Generative AI, Synthetic Child Exploitation, and the Structural Dissolution of *Ashcroft's* Production-Based Precedent[1]

By: Campbell Littlefield, Staff Writer, Vol. 33[2]

I. Introduction

The First Amendment broadly protects freedom of expression, prohibiting the government from "abridging the freedom of speech."[3] That protection however, is not absolute; the Supreme Court has recognized a few narrow categories of speech that fall outside its protection, including "obscenity, defamation, fraud, incitement, fighting words, true threats, speech integral to criminal conduct, and child pornography."[4] Child sexual abuse material ("CSAM") is among the most firmly established of these exclusions. In *New York v. Ferber*, the Supreme Court established that the compelling state interest in safeguarding minors justifies the categorical exclusion of CSAM from First Amendment protection, based on the physical harms inherent in the underlying production of CSAM.[5] Two decades later, the Court in *Ashcroft v. Free Speech Coalition* declined to extend *Ferber* to entirely "virtual" depictions of CSAM, assuming that purely synthetic images lack a cognizable, real-world victim and have

---

[1] This blog was written prior to the Seventh Circuit's decision in United States v. Anderegg, No. 25-1354 (7th Cir. Aug. 25, 2026). For a brief discussion of the decision, see *infra* note 44.

[2] Campbell Littlefield is a second-year regular division law student at Widener University Delaware Law School. Originally from Mineola, Texas, Campbell earned a B.A. in political science and pre-law from Stephen F. Austin State University, where, as a member of Alpha Chi Omega, Campbell volunteered with the Family Crisis Center of East Texas through the chapter's domestic violence philanthropy, an experience that inspired a passion for protecting children. Campbell currently serves as a staff writer for *Widener Law Review*, Volume 33. Campbell is interested in educational law and intends to clerk in Delaware after graduation and then practice in Texas.

[3] U.S. CONST. amend. I ("Congress shall make no law . . . abridging the freedom of speech . . . .").

[4] VICTORIA KILLION, CONG. RSCH. SERV., IF11072, THE FIRST AMENDMENT: CATEGORIES OF SPEECH (2026). Although courts, statutes, and older sources use the term "child pornography," this blog uses "child sexual abuse material" (CSAM), the term now preferred by advocates and law enforcement because it reflects that such material documents the abuse of a child. *See* INTERAGENCY WORKING GRP. ON SEXUAL EXPLOITATION OF CHILD., TERMINOLOGY GUIDELINES FOR THE PROTECTION OF CHILDREN FROM SEXUAL EXPLOITATION AND SEXUAL ABUSE 38–40 (2016), https://www.ohchr.org/sites/default/files/TerminologyGuidelines_en.pdf.

[5] New York v. Ferber, 458 U.S. 747, 749, 756–764 (1982). In *Ferber*, the Court upheld a bookstore owner's conviction for selling films depicting minor boys engaged in sexual conduct, though the films were not legally obscene. *Id.* at 751–753. The Court reasoned that CSAM is "intrinsically related to the sexual abuse of children" because it is a permanent record of that abuse, and that the market for it provides an economic motive for further production. *Id.* at 759–61. Given the material's minimal value, the Court held that it may be banned without satisfying the obscenity test of *Miller v. California*, 413 U.S. 15 (1973) (defining obscene material as material that, as a whole, appeals to the prurient interest, depicts sexual conduct in a patently offensive way, and lacks serious literary, artistic, political, or scientific value). *Ferber*, 458 U.S. at 762–64.

no impact on the illicit market.[6] Today, however, the accelerating advancement of generative artificial intelligence ("AI") has dismantled the factual basis of the *Ashcroft* decision.[7] Modern machine-learning models can generate high-fidelity imagery that erases the visual distinction between real and AI-generated photos, rendering the old, production-based standard set forth by *Ashcroft*, outdated.[8] When AI-generated images cause the same harms that were targeted in *Ferber,* the First Amendment standard must stop focusing on *how* an image was made and focus instead on the actual damage it inflicts. This blog will first examine the production-based standard for CSAM established by *Ferber* and *Ashcroft,* then explain how hyper-realistic AI-generated imagery inflicts the very harms *Ferber* sought to prevent, and finally propose a narrowly tailored, harm-based standard that addresses these harms while preserving First Amendment protections.

II. The Production-Based Standard

The Supreme Court's categorical exclusion of CSAM from First Amendment protection has historically relied on a strict standard focused entirely on how the material was produced.[9] In *Ferber*, the Court grounded its exclusion of CSAM on the physical abuse suffered by a real child during the production of CSAM, reasoning that the commercial market for CSAM incentivized and sustained child exploitation, linking the image directly to the underlying abuse.[10] When the Court later evaluated purely synthetic depictions of CSAM in

[6] Ashcroft v. Free Speech Coal., 535 U.S. 234, 249–56 (2002). The Child Pornography Prevention Act of 1996 banned any image that "appears to be" a minor engaging in sexually explicit conduct, including images created entirely by computer. *Id*. at 239–41. The Court held the ban unconstitutionally overbroad, reasoning that, unlike the material in *Ferber*, virtual material "records no crime and creates no victims by its production." *Id*. at 250, 258. The Court also rejected the argument that virtual images sustain the market for real CSAM, reasoning that if synthetic images were adequate substitutes, fewer producers of CSAM would risk abusing real children. *Id.* at 254–55.

[7] Seth Atisha, *Generative A.I., Virtual Child Pornography, and the First Amendment: Time to Reconsider Ashcroft v. Free Speech Coalition?*, 39 NOTRE DAME J.L. ETHICS & PUB. POL'Y 331, 358 (2025).

[8] Riana Pfefferkorn, *Addressing Computer-Generated Child Sex Abuse Imagery: Legal Framework and Policy Implications* 2, 7 The Digital Social Contract: A Lawfare Paper Series (Feb. 2024), https://www.lawfaremedia.org/article/addressing-computer-generated-child-sex-abuse-imagery-legal-framework-and-policy-implications.

[9] *Ferber*, 458 U.S. at 756–61; U.S. CONST. amend. I.

[10] *Ferber* 458 at 759–61.

Ashcroft, it maintained this strict focus on physical production.[11] Writing for the majority in 2002, Justice Kennedy held that virtual depictions remained protected because, unlike the material in *Ferber*, virtual depictions involved no real children in their creation.[12] The Court also rejected the government's argument that computer-generated images (CGI) would impede prosecutions involving real children, reasoning that "[t]he Government may not suppress lawful speech as the means to suppress unlawful speech."[13] Yet, at the time, as Justice Thomas observed, no defendant had successfully claimed that real images were computer-generated, and he suggested that the government's interest might become compelling if technology made that defense viable.[14] Together, *Ferber* and *Ashcroft* established a two-part, production-based rule under which images documenting real physical abuse are unprotected, while synthetic images remain protected speech.[15]

In the years after *Ashcroft*, federal courts relied on merely visual distinction between CGI and real photos; several circuits held that jurors could determine whether an image depicted a real child simply by viewing it, without expert testimony.[16] Today, generative AI has erased that boundary. Text-to-image models learn patterns from enormous datasets of real images and can then generate new, photorealistic results indistinguishable from actual

[11] *Ashcroft*, 535 U.S. at 241, 249–51. The Court maintained *Ferber's* focus on production for three reasons. First, it reasoned that *Ferber's* exclusion turned on how the material "was made, not on what it communicated," so its rationale did not reach images created without a real child. *Id.* at 250–51. Second, the Court noted that *Ferber* itself had pointed to simulated depictions and youthful-looking adults as a permissible alternative to using real children. *Id.* at 251 (citing *Ferber,* 458 U.S. at 763). Third, without that link to real abuse, the only remaining basis for a ban was obscenity, and sexually explicit images that did not meet *Miller's* test remained protected speech. *Ashcroft,* 535 U.S. at 240, 246–49.

[12] *Id.* at 250.

[13] *Id.* at 254–56. The government argued that because virtual images can look real, defendants charged with possessing actual CSAM could claim the images were computer generated, making it harder to prove that a real child was involved. *Id.* at 254–55. The Court rejected this as an attempt to ban protected speech to reach unprotected speech, an approach it said, "turns the First Amendment upside down." *Id.* at 255. It added that protected speech does not lose its protection merely because it resembles unprotected speech. *Id.* at 255–56.

[14] *Id.* at 259 (Thomas, J., concurring in the judgment).

[15] *See Ferber*, 458 U.S. at 759–61; *Ashcroft*, 535 U.S. at 249–51; *see also* Atisha, *supra* note 7, at 339.

[16] *See* United States v. Farrelly, 389 F.3d 649, 653–55 (6th Cir. 2004); United States v. Slanina, 359 F.3d 356, 357 (5th Cir. 2004) (per curiam); United States v. Kimler, 335 F.3d 1132, 1142 (10th Cir. 2003); *see also* Pfefferkorn, *supra* note 8 at 17.

photographs.[17] As a result, the visual test courts relied on after *Ashcroft* no longer works, and the premise Justice Thomas identified, that no defendant had successfully claimed a real image was computer-generated, no longer holds.[18] This leaves constitutional law largely unprepared for an era where the line between synthetic photographs and photographs of physical events has disappeared.[19]

III. Hyper-Realism and Harm

Modern generative AI produces images that are visually identical to real photos, and circulating this material fuels the same secondary market harms the *Ferber* Court sought to eradicate.[20] *Ashcroft* dismissed this concern, reasoning that if virtual images were adequate substitutes, few producers would risk abusing real children.[21] But that reasoning assumed synthetic images would replace real CSAM rather than circulate alongside it, when, in practice, the two are often traded together.[22] Additionally, this true-to-life synthetic media is hardly victimless, as it severely frustrates the investigative capabilities of federal and state law enforcement agencies.[23] These AI-generated images are forcing cyber-crime units to exhaust critical, time-sensitive forensic resources attempting to verify whether a depiction originates

---

[17] INTERNET WATCH FOUND., *How AI Is Being Abused to Create Child Sexual Abuse Imagery* 10–13 (2023), https://www.iwf.org.uk/media/q4zll2ya/iwf-ai-csam-report_public-oct23v1.pdf (noting that AI-generated abuse images are now "so convincing" that they are "indistinguishable from real images"); *see also* Atisha, *supra* note 7, at 336–38 (explaining how image-generation models are trained and describing a study of a single dark-web forum on which more than 20,000 AI-generated images were posted in one month, over 2,500 of which qualified as criminal "pseudo-photographs" under U.K. law); Pfefferkorn, *supra* note 8, at 2–3.

[18] *Ashcroft,* 535 U.S. at 259 (Thomas, J., concurring in the judgment).

[19] *But see* 18 U.S.C. § 2256(8)(B), (11) (defining child pornography to include computer-generated images that are "indistinguishable from" those of a real minor); *id.* § 1466A (prohibiting obscene visual depictions of minors whether or not the minor depicted actually exists); Pfefferkorn, *supra* note 8, at 6–8 (explaining that Congress enacted both provisions in the PROTECT Act of 2003 in response to *Ashcroft*, and that no court has ruled on whether the "indistinguishable from" definition is constitutional).

[20] *Ferber*, 458 U.S. at 759–61; *see also* INTERAGENCY WORKING GRP. ON SEXUAL EXPLOITATION OF CHILD., *supra* note 4 at 41 (noting that computer-generated CSAM contributes to maintaining a market for CSAM).

[21] *Ashcroft*, 535 U.S. at 254.

[22] Pfefferkorn, *supra* note 8, at 12 n.38, 22 (noting demand for AI-generated CSAM may be intermingled with the market for real CSAM).

[23] Jessica L. Terkovich, *The Emerging Danger of AI-Generated Child Sexual Abuse Material and an Unprepared Criminal Code,* 45 N. ILL. U. L. REV. 398, 400 (2025) (discussing the strain AI-generated CSAM places on law enforcement's forensic resources); *see also* Atisha, *supra* note 7, at 338–39 (describing how offenders deliberately create hyper-realistic images to divert investigators toward false leads).

from a physical crime scene or a generative model.[24] This dilution of law enforcement efficacy creates a severe backlog, which delays the identification and rescue of actual children who are currently suffering due to the production of CSAM.[25] Furthermore, offenders can weaponize lifelike synthetic media for blackmail, harassment, and grooming, creating real-world victims of CSAM.[26] This reality dismantles the core premise of *Ashcroft*: the assumption that synthetic material lacks a victim and inflicts no real-world harm.[27]

Beyond secondary harms, even images that appear to be entirely synthetic may trace back to real child victims.[28] In December 2023, researchers at the Stanford Internet Observatory found that LAION-5B, a widely used open-source dataset for training image-generation models, contained links to more than 1,000 known CSAM images, meaning that models trained on it were likely influenced by real abuse imagery.[29] Perpetrators have also fine-tuned diffusion models directly on photographs of existing abuse victims, so an image that looks entirely computer-generated can still trace back to a real child.[30] This is often done using small add-on

[24] Terkovich, *supra* note 23, at 400–02.
[25] *Id.*
[26] Offenders use AI tools to turn ordinary photographs of real children into sexually explicit images, which are then used to extort victims or circulated to harass them. *See* Fed. Bureau of Investigation, *Malicious Actors Manipulating Photos and Videos to Create Explicit Content and Sextortion Schemes,* INTERNET CRIME COMPLAINT CTR. (June 5, 2023), https://www.ic3.gov/PSA/2023/psa230605 (reporting that victims include minors). The National Center for Missing & Exploited Children has also received reports of offenders using AI-generated CSAM to extort children or their families for money. Nat'l Ctr. for Missing & Exploited Child., *Generative AI CSAM Is CSAM* (Mar. 11, 2024), https://www.missingkids.org/blog/2024/generative-ai-csam-is-csam. AI-generated images have also been used in attempts to groom children directly. In one federal prosecution, the defendant allegedly sent sexually explicit AI-generated images of minors to a fifteen-year-old boy. Press Release, U.S. Dep't of Just., Man Arrested for Producing, Distributing, and Possessing AI-Generated Images of Minors Engaged in Sexually Explicit Conduct (May 20, 2024), https://www.justice.gov/archives/opa/pr/man-arrested-producing-distributing-and-possessing-ai-generated-images-minors-engaged.
[27] *Ashcroft*, 535 U.S. at 250.
[28] Pfefferkorn, *supra* note 8, at 25.
[29] David Thiel, *Identifying and Eliminating CSAM in Generative ML Training Data and Models* 2, 7–8 Stan. Internet Observatory, (Dec. 23, 2023), https://purl.stanford.edu/kh752sm9123. LAION-5B is an open-source dataset of roughly 5.85 billion image-text pairs, compiled by scraping the internet, that has been used to train popular text-to-image models. The dataset does not contain the images themselves but links to where each image was found online. *See* Pfefferkorn, *supra* note 8, at 24 & n.85. After the researchers' findings were published, the organization responsible for LAION-5B took it down, but copies had already been downloaded and remained in private circulation. *Id*. at 24
[30] *Id.* at 2 & n.3.

files that let a perpetrator retrain part of an existing model on a narrow set of images, such as photographs of a specific victim.[31]

IV. A New Harm-Based Standard

The evolution of generative AI has broken the current production-based standard, and courts can no longer rely on an outdated visual test.[32] The law must instead develop a narrowly tailored standard that targets these novel harms without chilling protected expression. To survive strict scrutiny and avoid First Amendment overbreadth, that standard must shift from a production-based approach to a harm-based one. Where a production-based standard asks how an image was made, a harm-based standard asks what the image does: whether its creation or use inflicts concrete injury on real children, as when it is generated from real abuse imagery, depicts an identifiable child, or is used to extort or groom a minor.[33] This shift would not create a new category of unprotected speech, which the Court has refused to do;[34] rather, it would extend *Ferber's* own rationale, protecting real children from real harm, to the ways generative AI now inflicts that harm. To make the shift, Congress would enact a narrowly tailored statute targeting these specific harms, and the Court would need to narrow *Ashcroft's* holding as applied to such material.[35]

To meet this threshold, the state's compelling interest should be understood to encompass the concrete injuries synthetic material inflicts on real children: the harm to victims whose abuse imagery was used to train or fine-tune a model, the reputational and emotional harm to identifiable children depicted in sexually explicit images, and the harm to minors when

---

[31] *Id.* at 2 & n.3.
[32] *See* Pfefferkorn, *supra* note 8, at 7, 16–17.
[33] *See id.* at 10–12 (explaining that AI-generated images trained on real abuse imagery or depicting identifiable children likely fall outside First Amendment protection under *Ferber*).
[34] *See* United States v. Stevens, 559 U.S. 460, 472 (2010).
[35] *See* Atisha, *supra* note 7, at 354–57 (arguing that the Court should narrowly overrule *Ashcroft's* holding on virtual images).

such images are used to extort or groom them.[36] Generative AI also imposes serious burdens on investigators, flooding reporting systems and making it harder to identify real children in need of rescue.[37] Coupling a strict *mens rea* requirement with proof of these secondary harms keeps the standard narrowly tailored, addressing digital-age threats without infringing on the freedom of speech.[38]

V. Conclusion

The jurisprudence established by *New York v. Ferber* and *Ashcroft v. Free Speech Coalition* has reached a technological tipping point with the modern advancements of generative AI. For decades, federal courts were able to successfully manage the constitutional boundaries of child exploitation laws by relying on a clean, binary distinction: real photographs represented a physical crime scene, while primitive, computer-generated simulations remained protected speech.[39] However, generative AI has broken this production-based standard.[40] By generating images that look like real photos, modern AI inflicts the exact harms the Supreme Court originally sought to eradicate.[41] The foundational assumption of *Ashcroft* is no longer accurate.[42] To protect minors and preserve criminal investigations in the AI era without triggering First Amendment overbreadth, the courts and Congress must abandon their outdated reliance on how an image is made.[43] Instead, the law must shift to a narrowly tailored, harm-

[36] *See* Pfefferkorn, *supra* note 8 at 10–11 (explaining that AI-generated images trained on real abuse imagery or depicting identifiable children harm actual children); Thiel, *supra* note 29, at 10 (noting that repeated abuse images in training data reinforce the depiction of specific victims); *supra* note 26 and accompanying text (describing the use of AI-generated images to extort and groom minors).
[37] *See* Pfefferkorn, *supra* note 8 at 14–16; Atisha, *supra* note 7, at 338–39; Terkovich, *supra* note 20, at 400–02.
[38] *Mens rea*, Latin for "guilty mind," is the mental state the prosecution must prove a defendant had when committing a crime, such as acting knowingly or intentionally. *See Mens rea,* BLACK'S LAW DICTIONARY (12th ed. 2024). The Court has long required a mens rea element in laws restricting sexual depictions of minors to avoid punishing innocent conduct and chilling protected speech. *See Ferber*, 458 U.S. at 765; United States v. X-Citement Video, Inc., 513 U.S. 64, 78 (1994); *see also* Pfefferkorn, *supra* note 8, at 24 (stressing the need for "carefully tailored mens rea requirements" in any law addressing AI tools trained on abuse imagery).
[39] Atisha, *supra* note 7, at 338–39.
[40] *See supra* Part II.
[41] INTERNET WATCH FOUND., *supra* note 17, at 9, 10–13; Atisha, *supra* note 7, at 338–39.
[42] *Ashcroft*, 535 U.S. at 250.
[43] *See* Atisha, *supra* note 7, at 354–57.

based standard that targets the systemic injuries caused by synthetic media, thereby neutralizing this new threat while preserving First Amendment protections. The law must simply catch up.[44]

[44] Shortly after this post was written, the Seventh Circuit held that the First Amendment protects the in-home possession of obscene AI-generated CSAM that does not depict a real child. *Anderegg*, slip op. at 23–24. The decision reinforces rather than moots this post's argument. The court applied *Ashcroft* because the government's justifications were the same ones rejected in 2002, while acknowledging that AI can now "produce images 'virtually indistinguishable'" from real CSAM and expressing "misgivings" about the lines current precedent draws. *Id*. at 3, 19–23. Because the court concluded it was "not free to redraw" these lines, the change this post proposes must come from the Supreme Court. *Id.* at 3. A concurrence urged the Supreme Court to provide that guidance, noting, as Section III of this post does, that AI training datasets have been found to contain real CSAM. *Id.* at 25–27 (Lee, J., joined by Kolar, J., concurring).