com/2019/06/13/deepfakes-accountability-act-would-impose-unenforceable-rules-but-its-a-start/ (arguing that even the DEEPFAKES Accountability Act may not provide a sufficient remedy because the law would "require anyone creating a piece of synthetic media imitating a person to disclose that the video is altered or generated, using 'irremovable digital watermarks, as well as textual descriptions'" and because it would establish "a right on the part of victims of synthetic media to sue the creators and/or otherwise 'vindicate their reputations' in court").

[117] *See* Doron M. Kalir, *The Need for Principled Balancing When Constitutional Values Collide*, SCHOLARS STRATEGY NETWORK (Nov. 16, 2018), https://scholars.org/brief/need-principled-

threat deepfakes pose to the integrity of public discourse itself, in addition to the reputational harm it causes targeted individuals, may compel the Court's attention.[118]

A. U.S. Privacy Law's Origins in Protecting Human Dignity

Over the course of U.S. history, "the notion that every individual is entitled to the requisites of human dignity has been honored more in the abstract than in legal action."[119] In developing privacy tort law, however, U.S. legal scholars and courts were heavily influenced by the German right of personality.[120] In analyzing the sources that Samuel Warren and Louis Brandeis utilized in *Right to Privacy*,[121] Paul M. Schwartz and Karl-Nikolaus Peifer state that:

> [Warren and Brandeis] considered continental law, in particular, French copyright law and the nascent French right of privacy. At a key point in their argument, however, they turned, somewhat suddenly, to a concept from German philosophy: the right of personality. To some extent at least, Brandeis had been exposed to German thought through his family background, its cultural milieu, and his own education. Brandeis was born in Louisville, Kentucky to a family with German roots, and one that returned for business reasons to Germany while he was a teenager. Brandeis then spent three semesters, from 1873 to 1875, at the Annen-Realschule in Dresden.
>
> Warren and Brandeis drew on the concept of a personality interest to develop their right of privacy as more than a new property right. Of their 'right to be let alone,' Warren and Brandeis first noted its similarity with interests in being free from assault, false imprisonment, malicious prosecution, and defamation.[122]

---

balancing-when-constitutional-values-collide (suggesting that the Supreme Court avoids employing the constitutional balancing measures that are often used in other countries).

118 *But see* Chesney & Citron, *supra* note 1, at 1803." (quoting Brown v. Hartlage, 456 U.S. 45, 60 (1982) ("[T]he 'State's fear that voters might make an ill-advised choice does not provide the State with a compelling justification for limiting speech.' Not surprisingly, courts therefore have struck down periodic attempts to ban election-related lies. The entry of deep fakes into the mix will not likely change that result).

119 Kenneth S. Abraham & Edward White, *The Puzzle of the Dignitary Torts*, 104 CORNELL L. REV. 317, 319 (2019).

120 *See* Paul M. Schwartz & Karl-Nikolaus Peifer, *Prosser's* Privacy *and the German Right of Personality: Are Four Privacy Torts Better than One Unitary Concept?*, 98 CALIF. L. REV. 1925, 1937 (2010) (discussing the German influence on U.S. tort privacy law).

121 *See* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 213 (1890) (advocating for a U.S. right to privacy).

122 *Id.* at 1943 (footnotes omitted).

The German right of personality is rooted in the idea that privacy law should protect against the degradation of human dignity.[123] Initially, throughout the early 1900s, the German Federal Supreme Court and Federal Constitutional Court were reluctant to allow the legal protection of personality.[124] In addition, many legal commentators in Germany rejected the personality right as "too broad and vague to become part of the law."[125] After the horrors inflicted by Hitler and the German state, however, German courts opened up to the concept of protecting personality and dignity by enacting Germany's Basic Law and the Civil Code, *Bürgerliches Gesetzbuch* (Basic Law).[126] "The Basic Law grants protection of dignity to all humans . . . due to their unique individual status rather than their racial identity. The Basic Law also declares in its critical Article 1(1) that human dignity is 'inviolable,' which means that the State cannot take it away or destroy it."[127]

Over time, German federal courts strengthened privacy protections by rooting them in the dignity of the individual.[128] In a groundbreaking case, *Schacht*, the Federal Supreme Court held that German citizens have a "right of personality," defined as "the right of the individual to be respected in his dignity as a human being and to develop his individual personality."[129] Later, in 1958, the Federal Supreme Court identified a damage remedy for personality rights in the German Civil Code.[130] The Federal Constitutional Court finally ratified the Federal Supreme Court's decisions regarding the right

---

[123] *See id.* at 1946 (stating that there would be a "real difference [in U.S. privacy law] if 'the tort of invasion of privacy [were] taken to protect the dignity of man' as opposed to Prosser's four tort interests." (second alteration in original) (quoting Edward J. Bloustein, *Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser*, 39 N.Y.U. L. REV. 962, 1004 (1964))).

[124] *See id.* at 1947–48 ("German courts throughout the early 1900s were also reluctant to allow the legal protection of personality, apart from certain statutorily protected interests, such as the interests safeguarded through copyright law.").

[125] *Id.*

[126] *Id.*

[127] *Id.* at 1950.

[128] *Id.*; *see also* CHRISTIAN BUMKE, ANDREAS VOSSKUHLE, CASEBOOK VOSSKUHLE/VERFASSUNGSRECHT 76–84 (5th ed. 2008) (detailing developments in German privacy rights post-WWII).

[129] Schwartz & Peifer, *supra* note 120, at 1950 (citing Bundesgerichtshof [BGHZ] [Federal Court of Justice] May 25, 1954, Neue Juristische Wochenschrift [NJW] 334 (1954) (Ger) (*translated in* German Law Archive, https://germanlawarchive.iuscomp.org/?p=108).

[130] *Id.* at 1951 (citing Bundesgerichtshof [BGHZ] [Federal Court of Justice] Neue Juristische Wochenschrift [NJW]349 (1958) (Ger) (*translated in* German Law Archive, https://germanlawarchive.iuscomp.org/?p=113).

of personality in the 1973 case, *Soraya*.[131] And today, the German courts' decisions on the right of personality are "uniformly accepted."[132]

Prosser "stripped" the high-level, German-influenced concepts from Warren and Brandeis' writings to begin constructing the Four Prosser Torts, and, at that time, most other authors "rallied around the notion of the right of personality as the basis for a privacy tort."[133] When Prosser released his four categories, however, U.S. privacy and tort law "halted at the lines that [Prosser] drew."[134] While Prosser's four categories provided some clarity on issues of privacy law, it obscured the right to personality's classification as fundamental right under U.S. common law. Meanwhile, the superior "constitutional dimension" of German privacy tort law continued to safeguard "an absolutely protected sphere that is typically associated with highly intimate information or extremely sensitive aspects of private life."[135] As Schwartz and Peifer state:

> a unitary value grounded in the fundamental worth of human dignity has led to the protection of more interests than those covered by Prosser's four privacy torts . . . . At the same time, moreover, the German emphasis on the protection of dignity has not led to radically weaker protection for the freedom of expression. The German right of personality permits publication of newsworthy matters and strong criticisms of others that impinge on privacy.[136]

Thus, in concurrence with Schwartz and Peifer, instead of scrapping the Prosser Torts entirely, U.S. legal scholars and judges should continue analyzing and interpreting privacy protections with the historical link to a German unitary concept of human dignity in mind.[137] Specifically, courts

---

131 *See* Schwartz & Piefer, *supra* note 120, at 1951 n.179 ("[t]he systems of values of the Basic Law finds its focal point in the free human personality and his dignity, which develops in the social community. It is due respect and protection from all points of governmental power (Art. 1 and 2 para. 1 GG). Such protection extends to the private sphere of humans, the sector in which a person can remain alone, to reach decisions in his own responsibility, and to be free from invasions of all kinds[.]" (quoting Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court] Feb. 14, 1973, Neue Juristische Wochenschrift [NJW] 269 (1973))).

132 Schwartz & Peifer, *supra* note 120, at 1952.

133 *Id.* at 1944*; see also id.* ("[Prosser] transposed Warren and Brandeis's work into a comfortable middle range, one light on theory but heavy on doctrinal distinctions for judges and practitioners to follow.").

134 *Id.* at 1946*; see also id.* at 1944 ("For many years after Warren and Brandeis's contribution, other authors on the subject of privacy also rallied around the notion of the right of personality as the basis for a privacy tort. The sway of this idea remained unchallenged . . . until Prosser introduced his four-part classification.").

135 *Id.* at 1986–87.

136 *Id.* at 1947.

137 *Id.*

should view the reputation and personality tort principles underpinning U.S. defamation law as central, rather than tangential, to U.S. privacy law generally. Recognizing this history could be the first step in acknowledging the fundamental importance of protecting one's personality and how the bulk of U.S. privacy law was conceived for that exact purpose. Only then may courts consider how First Amendment limitations on privacy laws pose a significant threat to the most basic interests of U.S. citizens—as exemplified by the deepfake dilemma.

B. The Link Between the Constitutional Right of Autonomy and Protecting Human Dignity

A reinforced constitutional right to privacy that encompasses the human dignity factor protected under German constitutional law could be inferred from the right of autonomy under the 14th Amendment. It is important, however, to first explore how U.S. privacy *tort* law can be elevated to a *constitutional* protection of human dignity when, unlike in German law, the protection is not explicitly stated in the Constitution or its case law progeny. Indeed, where a significant interest or right is not included in the Constitution explicitly, the Constitution can be amended to include an explicit provision, or the right can be interpreted from existing constitutional protections. This comment address the latter approach.

The Supreme Court has showcased frequently its ability to interpret "a right of personal privacy, or a guarantee of certain areas or zones of privacy" in lieu of an explicit right in the Constitution.[138] Most notably, in *Roe v. Wade*, the Court held that the Due Process Clause of the Fourteenth Amendment creates a fundamental "right to privacy" that protects a woman's decision to have an abortion.[139] It held further that that right must be balanced against the government's interest in protecting women's health and potential human life.[140] The Court achieved this by piecing together various constitutional protections

---

[138] Roe v. Wade, 410 U.S. 113, 152 (1973). The Supreme Court held in *Roe* that the "right of privacy . . . founded in the Fourteenth Amendment's concept of personal liberty and restrictions upon state action . . . is broad enough to encompass a woman's decision whether or not to terminate her pregnancy." *Id.* at 153.

[139] *Id.*

[140] *Id.* at 153–54; *see also* Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 871–73 (1992) (modifying the balancing inquiry in *Roe* but maintaining the right to abortion).

and common law rights that, together, manifested an implicit right.[141] But only personal rights deemed "fundamental or implicit in the concept of ordered liberty are included in this guarantee of personal privacy."[142]

The right of autonomy was established, and is most often discussed, in reproductive rights and justice cases. In *Planned Parenthood v. Casey*, the court held that:

> Our precedents have respected the private realm of family life which the state cannot enter. These matters, involving the most intimate and personal choices a person may make in a lifetime, choices central to *personal dignity and autonomy*, are central to the *liberty* protected by the Fourteenth Amendment. At the heart of liberty is the right to define one's *own concept of existence, of meaning, of the universe, and of the mystery of human life*.[143]

The right of autonomy emphasizes the importance of protecting one's own bodily interests, including the "concept of existence" and "personal dignity."[144] Thus, the right of autonomy is directly linked to protecting personal or human dignity as a fundamental right—the root of Germany's constitutional right to personality and the common law origin of U.S. privacy protections. For this reason, the Court could reasonably bolster existing privacy tort protections with constitutional authority.

Furthermore, the Court would be similarly authorized to interpret the right of autonomy to include non-bodily dignity interests or interests not involving the physical invasion of one's person. As stated in *Casey*, "choices central to personal dignity and autonomy, are central to the liberty protected by the Fourteenth Amendment."[145] The "concept of existence" is also at the heart of liberty.[146] Furthermore, in *Meyer v. Nebraska*, the Court held that liberty is "not merely freedom from bodily restraint" but includes other privileges that

---

141 *See Roe*, 410 U.S. at 152–53 (explaining how different constitutional provisions and common law rights have led to the Court recognizing an implicit right to privacy); *see also* Griswold v. Connecticut, 381 U.S. 479, 484–85 (1965) (analyzing the First, Third, Fourth, Fifth, and Ninth Amendments to create the right to privacy in marital relations); Skinner v. Oklahoma ex rel. Williamson, 316 U.S. 535, 541–42 (1942) (finding a privacy right to procreate); Eisenstadt v. Baird, 405 U.S. 438, 453–54 (1972) (finding a privacy right to contraception).

142 *Roe*, 410 U.S. at 152 (internal quotation marks omitted) (citation omitted).

143 *Casey*, 505 U.S. at 851 (internal quotation marks omitted) (emphasis added) (citation omitted).

144 *Id.*

145 *Id.*

146 *Id.*

have been "long recognized at common law as essential to the orderly pursuit of happiness by free men[,]" such as the rights to contract and learn.[147]

The privacy interests threatened by nonconsensual deepfakes are non-bodily dignity interests that could be protected under the Fourteenth Amendment. The Fourteenth Amendment protects the fundamental right "to define one's own concept of existence,"[148] and nothing is more central to one's own concept of existence than their personality and reputation. Nonconsensual deepfakes deprive individuals of their autonomous right to build their own personality and reputation, especially when they depict individuals engaged in conduct that they, or any reasonable person, would find highly offensive. And like the rights to contract and learn, one's interest in protecting their reputation or personality, for private parties in particular, is long recognized in case law.[149]

Moreover, it is well understood that the Fourteenth Amendment limits state and government action only and cannot be invoked against private parties.[150] Nesting the personality harm created by deepfakes in the right to autonomy, however, would not create a new private right of action against private parties. Instead, it would establish a constitutional right to defend one's own personality and reputation, an interest at the heart of privacy common law. As a result, the right will bolster a plaintiff's case against a deepfake creator who raises a First Amendment defense, increase legislative and enforcement power against deepfakes, and force courts to engage in a well-reasoned balancing analysis of the competing constitutional interests.

Thus, federal courts could infer a 14th Amendment right to privacy that encompasses the human dignity factor at the root of U.S. privacy tort law and German constitutional law. This inference would significantly strengthen constitutional privacy protections against First Amendment defenses in cases concerning nonconsensual deepfakes. If the right of autonomy is recognized

---

147 Meyer v. Nebraska, 262 U.S. 390, 399 (1923).

148 *Casey*, 505 U.S. at 851.

149 *See* MCCARTHY & SCHECHTER, *supra* note 78, at § 1:20 (outlining cases that have recognized the right to privacy); *see also* KERR, *supra* note 91, at 248 (discussing federal criminal laws that recognize an individual's common law right to control images of themself in intimate settings).

150 *See* U.S. CONST. amend. XIV, § 1 ("No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.").

as protecting the dignity interests threatened by deepfakes, then the next step would be for federal courts to develop an analytical mechanism for balancing competing First and Fourteenth Amendment interests.

C. Germany's Example of Constitutional Balancing

Constitutional courts around the world use constitutional balancing or "proportionality" as the cornerstone for resolving constitutional conflicts.[151] Constitutional proportionality "requires that government intrusions on freedoms be justified, that greater intrusions have stronger justifications, and that punishments reflect the relative severity of the offense."[152] The U.S., however, has not standardized the practice.[153] For example, in *Masterpiece Cakeshop v. Colorado Civil Rights Commission*, the first instance where the U.S. Supreme Court engaged directly with a dispute between two constitutional rights, the Court declined to engage in constitutional balancing in favor of deciding the case on technical grounds.[154] But fortunately, if the Court decides to develop a carefully-reasoned balancing formula for lower courts, legislatures, and rights-advocates to use when considering their own policies, it would not have to start from scratch.

Germany utilizes constitutional balancing, and it could serve as the perfect model for U.S. courts, especially when addressing competing First Amendment and constitutional privacy protections. Not only do the German and U.S.'s privacy regimes share origins in protecting personal dignity and autonomy,[155] but Germany is also one of the most active countries in the fight

---

151 *See* Kalir, *supra* note 117 (listing Canada, Germany, South Africa, France, and Israel as some of the countries utilizing constitutional balancing).

152 Vicki C. Jackson, *Constitutional Law in an Age of Proportionality*, 124 YALE L.J. 3094, 3094, 3110-21 (2015).

153 *See id.* (arguing that while many areas of U.S. constitutional law include some elements of proportionality analysis, the U.S. could benefit from a greater use of proportionality principles and doctrine, especially on issues free speech).

154 *See* Masterpiece Cakeshop, Ltd. v. Colo. Civ. Rts. Comm'n, 138 S. Ct. 1719, 1724 (2018) (holding that the Colorado Civil Rights Commission's conduct in assessing a cake shop owner's reasons for declining to make a wedding cake for a same-sex couple violated the Free Exercise Clause); *see also* Kalir, *supra* note 117 (discussing the Court's tendency to decide cases on legal technicalities rather can addressing constitutional conflicts head-on).

155 *See* Schwartz & Peifer, *supra* note 120, at 1972–81 (detailing the common origins of U.S. and German privacy law).

against offensive uses of viral deception.[156] U.S. courts could adopt or modify the "significant impact" analysis employed by German courts to evaluate competing but equal constitutional interests. In *Generalverbot*, a 2009 German Federal Supreme Court case, the court relied upon the "significant impact" requirement to weigh the pros and cons of publishing photos of the children of Franz Beckenbauer, "a famous and successful soccer player[,] coach, and businessman."[157] In *Generalverbot*, the court acknowledged that "[t]he question of the permissibility of the publication of a picture requires in every individual case a weighing of the information interest of the public and the interest of the pictured party in the protection of his private sphere."[158] While ultimately finding that "a predominant interest in the information is . . . affirmed,"[159] the underlying interest in preserving human dignity and privacy allowed the court to engage in a comprehensive balancing inquiry that is more difficult under the free speech protections of the First Amendment.[160] While First Amendment protections create stringent actual malice requirements in defamation cases involving public figures or public matters, German courts, as in *Generalverbot*, are empowered to consider the harm done to a plaintiff's privacy right in their analysis. Furthermore, establishing a bright-line rule for balancing competing constitutional interests could assist legislators with crafting deepfake legislation that can survive First Amendment free speech restrictions.[161] Thus, upon surveying existing constitutional proportionality approaches, such as the "significant impact" test utilized by German courts, U.S. courts could develop an effective tool for reconciling disputes between privacy and First Amendment constitutional rights on a case-by-case basis. As a result, courts and legislatures will be better equipped to tackle unique

---

156 *See* Heidi Tworek & Paddy Leerssen, *An Analysis of Germany's NetzDG Law* 1–2 (Transatlantic Working Grp., Working Paper, 2019), https://www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf (discussing Germany's Network Enforcement Act (*Netzwerkdurchsetzungsgesetz* or NetzDG), an international test case in the regulation of defamatory deepfake material).

157 Federal Supreme Court, 112 GRUR 173, 173–74 (2010).

158 *Id.* at 174.

159 *Id.* at 175.

160 *See* Jackson, *supra* note 152, at 3094 (discussing the need to employ constitutional balancing to resolve conflicts involving free speech); *see also* Schwartz & Peifer, *supra* note 120, at 1954–55 (providing a detailed explanation of the "significant impact" analysis).

161 *See* Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019, H.R. 3230, 116th Cong. § 2(a) (2019) (granting the Attorney General with the authority to issue waivers from any requirements or liabilities if the producer demonstrates that compliance with the law would "impede their ability to engage in otherwise lawful activities protected by the First Amendment").

constitutional challenges, such as the ones presented by nonconsensual deepfakes.

CONCLUSION

There is an increasing need to protect public and private parties who are victimized in nonconsensual deepfakes. The First Amendment, however, poses a significant challenge to the U.S. government's ability to regulate nonconsensual deepfakes and individuals' ability to obtain redress for their harms. To remedy this issue, federal courts should infer a constitutional right to defend one's personality and reputation from the 14th Amendment right of autonomy. This interpretation would fortify plaintiffs against deepfake creators and compel courts to define clear rules for balancing competing constitutional interests. Ultimately, it would create space for lawmakers to develop more precise and innovative methods for discouraging outrageous attacks on human dignity while upholding core constitutional ideals.