Subcommittee on Cybersecurity, Information Technology, and Government Innovation
Committee on Oversight and Accountability

United States House of Representatives

Addressing Real Harm Caused by Deepfakes

March 12, 2024
2:00 PM
Room 2154 Rayburn House Office Building

Testimony of

Ari Ezra Waldman, JD, PhD
Professor of Law and, by courtesy, Professor of Sociology
University of California, Irvine School of Law

Chairwoman Mace, Ranking Member Connolly, and members of the Committee:

Thank you for the opportunity to provide testimony about the dangers of and possible responses to deepfakes. My name is Ari Ezra Waldman and I am a professor of law at the University of California, Irvine School of Law, where I research, among other things, the impact of new technologies on marginalized populations. Given my commitment to these issues and my own personal experience with image-based abuse, I also sit on the Board of Directors of the Cyber Civil Rights Initiative, or CCRI, the leading nonprofit organization dedicated to combatting image-based sexual abuse and other technology-facilitated harms. Although I sit on the Board of CCRI, I am here in my capacity as an academic and researcher.

In my brief remarks today, I would like to touch on three topics: the impact of technology, the unique and disproportionately borne harms of deepfakes, and what to do about it.

Let's start with the impact of technology. Many people, without the benefit of or need for a technical education, can crudely cut out someone's face in a JPEG or PDF and paste it onto someone else's body. We've had that technology for decades and although it still happens, it is also relatively easy to spot. But advances in machine learning have made highly realistic-looking deepfake images and videos much easier for anyone to create. And, again, with no technical expertise needed. Sometimes, all an AI needs is an instruction. Of course, that's where the portmanteau "deepfake" comes from: it's the words "deep-learning" (referring to a method of artificial intelligence) and "fake". So, we have a proliferation problem. Technology didn't create the problem of faked images and videos, but it has certainly made the problem bigger, harder to identify and dismiss, and vastly more common.

But common does not mean the harm is evenly distributed. Deepfakes cause unique harms that are disproportionately experienced by women, particularly those who are intersectionally marginalized, like Black women and trans women. So much of the history of modern technology begins with men wanting to objectify and sexualize women; it's no wonder that recent advances in deep learning technology is reflecting our cultural and institutional biases against women.

The people who create, solicit, and distribute deepfake porn of women and girls have many motives—to make money, to channel feelings of inadequacy or rejection, to gain the misguided admiration of their peers—but what they all have in common is a refusal to see their victims as full and equal persons. Like other forms of sexual exploitation, deepfake porn is used to punish, silence, and humiliate women, pushing them out of the public sphere and away from positions of power and influence. This form of image-based abuse inflicts particularly severe and unique dignitary and expressive harms on those targeted, hijacking their images and identities for entertainment and objectification. Let's be clear. This isn't mere speech. And it isn't protected by the First Amendment.

The harm caused by artificial nonconsensual pornography is virtually indistinguishable from the harm caused by actual nonconsensual pornography: extreme psychological distress that can lead to self-harm and suicide; physical endangerment that can include in-person stalking and harassment; and financial, professional, and reputational ruin. Fake sexually explicit imagery is also used to extort depicted individuals, including children, into sending actual sexually explicit imagery of themselves to their blackmailers.

There are new deepfake porn apps and web services that launch every month, and platforms don't seem willing to do anything about them. These services produce thousands of images every week, and those images are shared on websites that Google and other platforms list in their results, and prominently so. Deepfakes also go viral on social media platforms, as in the recent case involving Taylor Swift.

But Taylor Swift is a billionaire businesswoman who can pack an 80,000-seat stadium in LA night after night. She undoubtedly has a few lawyers who can threaten platforms and perpetrators, and use Swift's fame to protect their client from harm. Although so many people around the world relate to Taylor Swift's music, almost none of us have the same resources at our disposal as she does. If digital forgeries of us get out there, we are often powerless. That isn't just because we all can't afford lawyers. Nor is it just because we don't have lawmakers or platforms listening to us. It's because, just like with "real" nonconsensual pornography, it is extremely difficult to mitigate the harm of deepfakes after the fact. False images and videos that are virtually indistinguishable from real ones, especially when they involve graphic sexual activity, have an indelible impact the moment they are viewed. Then those images are shared, downloaded, uploaded, distributed, and all sorts of platforms benefit from their virality. What's more, it is practically impossible today to ever fully remove images or videos from circulation once they have been shared. I've seen this firsthand, not only in my own experience, but in experiences of clients at clinics supporting victims of nonconsensual pornography and the stories we hear at CCRI.

That means we need deterrence. We need to stop this, particularly nonconsensual deepfake pornography, before it starts. That's where Congress can step in. Along with everyone at CCRI and everyone who does work protecting victims of nonconsensual pornography, I am deeply grateful that the Chair called this hearing today. But we cannot sit here, talk about the problem, and say we care when there are committees down the hall and other branches of government taking away the government's power to do anything about it. The First Amendment does not stand in the way of Congress acting. There is longstanding precedent in First Amendment law for regulating false, harmful expression that is perceived by others to be true. While false expression that is clearly not intended or likely to be mistaken for real depictions of individuals, such as parody and satire, enjoy considerable protection under the First Amendment, defamation and fraud have historically been considered exceptions to full First Amendment protection, and criminal prohibitions against impersonation, counterfeiting, and forgery have never raised serious constitutional objection. There is also precedent for regulating harmful false expression regardless of whether it is likely to be

perceived as authentic, as demonstrated by the tort of false light and federal criminal legislation prohibiting "morphed" child pornography that combines the faces of real children with the bodies of adults.

I would argue that the intentional distribution of sexually explicit, photo-realistic visual material that appears to depict an actual, identifiable individual without that individual's consent should be prohibited. Civil penalties are helpful to achieve deterrence, but despite the risks of increased criminalization, the criminal law arguably has an important role to play. This is because at its core, deepfakes challenge and undermine our capacity to know what is true and what isn't. And the moment society tolerates falsehoods that are indistinguishable from truth, that is the moment democracy dies. Deepfakes offer a "liar's dividend," as the legal scholars Danielle Citron and Bobby Chesney have argued. In a world where we can't tell the difference between true and false, thoe that are lying have the leg up.

Thank you for your time and I look forward to your questions.