

FAKE IMAGES, REAL HARM: A CASE FOR CRIMINALISING NON-CONSENSUAL INTIMATE DEEPFAKES

*Claire Graham**

Advancements in artificial intelligence technologies have led to the perpetration of sexual violence in novel ways. Non-consensual intimate deepfakes (NCIDs) are a pernicious and pervasive form of sexual violence involving the creation, distribution or threatened distribution of digitally manipulated intimate images without consent. The law must be updated to effectively tackle harmful behaviours as they arise. This article aims to identify the most appropriate and effective way for New Zealand lawmakers to address the growing problem of NCIDs, in light of their significant individual and social harms. First, a positive case for criminalisation is made out through an examination of the individual and collective harms caused by NCIDs. Secondly, an analysis of the different available regulatory mechanisms establishes a negative case in support of criminalisation. Finally, this article offers specific recommendations for enacting distinct criminal provisions targeting NCIDs, as part of a broader integrated approach that is both proactive and reactive to the emerging NCID phenomenon.

I INTRODUCTION

In January 2024, a series of sexually graphic "deepfakes" depicting singer-songwriter Taylor Swift went viral on several social media platforms, with one post on X (formerly Twitter) amassing over 45 million views within just 17 hours.¹ This incident garnered widespread outrage from fans, the public and even the White House.² However, just two months later, Meta's platforms hosted an ad

* Submitted for the LLB (Honours) Degree, Faculty of Law, Te Herenga Waka—Victoria University of Wellington, 2024. I want to extend my deepest and sincerest gratitude to my supervisor, Dr Zoë Prebble, for her endless and unwavering support—both academic and emotional. Her thoughtful guidance, encouragement and generosity of spirit was a bright light in every moment of doubt. I am incredibly lucky to have had such a brilliant and compassionate mentor.

1 Vittoria Elliott "Celebrity Deepfake Porn Cases Will Be Investigated by Meta Oversight Board" (16 April 2024) [Wired](http://www.wired.com) <www.wired.com>; and Alice Witt and Nicola Henry "Taylor Swift deepfakes: New technologies have long been weaponised against women The solution involves us all" (1 February 2024) [The Conversation](http://www.theconversation.com) <www.theconversation.com>.

2 Witt and Henry, above n 1.

campaign featuring an underage Jenna Ortega, promoting a deepfake app that enabled users to digitally undress her using artificial intelligence (AI).³

Deepfakes are digitally altered images or videos created using AI "deep learning" techniques, where one person's face is transposed onto another's.⁴ Pornographic deepfakes first appeared in 2017, when Reddit user "u/deepfakes" posted several deepfakes created with open-source AI technology, which "stitched" the faces of female celebrities onto pornographic videos.⁵ While celebrity cases attract significant media attention, the increasingly sophisticated nature of AI means that anyone can be targeted.⁶ Deepfakes can cause significant individual, social and cultural harms. While several jurisdictions have responded to deepfake threats with legal reform, New Zealand is yet to take legislative action.⁷ However, New Zealand is not immune from this deepfake phenomenon; Netsafe increasingly receives reports of deepfake content.⁸ The critical question is: what is the most appropriate and effective way for New Zealand lawmakers to address the growing problem of non-consensual intimate deepfakes (NCIDs)?

To that end, Part II of this article contextualises NCIDs as an emerging form of technology-facilitated sexual abuse demanding urgent social and legislative attention. Part III explores the harms to, and personal consequences for, individual victims of NCIDs, and their broader societal impacts. Part IV analyses potential regulatory mechanisms and argues that distinct criminal provisions are necessary to effectively prevent and respond to the kaleidoscope of NCID experiences. Part V

-
- 3 Meg Walters "Amid Jenna Ortega underage deepfakes, why isn't social media taking it seriously?" (6 March 2024) *Glamour UK* <www.glamourmagazine.co.uk>; and Alex Ocho "Underage Jenna Ortega Deepfake Ads That Let Users 'Undress' Her Ran on Instagram and Facebook" *COMPLEX* (6 March 2024) <www.complex.com>.
- 4 Asher Flynn and others "Deepfakes and Digitally Altered Imagery Abuse: A Cross-Country Exploration of an Emerging form of Image-Based Sexual Abuse" (2022) 62 *Br J Criminol* 1341 at 1341; and Chidera Okolie "Artificial Intelligence-Altered Videos (Deepfakes), Image-Based Sexual Abuse, and Data Privacy Concerns" (2023) 25 *J Int Women's Stud* at 1–2.
- 5 Stephanie Tong "'You Won't Believe What She Does!': An Examination into the Use of Pornographic Deepfakes as a Method of Sexual Abuse and the Legal Protections Available to its Victims" (2022) 25 *UNSWLawJlStuS* 22 at 2; and Rebecca Umbach and others "Non-Consensual Synthetic Intimate Imagery: Prevalence, Attitudes, and Knowledge in 10 Countries" [2024] *Proceedings of the CHI Conference on Human Factors in Computing Systems* 1 at 1.
- 6 Witt and Henry, above n 1. See for example Noelle Martin "Image-Based Sexual Abuse and Deepfakes: A Survivor Turned Activist's Perspective" in Anastasia Powell, Asher Flynn and Lisa Sugiura (eds) *The Palgrave Handbook of Gendered Violence and Technology* (Springer International Publishing, Cham, 2021) 55.
- 7 Bella Stuart "Taylor Swift, Non-Consensual Deepfake Pornography, and What It Means for New Zealand" (7 February 2024) *Brainbox Institute* <www.brainbox.institute>.
- 8 Finn Hogan "'Our eyes now betray us': Experts concerned over rise of deepfake technology" *Newshub* (29 October 2022) <www.newshub.co.nz>; and Stuart, above n 7.

makes specific recommendations regarding the enactment of such provisions. This article argues that although the criminal law alone will not solve this socio-legal problem, it can shape social attitudes and norms before novel and harmful conduct becomes ingrained in the increasingly digital social world.

II THE ISSUE: WHY DEEPPAKES DEMAND URGENT ATTENTION

A The Issue in Context

The rapid development of digital technologies such as computers, smartphones and social media has revolutionised communication by breaking down traditional time and distance barriers.⁹ Technological advancements have had beneficial applications. However, they have also expanded and transformed how sexual violence can be perpetrated.¹⁰ Sexual violence is a globally significant human rights issue, with women overwhelmingly identified as victims and men as perpetrators.¹¹ Though sexual violence may typically be envisioned as involving physical force, another form of sexual violence has emerged in the digital sphere.¹²

The term "technology-facilitated sexual violence" (TFSV) describes the growing phenomenon of perpetuating sexually harmful behaviours via communication technologies.¹³ Online abuse is distinct from offline abuse in several ways: (1) the collapse of spatial-temporal barriers allows offenders to target victims faster and across transnational borders; (2) increased connectivity enables the targeting of multiple victims in rapid succession; and (3) online anonymity enables deception, and curbs detection and regulation.¹⁴ Consequently, a TFSV offender can, using comparatively minimal resources and efforts, significantly and detrimentally affect many victims.¹⁵

Image-based sexual abuse (IBSA) is an increasingly criminalised form of TFSV. It is more commonly and colloquially referred to as "revenge pornography," however, many scholars prefer the

9 Bridget A Harris "Technology and Violence Against Women" in Sandra Walklate and others (eds) *The Emerald Handbook of Feminism, Criminology and Social Change* (Emerald Publishing Limited, 2020) 317 at 317; and Nicola Henry and Anastasia Powell "Technology-Facilitated Sexual Violence: A Literature Review of Empirical Research" (2018) 19 TVA 195 at 305.

10 Nicola Henry and Anastasia Powell "Embodied Harms: Gender, Shame, and Technology-Facilitated Sexual Violence" (2015) 21 Violence Against Women 758 at 763; and Harris, above n 9, at 317.

11 Henry and Powell, above n 9, at 195.

12 Okolie, above n 4, at 2.

13 Henry and Powell, above n 10, at 759; and Henry and Powell, above n 9, at 195.

14 Majid Yar "The Novelty of 'Cybercrime': An Assessment in Light of Routine Activity Theory" (2005) 2 Eur J Criminol 407 at 410-411.

15 Henry and Powell, above n 10, at 763.

term "image-based sexual abuse" because it better reflects the various contexts and harms of the abuse.¹⁶ IBSA comprises three main behaviours: the creation of, distribution of, or threatened distribution of nude or sexual photos of someone without their consent.¹⁷ With recent advancements in artificial intelligence, an additional category of IBSA has emerged: the digital manipulation of imagery to create non-consensual intimate deepfakes.¹⁸

B Deepfakes: A Persistent and Progressing Problem

AI technological advancement has created novel and innovative tools that allow sexual violence to be perpetrated digitally with ever-increasing ease. Within just a few years, tools to create realistic deepfakes have been rapidly developed and simplified—a process that once took weeks or months now takes mere minutes.¹⁹ In 2019, just 15 seconds of an Instagram story was enough to create a reliable deepfake, and AI technology has only become more advanced.²⁰ Deepfakes create convincing illusions of a person's presence and conduct. Since faces are intrinsically associated with a person's identity, this can have serious consequences for the subjects of deepfakes.²¹

1 Narrowing the scope: Non-consensual intimate deepfakes

Complex legal issues continue to arise as a consequence of AI advancement. Deepfakes can engage numerous areas of law, impacting copyright and intellectual property, fraud and misrepresentation, privacy and security, and can affect democratic processes, such as election campaigns and public debates, due to their potential use for spreading misinformation. As AI continues to evolve, these issues will require legal regulation. However, this article is particularly concerned with how the law should respond to sexual violence facilitated by AI technologies.

Without commenting further on the need for broader deepfake regulation, this article argues that the creation, distribution and threatened distribution of a particular subset of deepfakes—

16 Clare McGlynn and Erika Rackley "Image-Based Sexual Abuse" (2017) 37 *Oxf J Leg Stud* 534 at 535.

17 At 3; Anastasia Powell and Nicola Henry "Rape Culture Unveiled" in *Sexual Violence in a Digital Age* (Palgrave Macmillan UK, London, 2017) 79 at 12; Nicola Henry, Asher Flynn and Anastasia Powell *Responding to 'revenge pornography': Prevalence, nature and impacts* (Criminology Research Advisory Council 2019) at 1; and Anastasia Powell and others "A multi-country study of image-based sexual abuse: Extent, relational nature and correlates of victimisation experiences" (2024) 30 *J Sex Aggress* 25 at 25.

18 Okolie, above n 4, at 3; and Powell and others, above n 17, at 26.

19 Tong, above n 5, at 3; and Flynn and others, above n 4, at 1342.

20 Elliott, above n 1.

21 Yuezun Li and others "Toward the Creation and Obstruction of DeepFakes" in Christian Rathgeb and others (eds) *Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks* (Springer International Publishing, Cham, 2022) 71 at 72; and Andreea Pocol and others "Seeing is No Longer Believing: A Survey on the State of Deepfakes, AI-Generated Humans, and Other Nonveridical Media" [2024] *Advances in Computer Graphics* 427 at 438.

non-consensual intimate deepfakes (NCIDs)—should attract criminal liability. NCIDs refer to intimate deepfakes created and/or shared without the subject's consent. This term focuses on the absence of consent, aligning with other sexual offences. Consent will be absent in two scenarios: where the subject has not consented to the deepfake's creation, or where they consented to its creation but not its dissemination. This term modifies earlier academic terms like "fake involuntary pornographic deepfakes"²² and "non-consensual synthetic intimate imagery".²³ Adaptation is suggested because "synthetic imagery" is now better known as "deepfakes", and the term "fake" is largely redundant given the inherent nature of deepfakes. Further, "pornography" inappropriately implies victim consent, or at least complicity, and concentrates on whether the images are sexually explicit, rather than whether they harm victims.²⁴ The term "NCIDs" is broad enough to capture content that is intimate but not explicitly sexual, such as partial or non-sexual nudity, which can still cause significant harm.²⁵

2 Significance and scale

In a 2021 survey of 6000 participants across the United Kingdom, Australia and New Zealand, 14 per cent disclosed that someone had either created or shared an intimate deepfake of them without their consent.²⁶ Measuring the prevalence of NCIDs is challenging due to the Internet's expansiveness, and the fact that most images are shared on secret "member only" forums,²⁷ meaning many victims may remain unaware of their exploitation, and even fewer will report it to police.²⁸ Recent research suggests that NCIDs account for approximately 96 per cent of all deepfake videos online, with numbers doubling every six months since their first appearance on Reddit in 2017.²⁹ Accordingly, NCIDs are the most common type of deepfake being requested, created and circulated online. The

22 Nikki Chamberlain "Privacy and Social Media" in Nikki Chamberlain and Stephen Penk (eds) *Privacy - A to Z of New Zealand Law* (online ed, Thomson Reuters, 2023).

23 Umbach and others, above n 5.

24 Clare McGlynn, Erika Rackley and Ruth Houghton "Beyond 'Revenge Porn': The Continuum of Image-Based Sexual Abuse" (2017) 25 *Fem Leg Stud* 25 at 38; McGlynn and Rackley, above n 16, at 535–536; and Anastasia Powell, Nicola Henry and Asher Flynn "Image-Based Sexual Abuse" in *Routledge Handbook of Critical Criminology* (2nd ed, Routledge, 2018) at 306.

25 McGlynn, Rackley and Houghton, above n 24, at 39.

26 Erika Rackley and others "Seeking Justice and Redress for Victim-Survivors of Image-Based Sexual Abuse" (2021) 29 *Fem Leg Stud* 293 at 308.

27 Powell and others, above n 17, at 28.

28 Nicola Henry and others "'Devastating, like it broke me': Responding to image-based sexual abuse in Aotearoa New Zealand" (2023) 23 *Criminol Crim Justice* 861 at 863.

29 Henry Ajder and others *The State of Deepfakes: Landscape, Threats and Impact* (2019) at 1; and Sensity *The State of Deepfakes 2020: Updates on Statistics and Trends* (2020).

unprecedented pace of technological advancement suggests that deepfake production will continue to increase, and its threats materialise, unless decisive action is taken.

III THE HARMS OF NCIDS

A Criminalisation on the Basis of Harm

Criminal offences should only be enacted when absolutely necessary,³⁰ because they significantly infringe on individual liberty.³¹ To justify criminalisation, Andrew Simester and Warren Brookbanks argue that a "positive" and "negative" case should be made out.³² This part deals with the positive case, which requires conduct to be "sufficiently serious", leading directly or indirectly to substantial *harm* to others.³³

This justification broadly aligns with John Stuart Mill's harm principle, which provides that the state's limits on individual liberty can only be justified insofar as they prevent harm to others.³⁴ Mill frames the harm principle as a negative constraint: in the absence of harm, the state is not entitled to use its coercive powers to intervene with individual freedom and liberty.³⁵ Conversely, Joel Feinberg frames harm as a positive ground for criminalisation (and speaks to the negative case):³⁶

It is always a good reason in support of penal legislation that it would probably be effective in preventing ... harm to persons other than the actor ... *and* there is probably no other means that is equally effective at no greater cost to other values.

Harm, or risk of harm, justifies criminalisation of certain conduct where it is the only available mechanism for effectively preventing and addressing such harms. Focusing on the impacts of NCIDs enables lawmakers to craft legislation to provide effective resolutions for victims. Moreover, a harm-centric approach centres victims' experiences, which is most appropriate for sexual offending.³⁷

30 Andrew P Simester and Warren Brookbanks *Principles of Criminal Law* (5th ed, Thomson Reuters New Zealand Ltd, Wellington, New Zealand, 2019) at 1004.

31 At 1003–1004.

32 At 1004.

33 At 1004.

34 John Stuart Mill *On Liberty* (2nd ed, John W Parker and Son, London, United Kingdom, 1859) at 22.

35 Andrew P Simester and Andreas von Hirsch *Crimes, Harms, and Wrongs* (Bloomsbury Publishing Plc, London, United Kingdom, 2014) at 35.

36 Joel Feinberg *The Moral Limits of the Criminal Law Volume 1: Harm to Others* (Oxford University Press, New York, 1984) at 26.

37 Anastasia Powell and Nicola Henry "Beyond 'Revenge Pornography'" in *Sexual Violence in a Digital Age* (Palgrave Macmillan UK, London, 2017) 117 at 119.

This article establishes a positive case for criminalisation on the basis of the significant individual, social and cultural harms caused by NCIDs.

B NCIDs: A Gendered Phenomenon

1 The gendered nature of NCIDs

NCIDs are a distinctly gendered form of IBSA. This is apparent in two main respects. First, unlike more traditional forms of IBSA such as sextortion, voyeurism or the non-consensual sharing of intimate images—which appear to have little to no distinctions between rates of victimisation for men and women³⁸—NCIDs almost exclusively target women and are predominantly produced by and for male audiences.³⁹ Recent studies indicate that approximately 99 per cent of individuals depicted in online NCIDs are women.⁴⁰

Secondly, studies on IBSA more generally have found significant differences between men and women regarding the *nature* and *impacts* of the harms suffered. Consistent with broader patterns of offline abuse, women victims experience particularly pernicious, long-lasting and all-encompassing harms as a result of IBSA.⁴¹ This suggests that the harm stems not only from the images themselves but also from the differential social meanings, and sexual double standards, attributed to men's and women's bodies that induce women to feel shame, blame and humiliation about their sexuality.⁴²

38 See Kate Walker and others "Nonconsensual Sharing of Private Sexually Explicit Media Among University Students" (2021) 36 *J Interpers Violence* NP9078; and Anastasia Powell and others "Image-based sexual abuse: The extent, nature, and predictors of perpetration in a community sample of Australian residents" (2019) 92 *Computers in Human Behavior* 393. Compare Christina M Dardis and Elizabeth C Richards "Nonconsensual Distribution of Sexually Explicit Images Within a Context of Coercive Control: Frequency, Characteristics, and Associations with Other Forms of Victimization" (2022) 28 *Violence Against Women* 3933; and V Karasavva and A Forth "Personality, Attitudinal, and Demographic Predictors of Non-consensual Dissemination of Intimate Images" (2021) 37 *J Interpers Violence* NP19265.

39 Ajder and others, above n 29, at 2; Sensity, above n 29; and Tong, above n 5, at 3–4. See also Flynn and others, above n 4, at 1345.

40 Security Hero *2023 State of Deepfakes: Realities, Threats, and Impact* (2023). See also Clare McGlynn and Rüya Tuna Toparlak "The 'new voyeurism': criminalizing the creation of 'deepfake porn'" (2025) 52 *Journal of Law and Society* 1; and Ajder and others, above n 29, at 2.

41 Powell, Henry and Flynn, above n 24, at 309; Flynn and others, above n 4, at 1355; Rackley and others, above n 26, at 299. See also Anastasia Powell and Nicola Henry "Technology-Facilitated Sexual Violence Victimization: Results From an Online Survey of Australian Adults" (2019) 34 *J Interpers Violence* 3637 at 3637; and Rackley and others, above n 26, at 299.

42 Powell, Henry and Flynn, above n 24, at 310; Rackley and others, above n 26, at 299; and McGlynn and Rackley, above n 16, at 544.

Taken together, these findings establish that NCIDs are distinctly gendered and arguably threaten women's rights and freedoms even more than more traditional forms of IBSA.⁴³

2 *A continuum of sexual violence*

NCIDs disproportionately harm women. Accordingly, they are situated on a continuum of sexual violence that encompasses different but related forms of violence against women,⁴⁴ including physical, emotional, symbolic and structural violence.⁴⁵ The continuum's meaning is two-fold: first, acts of sexual violence are not discrete—they are interconnected and cannot be easily distinguished; and secondly, they share a common underlying character of men using abuse, intimidation, coercion, intrusion, threats and force to control women.⁴⁶

The continuum model was first developed by Liz Kelly in 1987,⁴⁷ but has since been expanded to incorporate technology-facilitated behaviours that coerce and control women, such as IBSA.⁴⁸ In developing their own continuum concept, Clare McGlynn and colleagues identified the "common character" underlying the various forms of IBSA as follows:

- (i) the sexual nature of the imagery;
- (ii) the gendered dynamics (with women overwhelmingly victims and men perpetrators);
- (iii) the harm caused by the violation of fundamental rights to sexual autonomy and dignity; and
- (iv) the minimisation of this type of abuse in public discourse and legal policy.⁴⁹

Viewing sexual violence as a continuum brings into focus acts that are commonly minimised, overlooked or normalised.⁵⁰ It recognises that NCID victims and other sexual violence victims experience similar harms and consequences, and that IBSA may be experienced alongside other forms of abuse, resulting in cumulative harm.⁵¹ This perspective demonstrates the interconnectedness of

43 Shona Moreau and Chloe Rourke "Fake porn causes real harm to women" (8 February 2024) Policy Options <<https://policyoptions.irpp.org>>.

44 Henry and others, above n 28, at 867.

45 Henry and Powell, above n 10, at 759.

46 Liz Kelly *Surviving Sexual Violence* (Polity Press, Cambridge, United Kingdom, 1988) at 76.

47 Liz Kelly "The Continuum of Sexual Violence" in Jalna Hanmer and Mary Maynard (eds) *Women, Violence and Social Control* (Palgrave Macmillan UK, London, 1987) 46.

48 McGlynn, Rackley and Houghton, above n 24.

49 At 29.

50 Harris, above n 9, at 318.

51 Powell and others, above n 17, at 26.

online and offline abuse,⁵² and highlights IBSA as a form of gendered abuse that reinforces and derives from gender inequality.⁵³

C Underlying Harms

The criminal law exists in part to protect certain interests valued by society.⁵⁴ Under international humanitarian and criminal law, rape is framed as an infringement of several fundamental interests: autonomy,⁵⁵ dignity,⁵⁶ privacy⁵⁷ and choice.⁵⁸ Researchers have identified that IBSA's underlying harms include violations of those same interests.⁵⁹ The breach of these interests is also inherent in the creation and distribution of NCIDs. To the extent that an individual's rights are restricted, this constitutes a distinct and serious harm in and of itself,⁶⁰ regardless of any further consequential harm.

1 Sexual autonomy

NCIDs infringe on sexual autonomy, which some scholars have argued is distinct from bodily autonomy.⁶¹ Bodily autonomy generally refers to each person's right to make decisions about their own body,⁶² while sexual autonomy is the freedom to determine one's own sexual expression and experiences, including with whom those are shared.⁶³ This distinction is more pronounced in the

52 Abigail C O'Hara and others "Crime script analysis for adult image-based sexual abuse: A study of crime intervention points for retribution-style offenders" (2020) 9 *Crime Science* 26 at 14; and Harris, above n 9, at 318.

53 Powell and others, above n 17, at 26.

54 Nicola Lacey "Unspeakable Subjects, Impossible Rights: Sexuality, Integrity and Criminal Law" (1998) 11 *Can J L and Jurisprudence* 47 at 52.

55 At 52.

56 Anna High "Sexual Dignity and Rape Law" (2021–2022) 33 *Yale JL & Feminism* 1 at 3.

57 Dubravka Šimonović *Rape as a grave, systematic and widespread human rights violation, a crime and a manifestation of gender-based violence against women and girls, and its prevention* (Human Rights Council, 47th Session, 3) at [20].

58 Lacey, above n 54, at 59.

59 McGlynn and Rackley, above n 16, at 546.

60 At 548.

61 Mary Childs "Sexual Autonomy and Law" (2001) 64 *Mod L Rev* 309 at 311.

62 Musa Springer "What Does 'Bodily Autonomy' Mean?" (12 January 2024) ARC Southeast <<https://arc-southeast.org>>.

63 Lacey, above n 54, at 52.

context of sexual violation: non-consensual sexual acts harm women's autonomy in ways physical assaults do not, because they violate not just the physical body but also the emotional mind.⁶⁴

Stephen Schulhofer argues sexual autonomy has two aspects: the right to choose to have sex and the right to refuse.⁶⁵ The right to refuse is eviscerated by NCIDs, which completely circumvent consent. Samantha Cole asserts that deepfakes were created to "own women's bodies".⁶⁶ The mere act of creating an intimate deepfake is an invasion of sexual autonomy because it represents a person as engaging in sexual encounters or intimate scenarios that are completely fabricated and beyond the victim's control.⁶⁷

NCIDs also affect victims' positive sexual expression. Cultural conditions of gender inequality fuel the social shaming and victim-blaming of women who "allow" the creation of sexual images of themselves,⁶⁸ particularly women who choose to exercise sexual agency in ways contrary to dominant sexual scripts.⁶⁹ Against this backdrop, women who might like to create or share consensual intimate images might choose not to for fear of unauthorised use or exploitation, curtailing their right to sexual expression.⁷⁰

2 *Dignity*

The concept of dignity recognises the intrinsic worth of all individuals and their right to be treated with respect.⁷¹ Kantian philosophy dictates that people deserve to be treated as an end rather than a means;⁷² other people are not mere instruments for satisfying one's own desires.⁷³ This maxim is clearly violated by NCIDs, which deliberately infringe on the victim's dignity through sexual objectification.⁷⁴ Intimate deepfakes reduce women to mere body parts; interchangeable and

64 Childs, above n 61, at 311.

65 Stephen J Schulhofer "Taking Sexual Autonomy Seriously: Rape Law and beyond" (1992) 11 Law & Phil 35 at 70–71.

66 Samantha Cole "Deepfakes Were Created As a Way to Own Women's Bodies—We Can't Forget That" (18 June 2018) VICE <www.vice.com>.

67 Regina Rini and Leah Cohen "Deepfakes, Deep Harms" (2022) 22 J Ethics & Soc Phil 143 at 143.

68 McGlynn and Rackley, above n 16, at 548.

69 At 549.

70 At 549.

71 At 546.

72 High, above n 56, at 7.

73 Adrienne de Ruitter "The Distinct Wrong of Deepfakes" (2021) 34 Philos Technol 1311 at 1322.

74 McGlynn and Rackley, above n 16, at 546.

commodified sexual objects to be exploited and exposed.⁷⁵ Victim and activist Noelle Martin described her experience as "dehumanising, degrading, [and] violating to just see yourself being misrepresented and being misappropriated in that way".⁷⁶

The concept of dignity is expansive enough to encompass both individual and collective dimensions.⁷⁷ Rape is an affront to the dignity of all women because it establishes and entrenches the subordination of women to men as sexual objects.⁷⁸ Similarly, NCIDs communicate to all women that they are lesser, and should feel unsafe and insecure, especially online.⁷⁹

3 *Privacy*

Privacy generally refers to a person's right to control access to their personal data and preserve selected confidentiality.⁸⁰ Danielle Citron has advanced the concept of "sexual privacy": the social norms that govern access to, and information about, our intimate lives, including but not limited to communications about sex, sexuality and gender.⁸¹ However, privacy is not the same as privateness.⁸² While *privateness* stems from the information's secrecy or concealment, *privacy* is not lost even if information is shared with others.⁸³ Privacy is not protected by blanket suppression but by the ability to determine with whom information is shared and for what purpose. Consent to share or use information is contextual.⁸⁴ Information shared for one purpose (for example, an Instagram story describing your day) is not permission to share for other uses (creating a sexually explicit deepfake).

75 Danielle Citron "Sexual Privacy" (2019) 128 Yale Law Journal 1870 at 1925; and Rini and Cohen, above n 67, at 147.

76 Justin Sherman "'Completely horrifying, dehumanizing, degrading': One woman's fight against deepfake porn" (14 October 2021) CBS News <www.cbsnews.com>.

77 High, above n 56, at 9.

78 At 18.

79 McGlynn and Rackley, above n 16, at 546.

80 Tharun Anand Reddy Sure "Ethical Implications of Deepfake Technology in User Interfaces" (2023) 10 International Journal of Computer Techniques at tbl 1. Summary of core ethical principles relevant to deepfakes in UIs.

81 Danielle Keats Citron "The Roots of Sexual Privacy: Warren and Brandeis & the Privacy of Intimate Life Symposium: Owning Personality: The Expanding Right of Publicity: Session IV" (2018–2019) 42 Colum JL & Arts 383 at 385.

82 McGlynn and Rackley, above n 16, at 547.

83 Danielle Keats Citron and Mary Anne Franks "Criminalizing Revenge Porn" (2014) 49 Wake Forest Law Review 345 at 356.

84 At 355.

In the NCID context, the privacy violation stems from a breach of the victim's expectation that their personal data will not be used in unauthorised ways.⁸⁵ Citron expressly identifies NCIDs as a harm to sexual privacy, noting they are "an affront to the sense that people's intimate identities are their own to share or to keep to themselves".⁸⁶

4 *The centrality of consent*

Notably, the above underlying harms are all predicated on the absence of consent. Consent is "morally transformative", in that it transforms illegal into legal conduct.⁸⁷ From a harms-based perspective, NCIDs are dangerous not because of their content per se, but because of the severe infringements on the victim's sexual autonomy, dignity and privacy.⁸⁸ Whether creation and/or distribution of intimate deepfakes infringes on these interests, and thus causes harm, centres on whether or not the subject gave their consent.

D Individual Experiential Harms

While some political actors have asserted that digitally manipulated images cause lesser harms than "original" images,⁸⁹ technological advancements make it increasingly difficult to differentiate between them.⁹⁰ IBSA and NCIDs both inflict harm through non-consensual sexual objectification, affecting victims' health and wellbeing. Therefore, their consequential harms are comparable, with a few notable distinctions.⁹¹

NCIDs do not impact victims equally. In line with Kelly's continuum, which highlights the complex and non-linear relationship between abuse and harm, the form or context of abuse does not determine the nature or degree of the harms experienced by victims.⁹² Any consequential harms will vary by gender, ethnicity, sexuality, age, class and social positioning, indicating a range of supports and responses are required.⁹³

85 McGlynn and Rackley, above n 16, at 547.

86 Citron, above n 75, at 1921.

87 Vera Bergelson "Autonomy, Dignity, and Consent to Harm" (2007–2008) 60 Rutgers L Rev 723 at 723.

88 Amelia O'Halloran "The Technical, Legal, and Ethical Landscape of Deepfake Pornography" (BSc Thesis, Brown University, 2021) at 37.

89 Karen Bradley *Letter to Alistair Carmichael MP from Karen Bradley MP (then Minister for Preventing Abuse, Exploitation and Crime)* (2016).

90 Rackley and others, above n 26, at 308.

91 Flynn and others, above n 4, at 1343.

92 Clare McGlynn and others "It's Torture for the Soul!: The Harms of Image-Based Sexual Abuse" (2021) 30 Social & Legal Studies 541 at 549.

93 At 543.

1 *Psychological and physical harm*

Mental health consequences for IBSA victims are similar in nature to survivors of rape and sexual assault.⁹⁴ IBSA victims often experience shame, embarrassment, fear, humiliation and declining self-esteem.⁹⁵ Consequently, victims often engage in unhealthy coping mechanisms such as avoidance, denial and self-medication.⁹⁶ This may exacerbate their psychological harm, leading to more disruptive, severe and diagnosable mental health issues including depression, anxiety, panic attacks and PTSD.⁹⁷ Suicidal ideation and attempted suicide is relatively common, particularly for young victims.⁹⁸ The profound psychological trauma alone is a significant harm that warrants legal intervention.

For some, the psychological toll of victimisation can induce various stress-related conditions such as headaches, nausea, spasms and chronic pain.⁹⁹ Physical health impacts may be pronounced for victims with pre-existing disabilities and conditions, as victimisation exacerbates their symptoms and limits their ability to go about their daily lives.¹⁰⁰

NCID publication can also put victims at physical risk. Personal details and unsubstantiated allegations about the victim are sometimes published alongside their images, a practice known as "doxxing".¹⁰¹ Sometimes this is done intentionally to incite others to identify, harass and make sexual demands of the victim in person,¹⁰² putting them at increased risk of revictimisation, as well as offline

94 Samantha Bates "Revenge Porn and Mental Health: A Qualitative Analysis of the Mental Health Effects of Revenge Porn on Female Survivors" (2017) 12 *Feminist Criminology* 22 at 22, 33 and 39.

95 At 26, 33 and 38; and Robert Chesney and Danielle Keats Citron "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security" [2019] SSRN Journal at 1773.

96 Bates, above n 94, at 39.

97 At 27, 31 and 39; and McGlynn and Rackley, above n 16, at 545.

98 McGlynn and Rackley, above n 16, at 545; and Bates, above n 94, at 22. See for example Sam McPhee "Matilda Tilly Rosewarne: Teenage Suicide Victim's Torment Revealed by Family in Heartwrenching Essay" *Daily Mail UK* (24 March 2022) <www.dailymail.co.uk>; and Julie Cross "Matilda 'Tilly' Rosewarne's parents call for bullying consultants and principals to have power to punish" *The Daily Telegraph* (22 October 2022) <www.dailytelegraph.com.au>.

99 Rackley and others, above n 26, at 301.

100 At 301.

101 Powell and Henry, above n 37, at 138; and McGlynn and Rackley, above n 16, at 545. See also Nicola Henry and Asher Flynn "Image-Based Sexual Abuse: Online Distribution Channels and Illicit Communities of Support" (2019) 25 *Violence Against Women* 1932 at 1933.

102 Powell and Henry, above n 37, at 138.

harassment, stalking and assault.¹⁰³ Further, when IBSA occurs in a domestic relationship context, images may be used to coercively control victims, meaning they could be forced to engage in additional non-consensual acts, stay in abusive relationships or be prevented from disclosing their victimisation.¹⁰⁴

2 *Social rupture*

Clare McGlynn and colleagues developed the term "social rupture" to better encapsulate the totality of the diverse, subtle and pervasive harms experienced by IBSA victims.¹⁰⁵ Social rupture describes the complete devastation or disruption to the everyday lives of victims, and the way victimisation can radically alter victims' understandings of themselves and the world around them.¹⁰⁶ For some, the impacts of their experiences are all-encompassing and pervasive, causing harms which permeate their personal, professional and digital social worlds.¹⁰⁷

NCIDs can severely damage a victim's reputation due to false implications of sexual impropriety.¹⁰⁸ A single image or video can quickly dominate search engine results and be disseminated widely, reaching loved ones, potential employers and peers.¹⁰⁹ The personal and professional costs are potentially severe. Victims may suffer job loss, difficulty finding or sustaining employment and education, and challenges forming or maintaining relationships.¹¹⁰

IBSA victims often experience profound isolation, stemming from the breach of trust caused by the perpetrator, which can seep into their broader social circles, including family, friends, and peer

103 Citron and Franks, above n 83, at 350; and Adam Dodge and Erica Johnstone *Using Fake Video Technology to Perpetrate Intimate Partner Abuse* (2018) at 4–5. See also Flynn and others, above n 4, at 1343; Bates, above n 94, at 23.

104 Powell and Henry, above n 37, at 138.

105 Clare McGlynn and others *Shattering Lives and Myths: A Report on Image-Based Sexual Abuse* (2019) at 6.

106 McGlynn and others, above n 92, at 551. See also Flynn and others, above n 4, at 1351; and Henry and others, above n 28, at 867.

107 McGlynn and others, above n 105, at 6.

108 Joanne Worsley and Grace Carter "The Impact of Technology-Facilitated Sexual Violence: A Literature Review of Qualitative Research" in Anastasia Powell, Asher Flynn and Lisa Sugiura (eds) *The Palgrave Handbook of Gendered Violence and Technology* (Springer International Publishing, Cham, 2021) 261 at 269; and Kathie Reveley "'It's Been Devastating': An Interpretative Phenomenological Analysis of the Experience of being Cyberstalked" (University of East London, 2017) at 85.

109 Dodge and Johnstone, above n 103, at 4.

110 Worsley and Carter, above n 108, at 269; Citron and Franks, above n 83, at 352; Dodge and Johnstone, above n 103, at 4; McGlynn and Rackley, above n 16, at 545; and Powell and Henry, above n 37, at 137.

networks.¹¹¹ The anticipated or actual negative reactions from others—feeling judged, alone and unsupported—can further compound this isolation.¹¹² Their confidence in others is eroded, leading many to withdraw from social interactions, both online and offline.¹¹³ Some delete their social media accounts.¹¹⁴ Ortega recently revealed she left Twitter (now X) after seeing NCIDs of herself.¹¹⁵ These reactions, while aimed at reducing revictimisation and regaining control, can deepen their social isolation by cutting off positive social interactions.

3 *Constancy*

Participants in several IBSA studies have described the "constancy" of their abuse: the harms are "permanent", "ongoing" and "endless".¹¹⁶ Central to this issue is the concept of the "enduring image".¹¹⁷ Digital recordings or manipulations capture or create a particular event and extend it over time, allowing it to be relived in perpetuity.¹¹⁸ Once uploaded, images are nearly impossible to remove from the Internet, where they can be rapidly uploaded, replicated, downloaded and disseminated. Fresh harm is produced with each additional viewing or distribution.¹¹⁹ Even if images are removed, victims experience a "visceral fear" linked to uncertainty over who has seen or will see the images, and the ever-present possibility their images will be re-uploaded.¹²⁰ The constant fear of revictimisation complicates recovery and healing for victims, highlighting the need for legal remedies that recognise the lasting impacts on victims, and empowers them to move forward.

111 McGlynn and others, above n 92, at 554. See also Worsley and Carter, above n 108, at 268; Bates, above n 94, at 30; and Antoinette Raffaella Huber "Women, Image Based Sexual Abuse and the Pursuit of Justice" (PhD, Liverpool John Moores University (United Kingdom), 2020) at 126–127.

112 McGlynn and others, above n 92, at 554; and Henry and others, above n 28, at 868.

113 Worsley and Carter, above n 108, at 270; Bates, above n 94, at 23; and McGlynn and others, above n 92, at 554.

114 Flynn and others, above n 4, at 1352. See also Citron and Franks, above n 83, at 352; Bates, above n 94, at 23; and Nicola Henry, Asher Flynn and Anastasia Powell "Technology-Facilitated Domestic and Sexual Violence: A Review" (2020) 26 *Violence Against Women* 1828 at 1840.

115 Lana Andelane "Jenna Ortega quit Twitter after seeing explicit AI content of herself" *NZ Herald* (26 August 2024) <www.nzherald.co.nz>.

116 Rackley and others, above n 26, at 298; and McGlynn and others, above n 92, at 552.

117 Powell and Henry, above n 37, at 135.

118 Alexandra S Marcotte and Jessica J Hille "Sexual Violence and Consent in the Digital Age" in Anastasia Powell, Asher Flynn and Lisa Sugiura (eds) *The Palgrave Handbook of Gendered Violence and Technology* (Springer International Publishing, Cham, 2021) 319 at 320.

119 Rackley and others, above n 26, at 298; and Powell and Henry, above n 37, at 135.

120 Citron, above n 75, at 1925; and McGlynn and others, above n 92, at 553. See also Huber, above n 111, at 116.

4 *Harms unique to deepfakes*

Research demonstrates that IBSA is a potentially devastating form of abuse that causes significant, multi-faceted and enduring harm.¹²¹ While IBSA and NCID harms are broadly comparable, deepfakes also generate unique harms. First, NCIDs typically involve two victims: the person whose face is superimposed onto the image, and the performer, whose intimate images are used without their consent.¹²² Although the harm to performers is often overlooked, it can be significant:¹²³ the circumvention of their consent may leave them feeling violated, and involuntarily implicates them in the abuse of others.¹²⁴

Secondly, while the opportunity to create and distribute "original" images may depend on situational factors—for example, the existence of a relationship of trust in which images may be taken or exchanged—AI allows images to be produced on a whim, with access to an unlimited victim pool and endless publicly available information.¹²⁵ Deepfake technology obviates the need for victim-offender interaction. Additionally, while "original" images can be protected from malicious actors (albeit imperfectly), there is little protection against deepfake creation.¹²⁶ Very little data is required to create a deepfake and, in this increasingly digitalised world, it is nearly impossible to fully participate in society while also remaining assured of data privacy.¹²⁷

Finally, deepfakes have the potential to destabilise or even overwrite a victim's autobiographical memories.¹²⁸ In this way, deepfakes may be used to gaslight victims—that is, to emotionally manipulate them into believing that their perception of reality is mistaken.¹²⁹ The more regularly and persistently victims are exposed to deepfakes of themselves, the more likely their resistance will be worn down, until their memories are altered and their self-image is distorted.¹³⁰

121 Rackley and others, above n 26, at 298.

122 Tong, above n 5, at 7.

123 At 7–8.

124 Lux Alptraum "Deepfake Porn Harms Adult Performers, Too" (17 July 2024) WIRED <www.wired.com>.

125 Tong, above n 5, at 8.

126 Moreau and Rourke, above n 43.

127 Moreau and Rourke, above n 43.

128 Rini and Cohen, above n 67, at 153 and 154.

129 Kate Abramson "Turning up the Lights on Gaslighting" (2014) 28 *Philosophical Perspectives* 1 at 2.

130 Rini and Cohen, above n 67, at 155.

E Social and Cultural Harms

NCID impacts extend beyond individual harm—they have broader societal and cultural implications. Like physical forms of sexual violence, online perpetration stems from socially constructed beliefs and attitudes about gender and sexuality,¹³¹ which hugely impact how certain victim classes experience NCID abuse. A reductive focus on individual harm decontextualises and depoliticises sexual violence, disconnecting it from the patriarchy, misogyny and other forms of oppression that facilitate sexual violence.¹³² Persistent double standards for men and women about "appropriate" sexual behaviour are reflected both in the way that NCID abuse is perpetrated and responded to.

Women suffer more pervasive harms than men, partly due to an increased likelihood of experiencing online sexual harassment consisting of rape threats and gender-based hate speech, typically perpetrated by men.¹³³ Although popular discourse frames online harassment as acts by individual perpetrators, men often network harassment collectively.¹³⁴ Male-dominated online networks, such as the "manosphere" and similar subcultural platforms like Reddit and 4chan, unite men around anti-feminist ideologies.¹³⁵ These networks cultivate a culture that not only tolerates but *encourages* the sexual harassment of women.¹³⁶ Although men may be victimised, they are far less likely than women to experience coordinated campaigns of continuous intimidation and humiliation, facilitated by digital platforms.¹³⁷

By targeting feminist public figures, networked misogyny aims to silence women by forcing them out of online spaces.¹³⁸ For example, Indian journalist Rana Ayyub became the target of a NCID scheme after her attempts to expose government corruption.¹³⁹ Despite the images' artificiality, Ayyub faced widespread backlash which compelled her to move and temporarily stop working.¹⁴⁰ Activist Noelle Martin also experienced an onslaught of NCID abuse. After speaking out about her

131 Henry and Powell, above n 9, at 204.

132 McGlynn and others, above n 92, at 545.

133 Henry and Powell, above n 9, at 199; and Powell and Henry, above n 41, at 3655 and 3659.

134 Alice E Marwick and Robyn Caplan "Drinking male tears: Language, the manosphere, and networked harassment" (2018) 18 *Fem Media Stud* 543 at 543.

135 At 545 and 553.

136 Emma A Jane "Online misogyny and feminist digilantism" (2016) 30 *Continuum* 284 at 286.

137 Michael Salter and Thomas Crofts "Responding to revenge porn: Challenges to online legal impunity" in *New views on pornography: Sexuality, politics and the law* (Praeger Publisher, 2015) at 4.

138 At 4.

139 Moreau and Rourke, above n 43.

140 Moreau and Rourke, above n 43.

initial victimisation and successfully petitioning for law reform in Western Australia, Martin was targeted with another wave of sexually explicit deepfakes.¹⁴¹ She believes the second wave of videos were "weaponised to taunt, intimidate, and silence" her.¹⁴² These examples demonstrate the potentially significant harm to women's fundamental rights and freedoms by NCIDs, by preventing full civic participation.¹⁴³

Sexual double standards are also evident in responses to IBSA. Victim-blaming attitudes are commonly held by members of the public, including law enforcement, family and friends of the victim.¹⁴⁴ Victim-blaming attitudes persist even in the context of deepfakes where victims play no appreciable role in creating or disseminating their images.¹⁴⁵ Martin faced a barrage of victim-blaming rhetoric. Commentators said she was "asking for it", she was an "attention seeking whore" and "she should be flattered people would do that to her".¹⁴⁶ They warned her that she should be careful what she posted online, or get off social media completely.¹⁴⁷ The media frequently advocate risk management strategies, effectively holding women at least partially responsible for their victimisation.¹⁴⁸ These attitudes may be internalised by victims, thereby impeding their ability to seek support and justice.¹⁴⁹

Networked misogyny and victim-blaming narratives signify a broader rape culture, in which deeply embedded beliefs about gender and sexuality either expressly or impliedly condone men's violence against women.¹⁵⁰ Such a culture is harmful to all women. NCIDs are thus a self-perpetuating cycle of abuse, sustained by and conducive to culture that minimises and normalises non-consensual behaviour. Recognising the gendered dimensions of NCIDs emphasises social inequality as a source of harm and highlights the potential for prevention and mitigation through social change,¹⁵¹ which can be prompted and supported by changes to the law.

141 Martin, above n 6, at 67–68.

142 At 68.

143 Powell and Henry, above n 37, at 139.

144 Asher Flynn and others "Victim-blaming and image-based sexual abuse" (2023) 56 J Criminol 7 at 9.

145 Tong, above n 5, at 17.

146 Martin, above n 6, at 63.

147 At 63.

148 Bates, above n 94, at 25.

149 Flynn and others, above n 144, at 8.

150 Powell and Henry, above n 37, at 136.

151 McGlynn and others, above n 92, at 546.

IV ANALYSIS OF REGULATORY OPTIONS

The analysis above demonstrates that NCIDs cause significant individual, social, and cultural harms, providing a strong basis for criminalisation. However, it does not necessarily follow that criminalisation is the best response. Different modes of regulation serve different purposes, possess different benefits and limitations, and have differing impacts on those who are subject to them or seek justice through them. Any argument for legislative reform must identify the most appropriate mechanism to achieve its desired regulatory aims.¹⁵²

A Online Platform Regulation

Since NCIDs are primarily disseminated online, online platform hosts can adopt self-regulatory measures to prevent abuse occurring on their websites. Many mainstream platforms—including Google,¹⁵³ Reddit,¹⁵⁴ Facebook,¹⁵⁵ Instagram,¹⁵⁶ X,¹⁵⁷ Discord¹⁵⁸ and Pornhub¹⁵⁹—have already prohibited NCIDs.

Online platforms are also subject to legal regulation. Under New Zealand's Harmful Digital Communications Act 2015 (HDCA), the District Court can make orders against online content hosts to take down or disable access to material that has been posted or sent.¹⁶⁰ Non-compliance is a criminal offence.¹⁶¹ These mechanisms are important for facilitating content removal—a primary concern for victims.¹⁶² However, online platform regulation has several inadequacies.

152 Simester and Brookbanks, above n 30, at 1004.

153 Google Search Help "Remove explicit or intimate personal images from Google" <<https://support.google.com/>>; and Emily Price "Updated Google Ad Policy Bans Promotion of Deepfake Porn Apps" *PCMag Australia* (4 May 2024) <<https://au.pcmag.com/>>.

154 Reddit "Content Policy" <www.redditinc.com> at r 3; and Alex Hern "Reddit bans 'deepfakes' face-swap porn community" *The Guardian* (8 February 2018) <www.theguardian.com>.

155 Transparency Centre "Adult nudity and sexual activity" Meta <<https://transparency.meta.com/>>; and Transparency Centre "Adult sexual exploitation" Meta <<https://transparency.meta.com/>>.

156 Instagram Help Center "Terms of Use" <<https://help.instagram.com/>>; and Instagram Help Center "Reducing the Spread of False Information on Instagram" <<https://help.instagram.com/>>.

157 X Help "Our synthetic and manipulated media policy" <<https://help.x.com/>>.

158 Discord "Non-Consensual Adult Intimate Media Policy Explainer" <<https://discord.com/>>.

159 Pornhub Help "What sort of content is not allowed on the site?" Pornhub Help <<https://help.pornhub.com/>>; and Pornhub "Anonymous content removal" <www.pornhub.com>.

160 Harmful Digital Communications Act 2015, s 19(2)(a).

161 Section 21.

162 Henry and others, above n 28, at 873.

In all cases, the burden is on the victim to detect and report their images before they are taken down. These regulatory mechanisms are reactive: even if removal is successful, the harm caused by its posting has already occurred. Further, although many mainstream platforms have publicly committed to tackling NCIDs,¹⁶³ approximately 94 per cent of online deepfakes are hosted on underground websites dedicated to "deepfake pornography".¹⁶⁴ Since NCIDs are the core feature of their business, these websites lack any incentive to identify and remove non-consensual images which, due to the website's restricted nature, are unlikely to be detected by victims or law enforcement.

Website regulation is also ineffective when it comes to promoting perpetrator responsibility or deterring users from engaging in harmful conduct. While breaches can lead to content removal and temporary or permanent account suspension,¹⁶⁵ users who violate community guidelines can simply create new accounts with different details or move on to other platforms with less stringent rules. Finally, online platform regulation only addresses situations in which content is shared online; creation, threatened distribution and private sharing between peers cannot be captured by these regulatory mechanisms. Therefore, solely relying on online platform regulation would be an incomplete solution to the NCID problem.

B Civil Liability

1 Civil law advantages

In some circumstances, the extension of a common law tort or the creation of a statutory tort to capture a certain kind of harmful conduct may be a suitable alternative to criminal sanctions.¹⁶⁶ Enabling victims to harness the civil law has several advantages. Civil claims can help put victims back in control, which may be empowering in the NCID context.¹⁶⁷ Civil claimants need not rely on persuading others (such as the police) of the strength or seriousness of their case to bring a claim.¹⁶⁸ Moreover, the standard of proof, "on the balance of probabilities", is lower than for a criminal case, meaning that the prospects of success are higher.

163 See for example "The Code: Aotearoa New Zealand Code of Practice for Online Safety and Harms" (July 2022) The Code <<https://thecode.org.nz>> at app 1.

164 Ajder and others, above n 29, at 6.

165 Jesús C Aguerri, Fernando Miró-Llinares and Ana B Gómez-Bellví "Consensus on community guidelines: An experimental study on the legitimacy of content removal in social media" (2023) 10 *Humanit Soc Sci Commun* 1 at 2.

166 Simester and Brookbanks, above n 30, at 1024.

167 McGlynn and Rackley, above n 16, at 558.

168 At 558.

Civil claims can widen the net of potential defendants to include not only the primary offender but also hosts and successive distributors.¹⁶⁹ If a claimant is successful, civil remedies—such as damages and/or injunctive relief—may be more beneficial to someone trying to rebuild their life than a perpetrator's criminal sentence.¹⁷⁰ Arguably, civil remedies are more victim-centred; they aim to restore the victim to the position they would have been in had the wrong not occurred,¹⁷¹ and compensate the victim for the harms suffered.¹⁷² This contrasts with criminal sentences, which have several competing purposes, most of which are offender- and community-focused.¹⁷³

2 *New Zealand civil causes of action*

Before seeking to enact new legislation, it is important that all relevant existing laws are identified and considered.¹⁷⁴ There is a range of legal regimes touching upon the potential harms caused by NCIDs.¹⁷⁵ Therefore, it must be determined whether any existing civil causes of action satisfactorily deal with their creation, distribution and threatened distribution. It is not necessary to show that civil law could *never* provide redress to justify NCID criminalisation; there are many reasons a victim might opt to pursue a civil claim, outlined above. The following discussion focuses on problematic elements present across several causes of action, rather than separately assessing each potential cause of action, to demonstrate their limitations.

(a) Publication

Publication is a requirement for several statutory and common law torts. For example, in a defamation claim, the plaintiff must prove that the defendant *published* a defamatory statement to someone other than the plaintiff.¹⁷⁶ For breach of privacy, there must be *public disclosure* of private facts.¹⁷⁷ The publication requirement defeats any NCID claim not involving dissemination. Therefore,

169 At 560.

170 At 558.

171 *Livingstone v Rawyards Coal Company* (1880) 5 App Cas 25, at 39.

172 Stephen Todd *The Laws of New Zealand Nature and Definition of Tort: General* (online ed, LexisNexis NZ Limited) at [3].

173 Sentencing Act 2002, s 7.

174 Legislation Design and Advisory Committee "Legislation Guidelines" (2021) [ldac.org.nz](http://www.ldac.org.nz) <www.ldac.org.nz> at [3.1 and 3.4]. See also Tom Barraclough and Curtis Barnes *Perception inception: Preparing for deepfakes and the synthetic media of tomorrow* (2019) New Zealand at [17].

175 Barraclough and Barnes, above n 174, at [15].

176 Ursula Cheer *The Laws of New Zealand Regulation of Publication: Defamation* (online ed, LexisNexis NZ Limited) at [171].

177 *P v D* [2000] 2 NZLR 591 at [34].

instances where the defendant merely created an NCID, threatened to share it, or shared it only with the plaintiff, will not be covered.

(b) Intention to cause harm

To satisfy a claim for intentional infliction of emotional harm, the defendant must have *intended to cause harm* to the plaintiff.¹⁷⁸ However, NCID perpetrators may be motivated by a number of factors, including sexual gratification, extortion, control, proof of masculinity, consent circumvention, monetary gain or social notoriety.¹⁷⁹ Dependence on the perpetrator's primary intention will lead to inconsistent outcomes for victims and inappropriately detract from the harm caused to the victim.

(c) Serious or severe harm

A final problematic element is the "serious harm" threshold. In New Zealand, physical or bodily harm that results from severe emotional distress has become a requirement for intentional infliction of emotional harm.¹⁸⁰ Likewise, under the HDCA's civil regime, breaches of communication principles must cause or be likely to cause harm to an individual.¹⁸¹ Harm is defined as "serious emotional distress".¹⁸² The "serious harm" threshold invites evaluations about the level of harm a victim *actually* feels or *ought to* feel, rather than recognising the harm inherent in the wrongful act.

This point is illustrated by a New Zealand case. In *Police v B*, the victim's estranged husband threatened to post photographs online unless she stayed away from other men and cancelled the existing protection order.¹⁸³ Subsequently, he created a fake Facebook account posing as the victim, and posted semi-nude images of her that were obtained without her knowledge or consent.¹⁸⁴ However, charges under the HDCA were dismissed in the first instance because the judge was not satisfied that the victim's suffering reached the threshold of "serious emotional distress",¹⁸⁵ despite

178 *Bradley v Wingnut Films Ltd* [1993] 1 NZLR 415, at 422.

179 Nicola Henry, Asher Flynn and Anastasia Powell "Image-Based Sexual Abuse: Victims and Perpetrators" (2019) 572 *Trends & issues in crime and criminal justice* at 2; Okolie, above n 4, at 3–4; and Henry, Flynn and Powell, above n 17, at 58.

180 *Tucker v News Media Ownership Ltd* [1986] 2 NZLR 716, at 732; and *Bradley v Wingnut Films Ltd* at 421. See also Geoffrey Palmer "Freedom and dignity" [2024] NZLJ 212 at 215.

181 Harmful Digital Communications Act, s 12(2).

182 Section 4 definition of "harm".

183 *New Zealand Police v B* [2017] 3 NZLR 203 (HC) at [39].

184 At [40].

185 *R v Iyer* [2017] DCR 82 at [73].

evidence that the victim had become depressed, anxious and reliant on peer support.¹⁸⁶ The case was successfully appealed but was then retried by the same judge, who again dismissed the case.¹⁸⁷

Few cases fit neatly within the confines of a particular cause of action.¹⁸⁸ Consequently, one case may satisfy all elements while another similar case will not, notwithstanding the same harmful conduct involved. While some victims may be successful, the civil law in its current manifestation cannot comprehensively address all NCID experiences, leaving the law in an undesirably ambiguous state.

3 *Inherent limitations*

Beyond the inadequacies of existing civil laws identified above, there are inherent limitations that would prevent developments in the civil arena from effectively addressing NCIDs. First, the financial cost of civil litigation can be prohibitively expensive for many victims, even for plaintiffs with a high chance of success.¹⁸⁹ Secondly, success in court does not necessarily resolve the issue: an award of compensation is meaningless against a defendant who is unable to pay,¹⁹⁰ and an injunction against a defendant to remove images or stop further dissemination is not effective against any others who may possess and share the images.¹⁹¹ In fact, civil litigation may exacerbate the victim's harm because it can bring more attention to their images.¹⁹² This signals a third limitation: civil law plaintiffs do not have the benefit of automatic name suppression afforded to victims of criminal sexual offences,¹⁹³ which may deter them from bringing a claim.¹⁹⁴

Fourthly, efforts to adapt and extend various torts to provide the claimant with a remedy risk unacceptably distorting the harms of NCIDs;¹⁹⁵ for example, by re-characterising the harm as a breach of property rights or limiting the harm to just one aspect. Fifthly, while an award of damages may compensate the victim for the harm they have suffered, punitive or exemplary damages, which aim to

186 At [13]–[14].

187 Nicola Gavey and Jade Farley "Reframing sexual violence as 'sexual harm' in New Zealand policy: A critique" in M Gabriela Torres and Kersti Yllö (eds) *Sexual Violence in Intimacy: Implications for Research and Policy in Global Health* (Taylor & Francis Group, Oxford, United Kingdom, 2020) at 240.

188 McGlynn and Rackley, above n 16, at 559.

189 At 559; Rackley and others, above n 26, at 314–315; and Henry, Flynn and Powell, above n 114, at 1843.

190 Tong, above n 5, at 12; and McGlynn and Rackley, above n 16, at 21.

191 McGlynn and Rackley, above n 16, at 559.

192 Tong, above n 5, at 12; Rackley and others, above n 26, at 303–304.

193 Criminal Procedure Act 2011, s 203; and Rackley and others, above n 26, at 303–304.

194 See Martin, above n 6, at 68.

195 McGlynn and Rackley, above n 16, at 559.

punish the defendant, are rarely awarded.¹⁹⁶ Civil remedies effectively price, rather than prohibit, the wrongful conduct.¹⁹⁷ Finally, the civil law lacks the symbolic, condemnatory function of the criminal law,¹⁹⁸ significantly restricting the civil law's broader deterrent effect. After all, as Danielle Citron and Mary Anne Franks observe, "would-be perpetrators are unlikely to fear a course of action that is unlikely to materialise".¹⁹⁹

C Criminal Liability

1 Criminal law advantages

The criminal law has unique advantages over other regulatory methods. From a financial standpoint, the criminal law is a more viable option for many victims, because the state brings the case at no financial expense to them.²⁰⁰ It is appropriate for the state to assume the responsibility of regulating behaviour affecting the community collectively, rather than one-off disputes affecting a limited number of individuals.²⁰¹ NCIDs are harmful to society as a whole and are thus rationally positioned within the criminal law realm. A criminal provision would also correctly situate NCIDs alongside other similarly harmful sexual offences, accurately characterising their harms and endorsing comparable punishment.

The civil law also lacks the criminal law's preventative function. While criminal sanctions primarily seek to punish the offender for crimes already committed, they also aim to promote a sense of responsibility and acknowledgement in offenders for the harm they have caused,²⁰² assist in their rehabilitation and reintegration²⁰³ and deter both the individual offender and society generally from repeat offending.²⁰⁴ Accordingly, the criminal law has a distinct moral voice; it both reflects and influences social values through official condemnation of certain conduct.²⁰⁵

Martin discusses how, at the time of her abuse, she faced substantial public misapprehension and misunderstanding because NCIDs were not criminalised and community attitudes towards NCIDs

196 Salter and Crofts, above n 137, at 8–9.

197 Simester and Brookbanks, above n 30, at 1024.

198 Salter and Crofts, above n 137, at 8.

199 Citron and Franks, above n 83, at 358.

200 McGlynn and Rackley, above n 16, at 557.

201 Andrew Cornford "The Aims and Functions of Criminal Law" (2024) 87 *The Modern Law Review* 398 at 412.

202 Sentencing Act, s 7(1)(b).

203 Section 7(1)(h).

204 Section 7(1)(f).

205 Simester and von Hirsch, above n 35, at 19.

were not established.²⁰⁶ The criminal law's expressive and preventative functions are thus important tools for influencing social attitudes. Cultural practices are rapidly evolving in the technological space. The time is ripe for legal action to guide acceptable standards of behaviour in the digital realm.²⁰⁷

2 *Criminal law limitations*

Despite its many advantages, the criminal law has several limitations, particularly in the NCID context. First, victim-blaming and harm minimisation attitudes discourage victims from reporting victimisation to the police.²⁰⁸ Even when reported, law enforcement may experience difficulty attributing the offending to a specific perpetrator, particularly in cases where the perpetrator has no "real world" link to the victim.²⁰⁹ Perpetrators secure their anonymity by using virtual private networks (VPNs) that prevent IP addresses from being traced.²¹⁰ There are also significant cross-jurisdictional challenges to prosecuting NCIDs. Countries rarely mobilise in relation to technology-facilitated abuse against adults,²¹¹ and strategies for combatting NCIDs, including legislation (or lack thereof), is inconsistent across jurisdictions.²¹²

The law enforcement difficulties that technology presents justify considering additional interventions beyond legislation, particularly those focused on prevention.²¹³ However, the fact that enforcement might be difficult in some cases does not outweigh the need for criminal law reform.

3 *Criminal offences in New Zealand*

While NCIDs are not specifically criminalised under New Zealand law, several existing laws may at least partially cover their dissemination.²¹⁴ The Harmful Digital Communications Bill was first

206 Martin, above n 6, at 63–64.

207 See Danielle Keats Citron "Law's Expressive Value in Combating Cyber Gender Harassment" (2009) 108 Mich L Rev 373 at 409. See also McGlynn and Rackley, above n 16, at 552.

208 Tong, above n 5, at 17–18; and Powell and Henry, above n 37, at 135–136.

209 Tong, above n 5, at 15.

210 At 15.

211 Henry, Flynn and Powell, above n 114, at 1844.

212 See Tong, above n 5, at 16.

213 Henry and Powell, above n 10, at 773.

214 Simpson Grierson "Pictures don't lie, or do they?" (20 May 2022) <www.simpsongrierson.com>.

introduced in 2013 to tackle cyberbullying,²¹⁵ but since its enactment in 2015, it has been extended and amended to cover other harmful online behaviours such as IBSA.²¹⁶

Section 22 criminalises causing harm by posting digital communications. Three elements must be satisfied: (1) the perpetrator *intended* to cause harm to the victim; (2) the post would cause harm to an ordinary reasonable person in the victim's position; and (3) the post did in fact *cause* harm to the victim.²¹⁷ The broad definition of "digital communication" under the Act plausibly covers NCIDs.²¹⁸ However, the "motivation threshold" is unlikely to be proven except in paradigmatic "revenge porn" cases, and is particularly unsuitable where the perpetrator and victim are unknown to each other.²¹⁹ Even if intention is established, the victim must have experienced serious emotional distress,²²⁰ and that experience must have been objectively reasonable, which encourages scrutiny of the victim in a context in which victim-blaming is already prevalent. Moreover, because the original HDCA provisions were enacted with different behaviours in mind, their existence and applicability in this context are relatively unknown and underutilised by police and prosecutors.²²¹

In recognition of the high threshold that s 22 creates, the HDCA was amended in 2022 to improve access to justice, particularly for IBSA victims.²²² It is now an offence to post intimate visual recordings without consent, and no proof of intention or actual harm is required.²²³ Removing these requirements has recognised the variety of perpetrator motivations and inherent harms.²²⁴ However, the new provision still has several problems.

Section 22A only applies to "intimate visual recordings". The HDCA mirrors the Crimes Act 1961 definition: an intimate visual recording is widely defined as "made in any medium using any

215 See New Zealand Parliament "Harmful Digital Communications Bill—First Reading" (3 December 2013) <www.parliament.nz>.

216 New Zealand Ministry of Justice "Changes to New Zealand's Harmful Digital Communications Law" <www.justice.govt.nz>.

217 Harmful Digital Communications Act, s 22(1).

218 Section 4 definition of "digital communication".

219 McGlynn and Rackley, above n 16, at 555.

220 Harmful Digital Communications Act, s 4 definition of "harm".

221 McGlynn and Rackley, above n 16, at 17.

222 Harmful Digital Communications (Unauthorised Posting of Intimate Visual Recording) Amendment Act 2022; and New Zealand Ministry of Justice, above n 216.

223 Harmful Digital Communications Act, s 22A.

224 New Zealand Ministry of Justice, above n 216. See also "Harmful Digital Communications Act: Recent cases, changes and calls for review" (9 May 2022) New Zealand Family Violence Clearinghouse <<https://nzfvc.org.nz>>.

device".²²⁵ However, the recording must be of an individual "in a place which ... would reasonably be expected to provide privacy", and they must be naked, engaged in sexual activity, or "engaged in showering, toileting, or [an]other personal bodily activity that involves dressing or undressing".²²⁶ These requirements ostensibly focus on the real location and body of the subject.²²⁷ Though it is open to the courts to include manipulated imagery,²²⁸ this might conflict with statutory interpretation rules that mandate an enactment's meaning must be ascertained in light of its purpose.²²⁹ Parliament specifically considered and rejected suggestions to expressly incorporate deepfakes into the definition prior to the 2022 amendments.²³⁰ Although NCIDs were prevalent then, AI technologies have become more accessible and sophisticated since, necessitating reconsideration by Parliament.²³¹

Even if the definition of "intimate visual recording" extends to deepfakes, both ss 22 and 22A only apply to posted content. Under the HCDA, "post" means "to transfer, send, publish, disseminate, or otherwise communicate" the relevant content.²³² There is no other offence in the HDCA that applies to the creation or threatened dissemination of digital content. Therefore, as it currently stands, the HDCA cannot conclusively or comprehensively address NCIDs under one cohesive criminal regime.

4 *Other jurisdictions*

In response to growing concerns about the proliferation of technology-facilitated abuse, several international jurisdictions have enacted, or plan to enact, specific legislation to criminalise NCIDs.²³³

225 Compare Harmful Digital Communications Act, s 4 definition of "intimate visual recording", para (a); and Crimes Act 1961, s 216G(1).

226 Harmful Digital Communications Act, s 4 definition of "intimate visual recording", para (a)(i) and (ii).

227 Diana Clement "Legislative gap leaves New Zealand exposed to deepfakes" (18 April 2024) McVeagh Flemming Lawyers <www.mcveaghflemming.co.nz>.

228 See Barraclough and Barnes, above n 174, at [505]–[510].

229 Legislation Act 2019, s 10(1).

230 (15 February 2022) 757 NZPD" <www.parliament.nz> at 7485; and Supplementary Order Paper 2021 (83) Harmful Digital Communication (Unauthorised Posting of Intimate Visual Recording) Amendment Bill 2021, 305–2. See also Katie Harris "'It's urgent': Calls for protections for 'deepfake porn' victims" *NZ Herald* (7 December 2021) <www.nzherald.co.nz>.

231 Since the time of writing, a Member's Bill has been put before Parliament that would expand the definition of "intimate visual recording" in both the Crimes Act and the Harmful Digital Communications Act to expressly include visual recordings that are created, synthesised or altered to appear intimate, where this occurs without the knowledge or consent of the person depicted: see Deepfake Digital Harm and Exploitation Bill 2025 (213-1). See also Cassandra Mudgway "Deepfake Digital Harm and Exploitation Bill" [2026] NZLJ 6.

232 Harmful Digital Communications Act, s 4 definition of "post", para (a).

233 Flynn and others, above n 144, at 9.

Under the United Kingdom's Online Safety Act 2023, it is a criminal offence to share or threaten to share an intimate photograph or film of a person without their consent.²³⁴ Significantly, images "altered by computer graphics or in any other way" are covered.²³⁵ These offences were inserted into the Sexual Offences Act 2003,²³⁶ situating them alongside other sexual violence provisions. Motivations such as intention to cause distress or sexual gratification are not requirements, although if proven, can increase penalties.²³⁷ Only the absence of consent is required.²³⁸ Further, under the Criminal Justice Bill 2023, which was making its way through the House of Commons up until the 2024 UK election, it would have been an offence to *create* intimate images of an adult without their consent, regardless of whether the creator intended to distribute it.²³⁹ However, the proposed legislation did have limitations: the image must have been created with the intention to cause alarm, distress or humiliation, or for the purpose of sexual gratification. This would have created inconsistencies in the law.²⁴⁰

Australia's recently passed Criminal Code Amendment (Deepfake Sexual Material) Act 2024 creates a federal criminal offence for the non-consensual dissemination of sexual imagery, including deepfakes, carrying a 6-year penalty, or 7 years for offenders who created the images they disseminate.²⁴¹ However, this federal law does not establish creation as a standalone offence; to do

234 Online Safety Act 2023 (UK), s 188.

235 Sexual Offences Act 2003 (UK), ss 66A(5)(a) and 66D(4).

236 At ss 66A–66D.

237 Manasa Narayanan "The UK's Online Safety Act Is Not Enough To Address Non-consensual Deepfake Pornography" (13 March 2024) Tech Policy Press <<https://techpolicy.press>>; Haroon Siddique "Sharing deepfake intimate images to be criminalised in England and Wales" *The Guardian* (26 June 2023) <www.theguardian.com>; and Tamsin Nowell "New laws criminalise the sharing of intimate deepfakes without consent" (5 February 2024) Refuge <<https://refuge.org.uk>>.

238 Sexual Offences Act (UK), s 66B(1).

239 Ministry of Justice and Laura Farris "Government cracks down on 'deepfakes' creation" (16 April 2024) gov.uk <www.gov.uk>; and Christy Cooney "Creating sexually explicit deepfakes to become a criminal offence" (16 April 2024) <www.bbc.com>.

240 Since the time of writing, the UK Parliament has opted to change its approach and move to a consent-based model, no longer requiring proof of the perpetrator's intent. See Lucy Morgan "Government U-turns on problematic deepfake amendment, following pressure from GLAMOUR, survivors and Baroness Charlotte Owen" *GLAMOUR* (online ed, London, 27 January 2025).

241 Mark Dreyfus "New criminal laws to combat sexually explicit deepfakes" (5 June 2024) <<https://ministers.ag.gov.au>>; and Gilbert & Tobin Lawyers "Australian Government targets sexually explicit deepfakes" (29 July 2024) <www.gtlaw.com.au>.

so would require changes to state and territory law and is thus outside the Commonwealth's jurisdiction.²⁴²

Also in the process of enactment is the United States' Preventing Deepfakes of Intimate Images Bill 2023, an integrated civil and criminal regime that makes the intentional disclosure (or threatened disclosure) of NCIDs both a criminal offence and a right of private civil action.²⁴³ At time of printing, the Bill is before the House of Representatives' Committee on the Judiciary.²⁴⁴ Even if it makes headway, it still does not address NCID creation, leaving a large gap in the law.

V FINAL RECOMMENDATIONS: ESTABLISHING A CRIMINAL FRAMEWORK FOR NCIDS

No existing regulatory mechanism can be exclusively relied on to comprehensively and conclusively cover NCID perpetration in New Zealand. This article recommends lawmakers enact specific criminal provisions that prohibit the creation, distribution and threatened distribution of non-consensual intimate images, including those that have been digitally or otherwise manipulated. This is not to minimise the importance of civil law avenues in providing redress for individual victims. Given the diversity of victims' experiences, it is essential there is a suite of justice, redress and support options available to them.²⁴⁵ However, there remains a need for specific criminal offences, reflecting the urgency and weight of the individual and collective harms posed by NCIDs.

A Elements of the Offences

To effectively regulate conduct, the law must be expressed in clear and unambiguous terms, so that it can be widely known, understood, followed and enforced.²⁴⁶ Precision, however, must be balanced against the need for wording to be sufficiently flexible to extend to unforeseeable abusive behaviours which is crucial in this context, where modes of perpetration adapt rapidly alongside technological advances.²⁴⁷

1 Definitions

The proposed criminal offences should cover all "intimate images".

242 Karen Middleton "Jail time for those caught distributing deepfake porn under new Australian laws" *The Guardian* (1 June 2024) <www.theguardian.com>.

243 Joseph Morelle "Congressman Joe Morelle Announces New Support for his Legislation to Stop Deepfake Pornography" (24 April 2024) <www.morelle.house.gov>.

244 Congress.gov "HR1941 - Preventing Deepfakes of Intimate Images Act" (6 March 2025) <www.congress.gov>.

245 Henry and others, above n 28, at 874.

246 McGlynn and Rackley, above n 16, at 552; and Simester and Brookbanks, above n 30, at 1029.

247 Rackley and others, above n 26, at 311; and McGlynn and Rackley, above n 16, at 556.

(a) "Intimate"

Terms used in new legislation should cohere with existing legal concepts where appropriate.²⁴⁸ "Intimate" should be defined consistently with the "intimate visual recording" definition in the HDCA and the Crimes Act, which includes a range of sexual and non-sexual private activities.²⁴⁹ It should not require the images to be sexually explicit, which may unduly prevent the conviction of certain offenders.²⁵⁰

(b) "Images"

"Images" should be defined broadly to capture photographs, video recordings and digital images made in any medium using any device.²⁵¹ Crucially, to avoid any ambiguity, the definition should expressly include digitally or otherwise made or manipulated images.²⁵² A clear definition would signpost that Parliament intended to criminalise NCID abuse.

*2 Actus reus*²⁵³

(a) Distribution

The law should criminalise NCID distribution in any manner, including but not limited to the use of digital devices, online platforms and in-person communication; and to any person, including the victim themselves. The law should focus on the fact of distribution, regardless of whether the perpetrator is the creator of the image or merely came into possession of it.²⁵⁴ This approach would recognise that everyone involved in the image's circulation is responsible for, and contributes to, the victim's harm.

248 Simester and Brookbanks, above n 30, at 1029.

249 Harmful Digital Communications Act, s 4 definition of "intimate visual recording"; Crimes Act s 216G(1).

250 McGlynn and Rackley, above n 16, at 540.

251 See for example Harmful Digital Communications Act, s 4 definition of "intimate visual recording".

252 See for example Sexual Offences Act (UK), s 66A(5)(a).

253 Since the time of writing, additional literature has emerged addressing the solicitation of non-consensual intimate deepfakes; namely "assisting and encouraging another to commit an offence including creating or distributing of a NCID". In light of this, I consider that soliciting the creation of NCIDs should also be captured within the actus reus of the relevant offence. For further discussion, see Gemma Davies and Clare McGlynn "Soliciting the Creation of Sexually Explicit Deepfakes: Analysis of the current criminal law, loopholes and reform options" (2025). See also Non-Consensual Sexually Explicit Images and Videos (Offences) Bill 2024 (HL Bill 26), cls 2.

254 Salter and Crofts, above n 137, at 10.

(b) Creation

The law should criminalise NCID creation as a standalone offence. Although the impact of dissemination is arguably greater, a victim's sexual autonomy, dignity and privacy is nevertheless violated when intimate images of them are made without their consent. These victims are equally worthy of legal protection. An approach that only criminalises distribution would fail to dissuade potential offenders from creating images in the first place.²⁵⁵

(c) Threatened distribution

Threatening to distribute images can cause significant harm to victims, even if those threats are never carried out.²⁵⁶ Threats are not covered by the criminal law in many international jurisdictions, unless prosecuted under separate blackmail or extortion laws.²⁵⁷ However, victims have described the effects of threats as "paralysing",²⁵⁸ highlighting how threats can be detrimental to a victim's mental wellbeing. Further, threatened dissemination often occurs within a broader pattern of coercive control,²⁵⁹ used as a tool to perpetrate further harms. Therefore, the law should criminalise threats of distribution.

(d) Lack of consent

Lack of consent is central to all sexual violence criminal offences and is a necessary element of the proposed offences. It is the fact that consent is absent that makes the creation or dissemination of intimate images harmful and perpetrators who engage in such conduct culpable.

3 *Mens rea*

There ought to be no "motivation threshold". Focus should be on whether the offender knew or ought to have known that the victim did not consent, rather than an intention to cause harm to the victim, in line with other sexual offences.²⁶⁰ Including a "constructive knowledge" standard would cover situations in which the victim was unaware their data was being used, and where the offender ought to have known they would not consent to it being used, to create an intimate deepfake.

255 New Zealand's objectionable publications regime similarly criminalises both the making and distribution of objectionable material. See *Films, Videos, and Publications Classification Act 1993*, s 123(1)(a) and (d).

256 Rackley and others, above n 26, at 309.

257 At 297 and 310.

258 At 309.

259 At 310; and McGlynn and Rackley, above n 16, at 540.

260 See for example *Crimes Act*, s 128(2). See also *Sexual Offences Act (UK)*, s 66B.

4 *Result*

The proposed offences should not require that the relevant NCIDs cause serious harm to the victim or that they would cause harm to a reasonable person in the position of the victim. Exclusion would recognise that the harm to victims' sexual autonomy, dignity and privacy is inherent in NCID abuse, and protect victims from undue scrutiny.

B *Categorisation*

The proposed offences should be categorised in law as sexual offences.²⁶¹ There are three significant reasons for this. First, this would accurately reflect the nature of the harms caused by IBSA.²⁶² Secondly, it would encourage integration of prevention measures and community support services that aim to tackle violence against women more generally.²⁶³ Thirdly, in New Zealand, victims of sexual violence are afforded specific accommodations in court proceedings, including automatic name suppression.²⁶⁴ Providing anonymity may encourage more NCID victims to come forward and seek justice.

C *Penalties*

Understandably, the law should impose criminal sanctions on offenders found guilty under criminal provisions. Further, it may be appropriate to increase criminal penalties under certain circumstances where the offence is "aggravated",²⁶⁵ for example if the offender committed the offence with the intention to cause harm or for the purposes of sexual gratification.²⁶⁶ Like other sexual violence victims, some IBSA victims engaging with the legal system might want the perpetrator prosecuted and convicted, while others might prefer a restorative justice approach.²⁶⁷ There is scope under the criminal law to accommodate either preference.

V *CONCLUSION*

Non-consensual intimate deepfakes are a pernicious, complex and multifaceted form of sexual violence urgently requiring legislative and social intervention. Their significant harms justify enacting distinct criminal provisions that expressly prohibit the creation, distribution and threatened distribution of NCIDs. Enacting specific criminal legislation is the most appropriate and effective way

²⁶¹ McGlynn and Rackley, above n 16, at 556.

²⁶² At 556–557.

²⁶³ At 557.

²⁶⁴ Criminal Procedure Act, s 203.

²⁶⁵ Compare Films, Videos, and Publications Classification Act, s 132A, which takes a similar approach.

²⁶⁶ See for example Sexual Offences Act (UK), s 66B(2)–(3).

²⁶⁷ Henry and others, above n 28, at 872–873.

for the law to deal with the growing problem of NCIDs, as part of a broader integrated approach targeted at the underlying social attitudes and cultural conditions that condone men's violence against women, both online and offline. Lawmakers must heed this warning and act quickly to protect New Zealanders from this novel manifestation of sexual abuse before another sexual violence victim is left destitute by the criminal justice system.

