



This is a repository copy of *Pornographic deep fakes: liability for breach of privacy in cases of parody?*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/229598/>

Version: Published Version

Article:

Keren-Paz, T. (2025) Pornographic deep fakes: liability for breach of privacy in cases of parody? Journal of Tort Law. ISSN: 2194-6515

<https://doi.org/10.1515/jtl-2025-0030>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Tsachi Keren-Paz*

Pornographic Deep Fakes: Liability for Breach of Privacy in Cases of Parody?

<https://doi.org/10.1515/jtl-2025-0030>

Received September 24, 2025; accepted September 24, 2025; published online October 7, 2025

Abstract: The combination of technological advancement of generative AI with the ever-increasing importance of participation and presence in the digital world for one's overall quality of life makes the harm from deep fakes a pressing problem. There is a need to both regulate deep fakes and remedy the harm they cause. In the paper, I focus on tort law, privacy-based response to a subset of the deep fake problem: nonconsensual intimate fakes ('intimate fakes'). I make the following claims in the paper: (1) Those who distribute intimate fakes of identifiable real persons where the image appears to be real should be liable to the plaintiff for breach of their privacy. Those who make these images (with a self-use exception) and those who host and view them should be likewise liable, but I will focus on distributors here. This follows from the position that false private information, even if defamatory, implicates the plaintiff's privacy interest and should allow the plaintiff to sue for breach of privacy for reasons both pragmatic and conceptual (including an analogy I make between information and genome). This position is largely endorsed by existing English law. (2) There should (and probably could) be no tort liability for the distribution of intimate fakes of fictional characters not reasonably understood to refer to a real person. However, liability for breach of sexual privacy manifested in intimate fakes should extend also to cases of look-alikes (including of an invented character) and in this sense as well liability should be strict. (3) A strong case exists for imposing liability also in cases in which the intimate fake is a known fake: ostensibly fake, or accompanied by disclaimer that it is. I first explain why these known fakes merit liability, highlighting what makes their distribution both harmful

I would like to thank Jeevan Hariharan, Greg Keating, Clare McGlynn, Suzzane Ost, Paul Wragg and the participants of the JTL Symposium (August 2025) on this special issue and of the May 2024 Deepfakes and the Law symposium (organised by Rebecca Moosavian and Thomas Bennett and held in City University, London) for very helpful comments on previous drafts, Maria Sklavou, for directing me to relevant criminological literature and Sean Martin for his excellent research assistance.

***Corresponding author: Tsachi Keren-Paz**, Professor of Private Law, The University of Sheffield, Sheffield, England, E-mail: t.keren-paz@sheffield.ac.uk

and (civilly) wrongful. I then explain that liability could be conceptualised as (also) vindicating the claimant's privacy interest. (4) There should be no tort liability for the creation of an intimate fake for self-use. However, the creator should be strictly liable for the distribution of the intimate fake; moreover, there is much to support liability of the creator towards the subject of the image should the latter learn that the image exists even if the intimate fake itself was not distributed.

Keywords: deepfakes; defamation; privacy; gendered harms; freedom of expression; democratic participation

1 Introduction

The combination of technological advancement of generative AI with the ever-increasing importance of participation and presence in the digital world for one's overall quality of life makes the harm from deep fakes a pressing problem. There is a need to both regulate deep fakes and remedy the harm they cause. In this paper, I focus on a tort law privacy-based response to a subset of the deep fake problem: nonconsensual intimate fakes ('intimate fakes'). I do not argue that intimate fakes cannot be remedied by defamation (or other causes of action); rather, intimate fakes invoke a privacy interest and a privacy cause of action is likely to be more effective than a defamation claim.¹ The

1 Views in the literature are split between defamation-led and privacy-led conceptualizations and responses. For the former see Vasileia Karasavva & Aalia Noorbhai, *The Real Threat of Deepfake Pornography: A Review of Canadian Policy*, 24 CYBERSPSYCHOLOGY, BEHAVIOR AND SOCIAL NETWORKING 203 (2021); Benjamin T. Suslavich, *Nonconsensual Deepfakes: A "Deep Problem" for Victims*, 33 ALB. L.J. SCI. & TECH. 160 (2023); Abigail George, *Defamation in the Time of Deepfakes*, 45 COLUM. J. GENDER & L. 122 (2024); Emily Pascale, *Deeply Dehumanizing, Degrading, and Violating: Deepfake Pornography and the Path to Legal Recourse*, 73 SYRACUSE L. REV. 335, 341 (2023); Jon M. Garon, *An AI's Picture Paints a Thousand Lies: Designating Responsibility for Visual Libel*, 3 J. FREE SPEECH L. 425, 438–40 (2023). And see Lindsey Joost, *The Place for Illusions: Deepfake Technology and the Challenges of Regulating Unreality*, 33 U. FLA. J.L. & PUB. POL'Y article 7 (defamation inapplicable).

For privacy-led conceptualizations and responses see Benjamin C. Zipursky & John C. P. Goldberg, *A Tort for the Digital Age: False Light Invasion of Privacy Reconsidered*, 73 DEPAUL L. REV. 461 (2024); Danielle Keats Citron & Daniel J. Solove, *Privacy Harms*, 102 B U. L. REV. 793 (2022); Danielle Keats Citron, *Privacy Injunctions*, 71 EMORY L.J. 955 (2022); Dan Nikolin, *When Seeing Is No Longer Believing: Is Australian Privacy Law Ready for Deep Fakes?*, SSRN J (2023), <https://www.ssrn.com/abstract=4368696>; A O'Connell, *Image-Based Sexual Abuse and Gendered Conceptions of Harm in Tort*, in HORSEY, K (ED.). DIVERSE VOICES IN TORT LAW 221 (2024).

For privacy sceptical views (inapplicable, undesirable or under-inclusive) see Matthew B. Kugler & Carly Pace, *Deepfake Privacy: Attitudes and Regulation*, 116 NW. U. L. REV. 611 (2021) (intimate fakes invoke privacy interests but tort law is ill equipped to remedy); Emily Laidlaw, *Technology Mindfulness and the Future of the Tort of Privacy*, 60 OSGOODE HALL L.J. 597 (2023). For the potential and

argument is largely normative; what I say about the law focuses on English private law (given that private law responses to gendered harms is my area of interest and the focus of the special issue), although I engage to some extent with the criminal side of things, refer to potential case studies across the globe and refer to some non-English authorities.² I advance several lines of argument; most importantly, however, I argue that ‘known fakes’ – intimate fakes known by those viewing them, at least by the reasonable viewer, to be fake³ – are a multi-faceted phenomenon which merits liability and could be both theorised and remedied as an invasion of privacy. Known fakes are a gendered phenomenon, as they target mainly women, and often female politicians; and they are at once a breach of privacy, a form of defamation, harassment and discrimination, and a threat to democracy, freedom of expression and equal citizenship. While known fakes distinctively implicate interests in sexual privacy and gender equality, they also raise – as typical to other types of deep fakes – concerns about the integrity of the democratic process and the undermining of fair competition in the marketplace of ideas. The analysis also reveals the relationships between (a) false facts, opinion, fantasy and the justifications for not/imposing liability for each; and (b) deep fakes (in general) as a problem of disinformation and a threat to democracy and intimate fakes as a threat to privacy, dignity and women’s equal participation in the public sphere.

A useful starting point for the analysis distinguishes between the image and its subject. With regards to each, one could ask whether they are ‘real’ or ‘fake’, although what is meant by these terms varies. In terms of the *image*, English law (developed initially in the context of child sexual abuse images prior to the popularisation of ‘deep fakes’ as a concept) distinguishes between ‘photograph’ (including video) which is not manipulated so depicts an authentic image, ‘pseudo photograph’ which is a digitally manipulated image which looks like a real photograph (so a deep fake passing as true image in our context) and ‘prohibited image’ which is a non-photographic image (so a known fake) including Computer-Generated Images (CGIs), cartoons, manga images and drawings. Such images might not (but could) involve any real child nor look like a photographic image of a real child, but rather are fantasy visual representations of child sexual abuse.⁴

limitations of both causes of action see Eric Kocsis, *Deepfakes, Shallowfakes, and the Need for a Private Right of Action*, 126 DICK. L. REV. 621 (2022).

2 For US jurisdictions statutory civil responses to intimate (and other types of) deep fakes see Kadri & West in this special issue.

3 For reasons of brevity, I do not use ‘intimate known fakes’ to denote intimate fakes known to be fake, despite the fact that known fakes include of course also non-intimate deep fakes.

4 Protection of Children Act 1978, c. 37, § 7 (photograph and pseudo photograph); Coroners and Justice Act 2009, c. 25, § 62 (prohibited images); *Indecent and Prohibited Images of Children*, CPS LEGAL GUIDANCE (May 7, 2024), <https://www.cps.gov.uk/legal-guidance/indecent-and-prohibited-images-children>.

Turning to the image's *subject*, they can be either a real person (either a celebrity or not) or fictional. A real person could be identifiable or not. The identifiability of the person is separate from the type of the image: An intimate image of an identifiable real person (for example when a person is naked and their face is showing) could be real (if such photo was taken without their knowledge or consent), fake (if the image is manipulated but appears to be real, so a pseudo photograph), or a known fake (if the person is being drawn, if the quality of the intimate fake suggests that the image is not real, or if a pseudo photograph of a high quality is accompanied by text clarifying it is a fake). Moreover, the subject's identifiability is separate from the question whether the image is real or not (e.g., a cartoon of a naked POTUS identifies the president but the image is understood to be fake) and whether the person is real or not: a fictional character could be identifiable (e.g., Superman in an intimate setting) and a real person could not (as in a hacked 'dick pic' or 'pussy shot' where the identity of the subject whose genitals' images are distributed is unknown). Finally, an intimate fake (whether passing as authentic or fake) created with the intention of depicting an invented fictional character could be reasonably understood by the audience as referring to a real person.⁵

Space constraints prevent me from covering all the rubrics in a grid based on the above distinctions. Rather, I will limit the analysis to few examples which are either pragmatically important (as prominent) or theoretically interesting/contested. I make the following claims in the paper: (1) Those who distribute intimate fakes of identifiable real persons where the image appears to be real should be liable to the plaintiff for breach of their privacy. Those who make these images (with a self-use exception) and those who host and view them should be likewise liable, but I will focus on distributors here.⁶ This follows from the position that false private information, even if defamatory, implicates the plaintiff's privacy interest and should allow the plaintiff to sue for breach of privacy for reasons both pragmatic and conceptual (including an analogy I make between information and genome). This position is largely endorsed by existing English law. (2) There should (and probably could) be no tort liability for the distribution of intimate fakes of fictional characters not reasonably understood to refer to a real person. However, liability for breach of sexual privacy manifested in intimate fakes should extend also to cases of look-alikes (including of an invented character) and in this sense as well liability should be strict.

⁵ Similarly, look alike cases can exist where the creator of a deep fake intended to depict person A (with/out A's consent) but the image is understood to be B's; or, where one distributes A's authentic intimate image (with A's consent) but the image is understood to be B's.

⁶ In TSACHI KEREN-PAZ, *EGALITARIAN DIGITAL PRIVACY: IMAGE-BASED ABUSE AND BEYOND* (2023) I call for strict liability of platforms which host non-consensual intimate images and of viewers; and argue that strict liability for breach of privacy is compatible with the understanding of breach of privacy as an intentional tort; and that each unauthorised viewing is an intrusion into privacy.

(3) A strong case exists for imposing liability also in cases in which the intimate fake is a known fake: ostensibly fake, or accompanied by disclaimer that it is. I first explain why these known fakes merit liability, highlighting what makes their distribution both harmful and (civilly) wrongful. I then explain that liability could be conceptualised as (also) vindicating the claimant's *privacy* interest. (4) There should be no tort liability for the creation of an intimate fake for self-use. However, the creator should be strictly liable for the distribution of the intimate fake; moreover, there is much to support liability of the creator towards the subject of the image should the latter learn that the image exists even if the intimate fake itself was not distributed.

2 The Core Case: Intimate Fakes as a Privacy Invasion

The core case includes intimate fakes that are *not* known fakes. I deal with known fakes in Section 4. Intimate fakes are one instance of information which is both private and false. As the information is also defamatory, the question presents itself whether it should be remedied only by defamation or conceived also as misuse of private information (the English privacy tort). The starting point of English law is sound: what matters is whether the information is private, not whether it is true or false.⁷ Like many others, I think this is justified.⁸ If control over the information (autonomy) and dignity are the values underpinning privacy⁹ the dissemination of true or false information about the matter undermines the claimant's right and ability to be let alone and to keep other people out of this corner of their life. The arguments in support of such view are both conceptual and pragmatic. Conceptually,

7 *McKennitt v. Ash* [2006] EWCA Civ 1714 [86]. However, where the nub of the claim is reputational, a claim in privacy might be blocked as an abuse of process: *id* [79]. Some privacy scholars are both wedded to the idea that conceptually the core of privacy is and should be about true facts and that doctrinally, the proposition that *McKennitt* and other case law are a sound authority for false facts as private could be doubted, for example, since *McKennitt* dealt with many facts, some of them true and others not. see e.g., John Hartshorne, *An Appropriate Remedy for the Publication of False Private Information*, 4 J. MEDIA L. 93 (2012); Nicole. A. Moreham, *Privacy, Reputation and Alleged Wrongdoing: Why Police Investigations Should Not Be Regarded as Private*, 11 J. MEDIA L. 142 (2019).

8 Eg., Eric Barendt, *An Overlap of Defamation and Privacy?*, 7 J. MEDIA L. 85 (2015); Robert Craig & Gavin Phillipson, *Privacy, Reputation and Anonymity until Charge: ZXC Goes to the Supreme Court*, 13 J. MEDIA L. 153 (2021) (responding to Moreham's argument *id.*; and arguing that reputation, both conceptually and pragmatically is inseparable from privacy and misuse of private information should remedy harm to reputation from revealing true and sensitive information). For a recent support of the false light tort, also in the context of deep fakes see Goldberg and Zipurski, *supra* note 1.

9 *Campbell v. MGN* [2004] UKHL 22 [51]; HELEN FENWICK & GAVIN PHILLIPSON, *MEDIA FREEDOM UNDER THE HUMAN RIGHTS ACT* 804 (2006).

if the subject matter is private, all its potential instantiations are private, not only the actual one. So if a penis's size is private, statements that X's penis size is 8, 12 or 20 cm all reveal private facts. Part of the benefits afforded by privacy is exactly avoiding a seemingly informed discussion of private matters regardless of whether the alleged private fact happens to be correct or not. So if we analogise information with genes, if a certain gene is private, all of its alleles are. Since one's intimate image is private – it is a private gene – all non consensual depictions of it are breach of privacy, regardless of whether they are true or not. Indeed, there are two ways in which intimate images could be fake. They could be fake, in the technological sense, even if the generated image is identical to the plaintiff's body, if they were created by digital manipulation. They could also be fake in a second sense resembling common understanding of photoshop: if the generated image is different from the plaintiff's body. Regardless of whether the fake image is 'prettier' or 'uglier' than the corporeal body, it is fake. Both ways of producing fake intimate imagery infringe one's privacy. They reveal a private information about the plaintiff; and the fact the information is false is neither here nor there; it is one allele of a private gene, so revealing the private information without the plaintiff's consent is a breach of privacy. Nor does it matter that the circulation of the deep fake might be, and in all likelihood is, defamatory. To the extent that it is, the plaintiff can decide whether to sue in defamation. But as the information is obviously private, they should be able to sue in privacy. They could sue in both, provided they do not recover more than the harm they suffered.

That private false information – including fake intimate images – is private and hence plaintiffs should be allowed to sue for breach of privacy those who misuse their private information is, I think, true as a theoretical proposition regardless of resorting to policy or pragmatic considerations. But within the teleological legal realist and socio legal approach to law I espouse, there is even a stronger reason to hold false private facts as private.¹⁰ First, channelling plaintiffs to sue in defamation for false private facts and in privacy for true ones for itself undermines the plaintiff's interest in privacy; it forces litigants and the trier to litigate over the question whether the disputed fact is true or false. This in itself undermines one's privacy. On an understanding of privacy as the right to be let alone and to vindicate one's desired inaccess,¹¹ that division of labour is a non-starter: it puts at the centre of the *lis* the very issue the plaintiff wishes to exclude others from knowing or discussing. In

¹⁰ Elsewhere, I take this argument forward to claim that lying about a private matter should be permitted as lying might be the only effective way to preserve a zone of privacy and as there might not be a valid public interest in putting the record straight. Tsachi Keren-Paz, *Privacy-Preserving Lies*, in *DISHONESTY, LIABILITY AND THE LAW* 166 (2025).

¹¹ Nicole Moreham, *Privacy in the Common Law: A Doctrinal and Theoretical Analysis*, 121 *LAW Q. REV.* 628.

contrast, focusing on the character of the fact as private (regardless of whether it is true or false) respects the plaintiff's autonomy of whether to address the truthfulness or falsity of the alleged fact. It allows those who wish to assert that the fact is false to do so, without compelling those who prefer not to address this issue to litigate it as a condition to receive a remedy.¹²

That the court might give anonymity orders in privacy litigation cannot support the division of labor approach. First, depending on the jurisdiction, anonymity orders are not always available; second, sometimes anonymity orders will be ineffective so the plaintiff's identity will be exposed; finally, even under an anonymity order, a division of labor compels the plaintiff to litigate the fact in front of the adversary, adjudicators, legal representatives and at time people close to them who often are the people the plaintiff wishes most to keep uninformed.¹³ Consider a gay youth from a religious conservative community who is being threatened to be outed. An anonymity order would do little to protect them from their parents' opprobrium if they need to litigate the fact whether their same-sex sexual orientation is true or not. From expressive perspective too – as both the outing example¹⁴ and intimate fakes demonstrate – it is better to remedy these cases as a breach of privacy than as a defamatory publication, given the expressive ramifications of establishing that the behaviour is likely to lower the plaintiff's reputation in the eyes of the right-thinking people in society as a condition for defamatory meaning. Put differently, it avoids the need for the law to 'ratify', endorse, or to give its imprimatur to the view that being gay is something that will harm one's reputation in the eyes of *right-thinking* people in society. Rather, as the issue is private, it should be out of bounds and the damages awarded could compensate any reputational loss if such loss is likely to occur, and regardless of whether the statement is true or false. Policy and scholarly discussions of privacy (and defamation) litigation highlight that the litigation itself increases public attention to statements that undermine one's privacy or reputation.¹⁵ A bifurcated system only worsen this problem.

Secondly, as Nimmer observed long ago (in the US context in support of the false light tort¹⁶), the harm suffered by the plaintiff from revealing false or true private

12 Indeed, both conceptualising privacy this way and the choice this understanding of privacy gives its bearers can be linked with Graeme Laurie's realm of psychological spatial privacy in, *GENETIC PRIVACY: A CHALLENGE TO MEDICO-LEGAL NORMS* (2002) and with Lisa Austin's privacy as self-presentation in Lisa Austin, *Privacy and Private Law: The Dilemma of Justification*, 55 *McGILL L.J.* 165, 204 (2010).

13 While during litigation, the privacy interest of the claimant might be undermined even under 'privacy-encompasses-false-facts' rule, e.g., while litigating public interest in publishing the fact, for which truth or falsity is relevant, the alternative rule undermines to a greater extent the interest in privacy.

14 KEREN-PAZ, *supra* note 6, 204.

15 See e.g., Danielle Citron, *Sexual Privacy*, 128 *YALE L.J.* 1870, 1930 (2019).

16 Restatement (Second) of Torts § 652 E (Am. Law Inst. 1977).

information about them – in his example, real and fake non-consensual intimate images (NCII) – is identical,¹⁷ this is supported empirically, including by public attitudes surveys.¹⁸ Since the harm is identical, so should be the legal response.

Finally, as privacy law is more claimant friendly compared with defamation (at least in the UK and the EU legal orders) e.g., in terms of the availability of interim injunctions and a longer limitation period, the plaintiff should be afforded the more effective means for protecting this important interest. So regardless of the conceptual point, the fact that intimate fakes are both wrongful and harmful justifies an effective legal response. Given that privacy law is better equipped to remedy the harms from deep fakes than defamation law,¹⁹ false private facts should be actionable in misuse of private information even if there were some doubt about the conceptual fit. In this regard, the fact that sexual privacy is a gendered interest only bolsters this conclusion by adding to the mix egalitarian and distributive justice considerations. Like authentic non-consensual intimate images, deep fakes are a gendered phenomenon in terms of both perpetration and victimhood.²⁰ It is a systemically gendered phenomenon on the spectrum of sexual abuse. As such, egalitarian and distributive justice considerations support the more effective private legal response, which is a breach of privacy claim.

The account I offer explains why the common view according to which intimate deep fakes are not a breach of information privacy is unconvincing. The dominant version of this account maintains that the private facts whose public disclosure is prohibited need to be true so the fact that the image (the relevant fact) is fictional or publicly sourced makes it inactionable.²¹ Conceptually, however, if a naked body or a

17 Melville B. Nimmer, *The Right to Speak from Times to Time: First Amendment Theory Applied to Libel and Misapplied to Privacy*, 56 CALIF. L. REV. 935, 963 (1968).

18 Kugler and Pace, *supra* note 1, 643, report that intimate fakes were perceived by the public to be as harmful as and slightly more blameworthy than authentic nonconsensual intimate images. Another attitude study found that while the distribution of authentic NCII was seen as significantly more deserving of criminalization than sharing deepfake pornography, ‘the distribution of deepfake pornography was considered especially bad’ and ‘a general sentiment that people should get upset if they are the target of a deepfake pornography video’ in Rebecca Umbach et al., *Non-Consensual Synthetic Intimate Imagery: Prevalence, Attitudes, and Knowledge in 10 Countries*, in PROCEEDINGS OF THE CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS 1 (2024).

19 Still, the sticking point is the ability to clean the virtual space from the photos, or better, still, to ensure they are filtered out, which is currently not the law. For a call to have both hosts and viewers strictly liable for breaching the subject image privacy rights see KEREN-PAZ, *supra* note 6.

20 *Id.* 90–96; Clare McGlynn & Rüya Tuna Toparlak, *The ‘New Voyeurism’: Criminalizing the Creation of Deepfake Porn*, 52 J L. & Soc’y 204 (2025).

21 See e.g., Laidlaw, *supra* note 1, 616 (‘The intrusion on seclusion tort is inapplicable because it is the manipulation of an image, rather than intrusion into private spaces, that is the core mischief. The public disclosure tort likely does not apply to deepfakes because that which is publicized must be private, and the image manipulated might not be private information, unless the claim is by the person whose body has been used in the video’).

sexual content are a private subject matter, once they refer to an identifiable plaintiff, their circulation involves a private fact about the plaintiff; the fact that the private fact is false does not make it any less private.

Whether an intimate deep fake is conceptually an intrusion is a harder question. As I discuss in Section 5, its creation does not cross a physical boundary in the same way that taking an authentic intimate image does, which is the common argument against understanding the making of (or distributing) an intimate deepfake as intrusion.²² But on the functional equal harm finding, insisting that both victims and the public perceive both the harm from intimate deepfakes and the wrongfulness of disseminating it as identical or near identical to that involving authentic intimate images, and bearing in mind that forced nudification and sexual objectification are on the continuum of sexual abuse which necessarily implicates intrusion, it is easy to see how at least dissemination of intimate deep fakes but probably their creation as well (with the exception of creation for self-use discussed below) could, and indeed should be conceptualised as intrusion. Another way to put it is that on the desired inaccess model of a right to privacy and its connection with the ideas of self-presentation or psychological spatial privacy, the unauthorised gaze at the plaintiff's intimate image is intrusive even in the image if fake. Indeed, on Hariharan's account of physical privacy, an intimate deepfake might undermine physical privacy. According to his account, interference with physical privacy occurs, where someone's interest in bodily integrity is interfered with through the *senses*, in particular watching or listening to the claimant without their consent.²³ The creation of an intimate fake would undermine physical privacy to the extent that the construction and dissemination of a false image of the plaintiff necessarily requires using a true image of them without their consent, so would occur at least where nudification apps are used. I note, however, that this account much resembles a publicity right model, and as such is not necessarily limited to, or inextricably linked with the fact that the image is intimate.

Interestingly, in what seems to be the first intimate fake civil litigation in US jurisdictions (created and shared by a classmate of the 14 y/o plaintiff), the common law torts relied on (alongside statutory causes of action) were public disclosure and intrusion but not false light or defamation.²⁴

²² *Id.*

²³ Jeeven Hariharan, *Physical Privacy and Bodily Integrity*, in *NEW DIRECTIONS IN PRIVATE LAW THEORY* 227 (2023).

²⁴ *Jane Doe v. K.G.* (USDC NJ filed 5 Feb 2024). Litigation around intimate fakes is still nascent. In 2024, Italian Prime minister sued in a Sardinian court a father and son for making and uploading an intimate fake of her to a US pornography website. The cause of action is (criminal) defamation and Maloni sues for 100,000 Euros: *Giorgia Meloni: Italian PM Seeks Damages over Deepfake Porn Videos*, BBC NEWS, Mar. 20, 2024, <https://www.bbc.com/news/world-europe-68615474>.

3 Fictional Characters and Look-Alikes

In this section I argue that tort law's focus on remedying concrete harms to concrete individuals rather than on the regulation of risk, education, or promoting virtuous behaviour suggests that the creation and dissemination of non-photographic intimate images of fictional women should not lead to civil liability (and as an aside, probably not to criminal responsibility). However, as with the case of pseudo photographs,²⁵ there should be liability if the intimate image of the fictional character is reasonably understood to be the plaintiff's – a real look-alike woman.

The law regulating child sexual abuse images – in both the UK and the US – criminalises also the making and sharing of non-photographic images.²⁶ Such legislation, which is the subject of forceful critique,²⁷ cannot plausibly be justified by the harm principle. At best, it hinges on a contested empirical claim that exposure to such images (presumably also for self-use) is likely to cause those exposed to them to commit offences against real children (either directly sexually abusing them or consuming child's intimate image abuse). The contested empirical claim – which echoes the one about the relationship, if any, between pornography consumption and violence against women²⁸ – is also exposed to a critique coming from criminological quarters that proponents of prohibition of non-photographic images confuse paedophilia, which is a sexual attraction to pre-puberty children with child sexual abuse, which is the acting on this proclivity by sexually abusing (real) children.²⁹ More likely, the criminalisation of synthetic images is based on virtue ethics

25 In cases of pseudo photographs superimposing on X's naked body Y's face, a different question arises: whether X (and not only Y) could sue for breach of her privacy or in defamation. I do not discuss this issue here; the answer will hinge mainly on whether the plaintiff in a breach of privacy case needs to be identifiable but also on whether X had a reasonable expectation of privacy with regard to her intimate image prior to its use by the defendant.

26 *Supra* note 4.

27 Suzanne Ost, *Criminalising Fabricated Images of Child Pornography: A Matter of Harm or Morality?*, 30 *LEGAL STUD.* 230 (2010).

28 Gemma Mestre-Bach, Alejandro Villena-Moya & Carlos Chiclana-Actis, *Pornography Use and Violence: A Systematic Review of the Last 20 Years*, 25 *TRAUMA VIOLENCE ABUSE* 1,088 (2024) (concluding that '[A]n association between pornography use and nonsexual violence seems to exist, although the causality of this association remains unclear... The association between different types of violence and pornography use requires further in-depth research to better understand the specific link between both constructs'). For a powerful summary of the harms of pornography based partially on social scripts theory and pornography's impact on sexual autonomy and expression rather than causal effects see Clare McGlynn, Lorna Woods & Alexandros Antoniou, *Pornography, the online Safety Act 2023 and the Need for Further Reform*, 16 *J. MEDIA L.* 211, 215–17 (2024).

29 See e.g., Rebecca Lievesley et al., *Fantasy Sexual Material Use by People with Attractions to Children*, 25 *CURR PSYCHIATRY REP.* 395 (2023); Craig A. Harper & Rebecca Lievesley, *Exploring the*

(what critiques might refer to as an instance of moral panic): the idea that making, sharing, or possessing such images is detrimental to the leading of a well-rounded flourishing life.³⁰ In the context of adult NCII, initially and currently the amended offences have not criminalised the making of deep fakes,³¹ but the government intends to amend the law to criminalise deep fake ‘which appears to be a photograph or a film’ thus excluding non-photographic images.³² Whatever one’s view on the criminal side of things is, it is clear that tort law has no business imposing liability for the creation of intimate images of fictional characters, as there is no ‘real’ victim whose interests – in privacy, reputation or otherwise – have been undermined. Indeed, even when non-photographic intimate images are criminalised, as with the case of child sexual abuse imagery, it is impossible to find and hard to imagine civil claims in the absence of victims.

The position with respect to look-alike should be different,³³ and follows from my position that liability for breach of (at least sexual) privacy should be, and is likely to be, strict.³⁴ Note, however, that an English defamation High Court decision might suggest otherwise. In *O’Shea v. MGN*³⁵ Morland J dismissed a claim brought by the claimant, a look alike of a model posing for a pornographic ad. The court reasoned that article 10 ECHR right to freedom of expression should prevent applying to images the entrenched common law rule (developed with regard to publications involving text) according to which the ‘reference to the plaintiff’ test in defamation law is strict. Beyond the facts that the cogency and desirability of the result in *O’Shea*

Ownership of Child-Like Sex Dolls, 51 ARCH SEX BEHAV. 4,141 (2022); Sara Jahnke, *The Stigma of Pedophilia: Clinical and Forensic Implications*, 23 EUR. PSYCH. 144 (2018).

³⁰ My critique here is similar to Ost, *supra* note 27. It might go beyond, in that Ost might be read as justifying a prohibition on *creation* of fabricated images on the harm principle. However, Ost does not focus explicitly on creation for self-use and implicitly finds criminalising the creation of completely fabricated images of fictional/fantasy children hard to justify. Indeed, supporting criminalisation of the making of fabricated images for self-use (as distinct from creation accompanied with an intention to distribute) is hard to reconcile with Ost’s critique of the Act’s criminalisation of possession.

³¹ Sexual Offences Act 2003, c. 49, § 66 B-66D (UK). This followed stakeholders’ disagreement in response to the consultation paper. The recommendation was not to criminalise the making of an intimate image without the consent of the person depicted. Law Commission, *Intimate Image Abuse: Final Report* (HC 326, 2022) 4.176–4.220.

³² Crime and Policing Bill 2024–2025, HC Bill [187], sch. 8, part 1, 5(3) (as introduced) purports to add to Section 66D of the Sexual Offences Act 2003 sub-section (4 B) which will include in the relevant definition of ‘photograph’ deep fakes. See also Ministry of Justice & Alex Davies-Jones MP, *Government Crackdown on Explicit Deepfakes*, GOV.UK (Jan. 7, 2025), <https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes>.

³³ Suzanne Ost’s position, *supra* note 27 is similar, in the context of non-photographic child image abuse depicting real children.

³⁴ KEREN-PAZ, *supra* note 6, 54–85, 110–131, 175–93.

³⁵ [2001] EWHC 425 (QB).

could be doubted,³⁶ and that more recently, the Supreme Court in *Lloyd v. Google* mentioned in obiter that liability in misuse of private information is strict,³⁷ O'Shea could be distinguished. There, a real person existed (the model) who was the foreseeable target of the publication and with respect to whom the publication was neither defamatory nor false. So a publication which was intended to be about X (who was hence the foreseeable reference), consented to by X and benefited her did not lead to liability towards an unforeseeable Y (the claimant). In the typical fictional character intimate fake case there is no X with respect to whom the image is true – as by definition the image is created – so there is no freedom of expression interest in protecting speech which is non-tortious towards a *real* X (and in fact benefits them) and is *true*: the publisher intends to depict X as a *fictional* nude character.

4 Known Fakes: the Case for Privacy-Based Liability

I turn now to the most controversial proposition (especially for American ears) which I make somewhat provisionally: that known fakes should trigger liability and that such liability could be understood as grounded in an extended notion of privacy. The questions 'whether certain harms should be remedied' and 'if so, how' are distinct. The latter question might involve both fit and justification considerations³⁸ which could roughly be translated to doctrinal and conceptual considerations. Part of the issue is that known fakes are a complex phenomenon which could not easily be compartmentalised as belonging to a certain analytical or conceptual box. The normative question – in our case, whether the sharing (and possibly viewing) of known fakes should lead to liability – is more important than the question, if so how. If liability is justified, one can opt for a *sui generis* kind of liability which will usually involve a statutory civil liability cause of action, courts developing a new common law tort, or courts expanding (or re-interpreting) one or more existing causes of action to provide a remedy to the new situation. And the choice between these options could take into account (1) access to justice and costs, (2) likely effectiveness, (3) conceptual neatness and coherence; and (4) law's expressive message.

³⁶ The reasoning's cogency is limited, as the distinction between images and text could have and has worked in the opposite direction of increasing protection to plaintiffs in cases of images. Eg., *Terry v. Persons Unknown* [2010] EWHC 119 (QB) [55]; Rebecca Moosavian, *Stealing 'Souls'? Article 8 and Photographic Intrusion*, 69 N. Ir. LEGAL Q. 531 (2018). The result is undesirable precisely as, at least in the sexual context, a gendered harm is left without an adequate remedy.

³⁷ [2021] UKSC 50 [133].

³⁸ RONALD DWORKIN, *LAW'S EMPIRE* 53 (1986).

4.1 Known Fakes – Harmful and Wrongful

So why are known fakes harmful and wrongful? I will start with an example which happened in Israel and was never litigated. For reasons explained below, it is a hard case as it gets to impose liability since (1) it involves political speech; and (2) the level of intimate intrusiveness was relatively low. In January 2015 a promotion leaflet for fringe rock bands performance appeared on Facebook.³⁹ It depicted the then member of parliament (and later a minister) Ayelet Shaked, a right wing conservative female politician, sitting semi-clad, underwear on her shins and her breasts half covered by a dress. She wore an armband with some resemblance to a swastika (which was described in an article covering the controversy as a hybrid of the Jewish police in Ghetto and the image of Shaked's political party). The posture and context could be interpreted as suggesting that her consent to sex is in doubt – she is potentially under substance-influence. Alternatively, the context suggests that she is 'a space whore', as the headline reads 'A Space Whore presents:' which is followed by a deliberate use of a homophone changing the conventional 'Young Bands Evening' to 'An Evening to Beat Up Young Women' [*Erev Lehakot Tseh-ee-rot*].⁴⁰ The sex with aliens meaning is strengthened by an image of variety of characters – some human looking, others not – standing on a phallic looking spaceship/drill in proximity to Shaked. The defamatory meaning of Shaked as a whore is strengthened by a sub-title using another pun to describe the stage [*Bama*] as either 'unguarded/penetrable/open to all' or 'a prostitute' [*prutsa*].

The pamphlet/ad is problematic on several grounds and, I argue, is harmful to both Shaked and women in general (including other female politicians or public figures). First, building on societal double standard about sexuality it uses Shaked's gender in order to humiliate and silence her. Second, this speech is discriminatory since a possible effect and purpose of sexualising and 'nudityfying' women is deterring women from equal civil and political participation.⁴¹ It is a tactic that is used disproportionately against women (and other minorities);

³⁹ Not without hesitation, I've opted for not referring to the news article covering the controversy, as it includes the offending image.

⁴⁰ The noun 'bands' [*lehakot*] is spelt in Hebrew with פ, while the verb 'to beat' [*lehakot*] is spelt with כ (both letters create the phoneme k). The leaflet used the latter spelling, making the qualifier [*Tseh-ee-rot*] a noun/direct object – young women – rather than an adjective for bands – 'young' or 'new'.

⁴¹ SARAH SOBIEAJ, CREDIBLE THREAT: ATTACKS AGAINST WOMEN ONLINE AND THE FUTURE OF DEMOCRACY (2020); *online Violence against Women Parliamentarians Hinders Democracy, and All Parliamentarians Are Responsible for Addressing It*, COMMONWEALTH PARLIAMENTARY ASSOCIATION UK (Jun. 14, 2021), <https://www.uk-cpa.org/news-and-views/online-violence-against-women-parliamentarians-hinders-democracy-and-all-parliamentarians-are-responsible-for-addressing-it>.

and when it occurs, it harms women more than men (due to double standards about sexuality) and it targets women *because* they are women and based on a gendered distinction. To understand this, one needs only contrast the swastika-like imagery and the nudification. The former is a fair (if vile) attack both because Shaked's political views could be honestly believed to be fascist and – important for current purposes – the evocation of Shaked's alleged fascism is not discriminatory based on her gender. There is no ground to believe that right wing women are subject more than men to accusations they are fascist and even less, that the accusation made against Shaked that she is fascist is linked to her sex.

Third, both the semi-naked image and the posture alone and the innuendo that she is either a whore or was subject to sexual abuse is defamatory in the sense that people are likely to think less of her (even if they ought not to). I gloss over here three important issues revealing the potential inadequacy of defamation law to deal with these scenarios: (1) To the extent that the test for defamatory meaning is strictly normative ('right thinking people') rather than empirical (what people really think), harm from existing prejudice which is prevalent but unjustified – such as double standard about sexuality – might leave a claimant without a remedy. (2) Cognitive psychologists found that the statement 'Jeanie did not slap her little brother' caused comparable reputation loss to the statement 'Jeanie slapped her little brother', and that innuendos based on concrete statements ('did not hold up a gas station') were regularly more effective than those based on abstract ones ('was not cruel').⁴² Defamation law has no solution for this problem both because the statement 'did not slap her little brother' does not have defamatory meaning and also because it is likely to be true.⁴³ (3) Images like Shaked's and more generally known fakes are likely to be considered as parody and hence as carrying no defamatory meaning since the audience is not likely to think of them as (true) facts.⁴⁴ A recent study has found, however, that participants evaluated a criticized individual more negatively following satire compared to direct criticism and when leaving comments in

42 Daniel M Wegner, *Innuendo and Damage to Reputations*, 11 *ADVANCES IN CONSUMER RSCH.* 694, 694–5 (1984).

43 If the statement is understood by the readers to convey the reverse of the literary meaning, it could lead to liability based on false (or true) innuendo. But the gist of Wegner's findings is that the subject's reputation was affected despite that the readers believe the negation qualifier.

44 In England see the discussion below of *Charleston v. NGN* [1995] 2 AC 65. In the US, *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46 (1988). But see George, *supra* note 1 who, based on *Burton v. Crowell Publ'g Co* 82 F.2 d 154 (1936) argues that liability in defamation for known fakes is possible, notwithstanding viewers' knowledge that the image is synthetic and that, as long as the image is realistic, it ought not be considered as parody. In some US jurisdictions, general deep fake statutes protect deepfakes made for satire of parody. See Kadri & West in this special issue.

YouTube, they used more dehumanizing language in response to satirical versus critical videos.⁴⁵

Fourth, the publication is at best flippant about violence against (young) women; at worst it incites such violence. The context also supports the interpretation that sexual violence is condoned. Whether this condones violence against Shaked herself is debatable but at least possible.

Shaked's case study is atypical in that the publication did not trigger a pile-on harassment (group of harassers targeting the same victim, often following a 'trigger' harassment).⁴⁶ It did not seem to have negatively affected her political career (a few months later she became a minister) and her spokesperson's response was measured and terse: 'I wish [the publishers of the pamphlet] good mental health'. But pile-on harassment often occurs after the circulation of intimate fakes (or authentic NCIs) and there is no indication that it is limited to videos or images perceived to be authentic. The use of online harassment with a sexual nature and the use of rape and death threats or sexual epithets – often insinuating that the target is a whore sleeping with the enemy is a sad reality of female politicians, journalists and public figures.⁴⁷ Intimate fakes are both one manifestation of and a trigger for further pile-on harassment. In the high-profile case of Indian investigative journalist Rana Ayyub, an intimate fake became viral with the intention and partial effect of silencing her critique of the government (and with considerable pile-on harassment) despite the fact that while her face was shown 'I could tell it wasn't actually me'. Ayyub was hospitalised with heart palpitations and anxiety and consequently became 'much more cautious about what I post online. I've self-censored quite a bit out of necessity'.⁴⁸ Indeed, one recent attitude study found that 'Even deepfakes labeled as deepfakes were viewed as blameworthy, harmful, and deserving of punishment'.⁴⁹ As early as 1968 Nimmer observed astutely in the context of intimate fakes that:

45 Hooria Jazaieri & Derek D Rucker, *Softening the Blow or Sharpening the Blade: Examining the Reputational Effects of Satire*, 154 J. EXPERIMENTAL PSYCH.: GEN. 1,201 (2025). Another research found that more informed people are likely to be influenced less from a satire than the less informed. Meital Balmas, *When Fake News Becomes Real: Combined Exposure to Multiple News Sources and Political Attitudes of Inefficacy, Alienation, and Cynicism*, 41 COMMUN RSCH. 430 (2014).

46 The Law Commission, *Modernising Communication Offences: Final Report* (HC 547, 2021) 5.70–5.90 recommended against criminalising knowing participation in pile-on harassment.

47 See e.g., S STEVENS, *DIGITAL HARASSMENT OF WOMEN LEADERS: ISSUES AND SOLUTIONS FACTSHEET* (2018) (common experiences include death and rape threats, gender based derogatory comments or insults, 'revenge pornography', cyberstalking, doxing and trolling. Impacts include withdrawal from politics and from online dialogue, emotional distress, economic harms, reduced access to technology and limitations on digital or actual mobility).

48 Rana Ayyub, *I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me*, HUFFPOST (Nov. 21, 2018), https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316.

49 Kugler and Pace, *supra* note 1, 643.

The sensibilities of [a] young lady whose nude photo is published would be no less offended if it turned out that her face were superimposed upon someone else's nude body. The resulting humiliation would have nothing to do with truth or falsity. The unwarranted disclosure of intimate 'facts' is no less offensive and hence no less deserving of protection merely because such 'facts' are not true'.⁵⁰

What both attitude studies and the case of Ayyub teach us is that even if the intimate image is known to be fake, it is still both wrongful and harmful.

4.2 Known Fakes, Parody and Defamation

Known fakes combine harassment, emotional distress, reputational harms, intrusion into private sphere, gender-based (and often intersectional) discrimination, and interference with political participation (broadly defined), and hence also with democracy. The common argument (at least in the USA) against providing legal remedy against harassment of public figures, including specifically, known fakes is the marketplace of ideas justification to freedom of expression: that the test of the truth or acceptance of ideas depends on their competition with one another and not on the opinion of a censor.⁵¹ Arguably, since such images are a parody, they are both an opinion and understood by the audience not to be a (true) fact and hence should not lead to liability under defamation, which is indeed the law.⁵² There is a relative dated English defamation authority – *Charleston v. NGN*⁵³ – suggesting that known fakes have no defamatory meaning where the image was accompanied by text clarifying it is neither real nor done with the claimants' (actors known for their role in a popular TV show) consent. The image was produced by the makers of a pornographic computer game by superimposing the faces of the plaintiffs without their knowledge or consent on the bodies of others and the article as a whole was critical of the makers of the game. It was conceded by the plaintiffs that if considered as a whole the publication is not defamatory, so the case turned on the court's refusal to give legal effect to the fact that sub group of readers would not read the text so for them the publication might bear a defamatory meaning. However, the protected interest in

⁵⁰ *Supra* note 17.

⁵¹ *Abrams v. US* 250 U.S. 616 (1919). The truth defence is not commonly cited in the UK and where it is, is generally taken to be Mill's version rather than Holmes's conception (although in *R v. Secretary of State for the Home Department* [1999] 3 WLR 328 both are mentioned). See Paul Wragg, *Mill's Dead Dogma: The Value of Truth to Free Speech Jurisprudence*, PUB. L. 363 (2013). Another justification to reduced protection is that politicians must demonstrate a higher level of tolerance to criticism than ordinary individuals.

⁵² See *supra* note 44 above.

⁵³ *Id.*

Charleston is different, the decision was pre Human Rights Act 1998 incorporation the European Convention on Human Rights as relevant in developing English privacy and defamation law in a compatible way,⁵⁴ the court mentioned that it was not invited to consider whether the publication of the photographs by itself constituted some novel tort and English law shows both a trajectory of increased protection to privacy in balancing it against freedom of expression and a shift of mode of protection from defamation to privacy. Beyond these factors, an additional point in *Charleston* was that the publisher, albeit somewhat hypocritically, criticised the creation of these images by the game makers, so the point was not to criticise the plaintiffs, who were indeed captioned as ‘victims’ by one of the sub-titles. This has bearings to defamatory meaning and could be compared with the relevance of tone and purpose of specific uses of otherwise tortious text or images in determining whether the publication is justified on grounds of public interest in the contexts of defamation,⁵⁵ privacy⁵⁶ and data protection.⁵⁷

4.3 Consumer Deception in the Marketplace of Ideas

The marketplace of ideas analogy alerts us to the fact that regulating speech – including the more specific questions of platforms responsibility and intimate fakes – is yet another site of the ideological battlefield between a right wing laissez faire approach and a left-leaning pro regulation approach. If indeed a lie can travel halfway around the world before the truth puts on its shoes – as is echoed in research about mis- and disinformation⁵⁸ – and if market failure is a justification for regulation, then a case for regulating some sorts of ‘unfair’ speech emerges. The

54 In *VBK v. Austria* (2007) ECDR 7 a politician’s face was depicted on the body of a man gripping an ejaculating penis – and made to look as if ejaculating on the face of Mother Theresa. The domestic decision in favour of the politician was held to be in breach of Art 10 although it is not entirely clear whether this is a principled decision on the facts or else a question of proportionality.

55 *Economou v. de Freitas* [2018] EWCA Civ 2,591 (focus of publications on criticising the prosecution which only tangentially (but inevitably) was defamatory towards the plaintiff relevant for affording defendant with public interest defence).

56 *Campbell*, *supra* note 9 [59] (Lord Hoffman dissenting) (relevance of tone); Law Commission IIA, *supra* note 31, 11.49, 11.61 (legal proceedings), 11.97 (administration of justice), 11.107 (genuine medical, scientific or educational purpose) 11.177 (public interest).

57 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), L 119/1 O.J. (2016) § 85.

58 Soroush Vosoughi, Deb Roy & Sinan Aral, *The Spread of True and False News online*, 359 Sci. 1,146 (2018).

marketplace of ideas rationale alerts us to the truism that ideas compete against each other. In commercial settings, liberal legal systems developed a well-established distinction between fair and unfair competition, subjecting the latter to a rigorous set of regulation and potential liability. Importantly, such regulation is based on two distinct rationales: the prevention of consumer deception and curtailing monopolistic power.⁵⁹ Known fakes belong to the former category.

On the non-privacy justification for regulating known fakes there are several ways to theorise, conceptualise and justify such regulation. First, competition between ideas needs to be fair. If, as the findings of Wegner and others suggest, associating claimants with nudity or with sexual activity debases them, and if, further, there is nothing in their approach (e.g., previous statements on the matter) which merits such association, the practice is unfair.⁶⁰ One way to look at it, is that we all have an interest that opinions will be formed on a rational basis. Debasing a person in a way which is not based on a relevant view undermines rational forming of opinion and is as such undesirable and unfair.⁶¹ Arguably, this is especially problematic since the attack is *ad hominem* (or in the intimate fakes context, mainly *ad feminam*) and even more so, since it relates to immutable characteristics and is, in a sense, embodied. Secondly, as ideas compete against each other, debasing a speaker based on irrelevant factors will always give an unfair advantage to ideas competing to those held by the speaker, and might also provide an unfair advantage to a competing speaker. For example, a politician who is the subject of a known fake might lose popularity or even an election campaign to a rival politician. Thirdly, we should not lose sight of the fact that both the phenomenon and the harm are systematically gendered; leaving the problem unaddressed is a form of gender-based discrimination both in general and in terms of political participation. Next, freedom of expression and especially political expression are civil and political rights. But at stake here are not only the civil rights of those disseminating intimate fakes but also of those depicted in them. To the extent that female politicians, for example, fail to get elected because of an intimate fake, or decide to quit, their civil rights are undercut.

59 See Dev Saif Gangjee, *Law(s) against Unfair Competition: The Legitimacy of the UK Approach*, 19 J. INTELL. PROP. L. & PRAC. 90 (2024). Analytically, most relevant is the malicious falsehood tort which is both an instance of business libel (although harder to prove) and the functional equivalent of Section 6 of the Restatement (Third) Unfair Competition. The idea that the harm caused from deep fakes is not only to the person depicted but also to consumers of the deep fake features in some general statutory responses to deep fakes in some US jurisdictions. See Kadri & West's contribution in this special issue.

60 Compare Paul H Brietzke, *How and Why the Marketplace of Ideas Fails*, 31 VAL. U. L. REV. 951, 966 (1997).

61 On the narrow account, the relevant wrong is the debasing (leading to an irrational forming of opinion). On the broader account, the wrong is the undermining of rational decision making. Both versions are incompatible with prevailing freedom of expression jurisprudence.

Broader still, a failure to regulate known fakes is a form of backlash against all high profile and indeed all opiniated women; as such known fakes create a strong (and gendered) chilling effect on women's participation in public life and on women's freedom of expression. Finally, when known fakes cause pile-on harassment such harassment exacerbates the problems noted above; this provides a further justification to regulate known fakes.

4.4 Private Law Remedy for Known Fakes? A Matter of Privacy?

From the fact that known fakes should be regulated, it does not necessarily follow that there should be a private law remedy. However, when one bears in mind that the harm from known fakes and other instances of intimate image abuse is similar, the case for affording a remedy for the former is strong. Moreover, while some of the harms are abstract, diffused and arguably affect other individuals, the plaintiff depicted in the known fake suffers the main bulk of the harm. This justifies a private law remedy. The harm is dignitary and as such is often presumed and is given as general damages (as in damages at large in defamation). To the extent that specific harms can be proven, special damages can be awarded as is the case in both defamation and breach of privacy.⁶² If at all, the fact that societal harm exceeds the claimant's should lead to the award of punitive damages; it should certainly not serve as a reason to deprive the principal victim from the activity from receiving a remedy.

I tried to establish that there are good reasons to remedy the harms caused by known fakes to those depicted in them. But can these harms be conceptualised as undermining an interest in privacy? On the gene/allele understanding of privacy I think the answer is yes, although I am not certain. When a private 'fact' is false, but is not declared as such, it is likely to be understood as true. So in this sense, it lifts the veil behind which the private information is supposed to be inaccessible to others.⁶³ As, on this understanding, the topic of discussion itself is private, any instantiation given to the topic is a breach of privacy, at least when it is not presented as a false claim. The problem with extending the argument to known fakes is that such liability might serve as a form of censorship and undermine the important distinction between facts and opinions. Strictly speaking, a fact qualified by the speaker as false is not an opinion, but the rationale behind curtailing liability with respect to opinions compared to liability for facts is (at least partially) relevant here as well. First, a fact known to be false presumably has lesser potential to harm its subject. Indeed, to the extent that Wegner's findings cast doubt on such assumption, the rationale's cogency

⁶² ReachLocal UK Ltd v. Jamie Bennett [2014] EWHC 3405 (QB); Richard v. BBC [2018] EWHC 1837 (Ch).

⁶³ See *supra* note 12.

for exempting false facts announced as such from triggering liability is significantly lessened. Secondly, since opinions are necessarily subjective, they seem to implicate one's interest in freedom of expression, dignity and autonomy to a greater extent; therefore, repressing opinions seems less justified and more harmful, or even oppressive.

It is here that the comparison with facts declared to be false is contested. On one understanding, images declared (or obviously passing) as fake belong to the realm of opinions. They either manifest wishful thinking (or a form of fantasy) or express a negative opinion, perhaps by the use of innuendo, on the depicted person. Elsewhere, I explain that the law of privacy should avoid the pitfall of using the public interest in commenting about the desirability of an otherwise private behaviour as a reason to 'out' that behaviour. To the same extent that in defamation law an opinion would be defamatory if it is based on defamatory facts (so unless the underlying facts are true of privileged), so an opinion which is based on private facts should lead to liability for misuse of private information (unless the interest in the publication of the underlying fact is different from that of merely criticising that fact).⁶⁴ It follows that even on the understanding of known fakes as a form of opinion, to the extent that these opinions are based on private facts – the claimant's naked body – and bearing in mind that private facts include also false facts, known fakes should be considered as a breach of privacy. Given social taboos about nudity and sexuality, and when the interest in privacy is understood to encompass reputation and freedom from harassment⁶⁵ there is a considerable overlap between (1) the policies discussed above justifying remedying known fakes on grounds of anti-harassment, public participation and anti-discrimination and (2) the policies behind treating known fakes as a privacy breach. Understood this way, it is easy to see why one can consistently oppose regulating intimate fakes created for self-consumption since they amount to a form of opinion (fantasy) while supporting regulation of known fakes. The difference is that publication exists in the latter but not in the former.

Arguably, the gist of parody intimate fakes is in insulting their subject and reducing their reputation and not in revealing a private fact about them or intruding into their physical privacy so any response should lie not in privacy but possibly in defamation, anti-harassment or infliction of emotional distress. The account I offer explains why this conclusion should be resisted. First, that the publication is also harassing, defamatory, and an instance of gender discrimination does not prevent it from being also a breach of privacy. Authentic non-consensual intimate images are also a multifaceted phenomenon involving sexual abuse, discrimination, harassment, reputational losses and breach of privacy but are accepted as a serious privacy

⁶⁴ Keren-Paz, *supra* note 10,180–84.

⁶⁵ *Richard*, *supra* note 62; *Von Hannover v. Germany* (no 1) [2004] EMLR 379.

invasion. Ultimately, on the equal harm (and wrongfulness) understanding of intimate fakes⁶⁶ it makes little sense to regulate authentic images as a breach of privacy (despite the other motives, effects and ways to conceptualise it) but refuse to do the same for inauthentic images.

Secondly, the harassing, defamatory and discriminatory potential of the publication is ingrained exactly in the fact that the issue depicted is private and taboo. We are all naked at some point of our lives and almost of all of us (if lucky) share the joys of consensual sex. But these moments are private and outing these private moments are detrimental and embarrassing for many; and in our sexist society they are especially detrimental to women and to LGBTQIA + s. Recall the following two empirical findings: (a) associating the subject of the comment with negative activities sticks even when the audience know the statement is false (Wegner); (b) parody harms to a greater extent one's reputation than direct criticism (Jazaieri & Rucker). Recall also the analytical observation that a false fact about a private matter is still a fact about a *private* matter. Combined, these findings and observation support the following conclusion: Debasing someone – especially a woman and especially a female celebrity – based on their sexuality is wrong and harmful to a large extent since it deals with a private matter. Remedying this is the concern of privacy law, notwithstanding that reputational concerns are also paramount. Indeed, as I have recently argued, privacy law should – and in England is settling to do – compensate plaintiffs for their reputational loss even where the private facts are true.⁶⁷ That in intimate deepfakes the private fact is false only makes the award of reputational loss as part of a privacy claim even easier. That deepfakes do not need to rest on true information about their subject to harm them is inconclusive given that the subject matter is private.

Finally, that in some cases we feel that harsh critique and parody are justified even with respect to a private matter does not provide a good reason to oppose a remedy in cases in which the critique is not a fair game. For example, given Trump's predatory sexual behaviour, depicting him naked or engaging in crass or in predatory sexual behaviour is fair. But this goes to the second stage of privacy analysis – of balancing privacy interests with freedom of expression. As the previous discussion of Shaked demonstrates, the Swastika is a fair critique but the sexualisation, less so. The situation might have been different, had Shaked's past behaviour or statements were relevant for the issue of her sexuality.⁶⁸ Had Shaked made a specific statement to the effect that women should be ashamed if they had sex with foreign men, this would arguably make a cartoon of her as potentially having sex with aliens a proper

⁶⁶ See *supra* note 18.

⁶⁷ KEREN-PAZ, *supra* note 6, 200–206.

⁶⁸ See Keren-Paz, *supra* note 10.

satirical response. However, even under this hypothetical, the facts that no hypocrisy is involved (namely, that her behaviour is consistent with her views), the cartoon is more likely to target Shaked because she is a woman (so a male politician making the same comment is less likely to face such a cartoon) and the cartoon is likely to be more harmful to her political career since she is a woman make the cartoon and its effect problematic; and the cartoon still implicates Shaked's privacy interest in her sex life. But the case for liability in this hypothetical is significantly less strong given the public interest in criticising her views.

5 Creation for Self-Use

Where authentic intimate images are concerned, unless the plaintiff entrusted the disseminator with the image (but did not consent to the dissemination) accessing the intimate image is in itself an intrusion into (physical) privacy, whether it is done by hacking, voyeurism or taking the photo without the plaintiff's consent and in circumstances she had a reasonable expectation of privacy. It is for this reason that the viewing of NCII, including fake, should be – and probably is – a breach of privacy in its own right.⁶⁹ But should the creation of fake intimate images for self-use be considered as a breach of privacy and lead to criminal or civil liability? Views of those responding to the Law Commission Intimate Image Abuse Consultation Paper (which focussed on criminalisation) differed. The online Safety Act 2023 initially did not criminalise the making of such images but parliament is in the process of criminalising the making of intimate fakes.⁷⁰ Both conceptually and normatively, the case against regulating the making of such images for self-use is less compelling than in cases of authentic images. Analytically, it is harder to conceptualise intrusion. When a phone is hacked, a clear physical boundary is crossed. Upskirting or downblousing also crosses a physical boundary beyond which some body parts or clothing items are not ordinarily observable. On the other end of the scale, undressing someone you know in your mind (without gazing, and definitely when they are not present) does not cross a physical boundary. Moreover, for those who think that an image which passes as fake does not invoke an interest in privacy – a view I challenged above – the case for recognising a reasonable expectation of privacy in fantasies is difficult, as by definition, the person creating the image in their mind is aware that the image is not real. Making intimate images of another, whether by drawing or digitally, is

⁶⁹ KEREN-PAZ, *supra* note 6, 135–9. As I mention there, this point is lost on some US commentators mistaking the viewing of non-consensual of intimate images only as a matter of public disclosure, rather than of intrusion.

⁷⁰ See *supra* notes 31–32.

somewhere in between. There is a corporeal manifestation – a physical existence of the image – but it is far from clear that the plaintiff's physical boundary was crossed.⁷¹ The stronger objection is normative which is related to the conceptual challenge. The creation of fake intimate images is a form of fantasy and a manifestation of both sexual autonomy and freedom of expression. Moreover, it strongly engages the defendant's privacy both because fantasies in general, and sexual fantasies in particular are a private, intimate and sensitive matter and because the physical manifestation of the fantasy is usually done in a private space. Regulating (sexual) fantasies is both censorial and oppressive and the fact there is a physical manifestation of the fantasy is insufficient.

Most would agree, that concerns about objectification aside, no one has a reasonable expectation of privacy not to be subject of erotic or sexual fantasies even when these fantasies are physically manifested in masturbation and ejaculate. So the fact that the fantasy is physically manifested in an image, rather than in sperm, should arguably not matter. This relates also to the issue of harm. As long as the subject of the fantasy does not know she is such (whether just in the defendant's mind, or with the aid of analogue or digital visualisation) no (serious?) harm is done. Before defending this claim, I would like to highlight the difference between intimate fakes and objectifications which more clearly involve the crossing of physical boundaries.⁷² The latter are certainly wrongful, but if unknown to the victim they are not harmful, or at least much less harmful. For example, someone who unbeknown to them was sexually assaulted while unconscious is clearly wronged. It is often only when they are told (for example by the police) that they were victimised that harm occurs.⁷³ To my mind there is a question here whether there is, in such circumstances, an interest in not knowing.⁷⁴

On the (contested) account that breach of privacy (like the trespass torts) is actionable *per se*, the fact (or assumption) that no harm occurs if the victim is unaware does not prevent the constitution of the cause of action. Moreover, even assuming that being subject to voyeurism or to unconsented bodily contact is harmful even without awareness of the breach of privacy or bodily integrity,

⁷¹ Cf Law Commission IIA, *supra* note 31, 4.214–4.219; contrast with McGlynn and Toparlak, *supra* note 20.

⁷² Recall, that as discussed in Section 2, the creation of non-photographic images might still implicate an interest in physical privacy on a certain account.

⁷³ See Unsigned, *Reynhard Sinaga: "Evil Sexual Predator" Jailed for Life for 136 Rapes*, BBC NEWS, Jan. 6, 2020, <https://www.bbc.com/news/uk-50987823>.

⁷⁴ See Suzanne Ost & Alisdair A Gillespie, *To Know or Not to Know: Should Crimes Regarding Photographs of Their Child Sexual Abuse Be Disclosed to Now-Adult, Unknowing Victims?*, 25 INT'L REV. OF VICTIMOLOGY 223 (2019); and *Recognizing the Paradigm of the Unknowing Victim and the Implications of Liminality*, 64 BRIT. J. CRIMINOLOGY 194 (2024).

communicating to the victim that their right was breached *significantly increases* the harm from the breach. Hence the dilemma whether to communicate this fact to the victim and victims' potential interest in not knowing. The interest in not knowing might even be more pronounced in cases of serial predators (like Sinaga⁷⁵) whose conviction is secured even without notifying unaware victims and in which neither additional imprisonment time nor compensation from the tortfeasor/offender are likely.

The case in support of making the creation of intimate fakes for self-use either a criminal offense or a civil wrong is even more contested, as the objectification there is a manifestation of *fantasy* which less clearly crosses a physical boundary, so its culpability is more contested, compared to cases of voyeurism and physical sexual abuse in which the victim is unaware.⁷⁶ What makes sexual objectification in the form of *fantasy* both wrong and harmful is the *communication* of the fact the person is objectified either to the plaintiff herself or to third parties. So a man fantasising on bedding a colleague (even a subordinate) is neither culpable nor harmful in itself; however, telling her (or him) or another colleague about the fantasy might be harmful and wrongful. There might be another distinction at work here: where a seemingly consensual sexual relationship is nonetheless socially or legally unacceptable, knowledge that the plaintiff is a subject of sexual attention might be more harmful, disrespectful, unprofessional, distressing or undignifying so the communication of the interest might justify some legal response. But here too, it is the communication of the fantasy, rather than its existence which is the problem.⁷⁷

If I am right about that, the creation of deep fakes for self-use should not be criminalised. The dissemination of deep fakes is already criminalised, so if the creator is also a disseminator, he should be both criminally and civilly liable for the dissemination, so having a stand-alone responsibility for the creation is not particularly needed. But my analysis has two important qualifications. First, for the civil side of things I would support strict liability. So if the creation of the deep fake was intentional but the dissemination was not (due to the creator's mistake and arguably also due to a third party's acts) the creator should be liable for ensuing harm (alongside the distributor). I have defended strict liability for breach of sexual privacy in the NCII context elsewhere.⁷⁸ Of importance here, is that intimate fakes have the potential to create devastating harms to their subject if distributed, so the creator

⁷⁵ *Supra* note 73.

⁷⁶ For McGlynn and Toparlak, *supra* note 20, the creation of deep fakes is harmful, wrong and should be criminalised 'regardless of victim knowledge or online distribution'.

⁷⁷ I would apply the analysis in this section also to situations in which a high school student makes a deep fake of his teacher. I thank Suzzane Ost for raising the question while reading an earlier draft of this paper.

⁷⁸ *Supra* note 6.

of the risk (the maker) should be held to account if they were. In other words, those who bring on their devices something which is likely to cause mischief if it escapes should be strictly liable if it does.⁷⁹ Elsewhere I have argued that one should have no lesser remedy with respect to her stolen images than she has with respect to her stolen car. We can add to that, that one should be protected from the harmful consequences of escaped intimate images no less than from the escape of water from neighbouring reservoirs.

The second qualification is that if indeed what is problematic about objectification is not the fantasy but the fact it is communicated to the plaintiff or to a third party, it might be that merely communicating to the plaintiff (or to the third party) that such images were created might justify liability.⁸⁰ Currently the Sexual Offences Act criminalises a threat to disseminate deep fakes⁸¹ but the ramifications of the above analysis go potentially much further in that (1) responsibility is not limited to a threat to share; (2) nor to a motivation to cause fear; and (3) most importantly, the plaintiff's relevant knowledge should be about the creation and existence of the deep fake, not about plans to disseminate it. Whether this interest in not being informed that such an intimate image exists – assuming it is worthy of protection – is about vindicating privacy, freedom from harassment or reputation is a question I cannot resolve here, but it bears resemblance to the brief discussion above in the context of imposing liability for disseminating known fakes.

⁷⁹ Rylands v. Fletcher (1868) LR 3 HL 330.

⁸⁰ Cf George, *supra* note 1, 167.

⁸¹ Sexual Offences Act 2003, c. 49, § 66 B-66D (UK).