Examining the Social and Psychological Impact of Deepfakes

Rapid Evidence Review

July 2025

About Crest Advisory

We are crime and justice specialists - equal parts research, strategy and communication. From police forces to public inquiries, from tech companies to devolved authorities, we believe all these organisations (and more) have their own part to play in building a safer, more secure society. As the UK's only consultancy with this focus, we are as much of a blend as the crime and justice sector itself.

Crest Advisory (UK) Ltd is a company registered in England and Wales (08181317)
2 Bath Place, Rivington Street, London, EC2A 3DR
www.crestadvisory.com

Contents

Overview and Scope

Crest Advisory was commissioned by the Accelerated Capability Environment (ACE) on behalf of the Office of the Police Chief Scientific Adviser (OPCSA) to conduct research examining the social and psychological impacts of deepfakes on victims, with a focus on violence against women and girls (VAWG).

This rapid evidence review compiles relevant literature which informed our lines of enquiry and refined the scope of our primary research and engagement, including a public attitudes survey. This document has been iterated throughout the commission to ensure it is up to date at the time of writing (July 2025) and captures relevant emerging literature.

Deepfakes refer to any audio, image or video which has been digitally altered using machine learning methods. This includes fraudulent, political, or humorous content, as well as intimate images and pornography. However, in line with the focus of this commission, our evidence review focuses on deepfake violence against women and girls (VAWG). This focus reflects evidence that the vast majority of deepfake videos are sexualised in nature, with women being the disproportionate target of this abuse.

I: The scale and nature of deepfakes and violence against women and girls (VAWG)

The scale and nature of deepfakes

Summary

- First defined in 2017 on Reddit to mean artificially generated or manipulated nude images of celebrities, the term 'deepfakes' has been firmly entwined with image-based sexual abuse and synthetic sexual imagery since its creation
- Since 2017, the technology used to create deepfakes has exponentially developed, correlating with an explosion of interest from users. The accessibility and democratisation of this technology has meant that any individual can create, or commission the creation, of deepfakes at a very low cost and in a very short space of time
- Now there are dedicated apps and platforms dedicated to creating sexual deepfakes, and mainstream porn sites continue to host deepfake content despite its illegality
- The scale of deepfake violence against women and girls (VAWG) is difficult to calculate due to poor data collection and underreporting: the Crime Survey for England and Wales does not explicitly ask about deepfake victimisation (of any form); and many will not be aware that they are victims, leading to underreporting where this data in collected. Therefore the scale of the issue is likely to be significantly underrepresented
- The law has been slow to keep pace with the rapid rise of deepfakes. However, a surge in legislative activity over the last year, driven by campaigning efforts, and the approval of the Online Safety Act, has brought the UK closer to a fit for purpose legislative framework that reflects the accessibility and popularity of deepfakes

Definition and nature of 'deepfake' VAWG

The term 'deepfake' was first coined in 2017 on Reddit in a forum dedicated to developing synthetic (artificial) sexual imagery of celebrities (Deeptrace, 2019; Maddocks, S., 2020). Since then, its definition and name have been disputed. For many, 'deepfake' is a catch-all term to mean any audio,

image and video content that has been manipulated or created using machine learning methods to alter how a person, object or environment is presented (Centre for Data Ethics and Innovation, 2019; Sippya et al, 2024). This can include fraudulent, political, or humorous content, as well as intimate images and pornography. As the term was coined by a perpetrator of non-consensual synthetic sexual imagery, many organisations and academics have moved to rename 'deepfakes', offering alternatives such as AI-generated sexual images, or synthetic sexual imagery (Henry & Beard, 2024). We use 'deepfakes' throughout this report to reflect the wording of the commission.

From their inception, sexual/intimate deepfake videos have disproportionately impacted women and girls. A 2019 report based on analysis of 14,678 deepfake videos across a number of platforms and sites found that 96% were sexual and of women (Deeptrace, 2019). A further 2023 study identified that deepfake pornography makes up 98% of deepfakes online, and 99% of those are of women (Security Hero, 2023). Given that sexual deepfakes – almost all non-consensual – comprise the vast majority of all deepfake videos on the internet, there is an urgent need to focus research and response on this area. This rapid evidence review is focused on deepfakes as they intersect with violence against women and girls (VAWG), the definition of which is explained in the next section.

Scale of 'deepfake' VAWG

Initially created and disseminated through Reddit, the quantity and ease of access to deepfakes has drastically increased. As of 2019, deepfakes were hosted by 8 in 10 of the top porn websites, alongside a swathe of bespoke deepfake pornography sites (Deeptrace, 2019). Deepfakes are also accessible on mainstream platforms via Google, and can be distributed through Telegram, on X, and in apps. Research from the Oxford Internet Institute found that there are nearly 35,000 available AI tools, designed specifically to create deepfake images of identifiable people, which have been downloaded nearly 15 million times since late 2022 (Hawkins et al., 2025).

Within this ecosystem, there is an economy around the commissioned creation of deepfakes (Tenbarge, 2023). In various dedicated forums and platforms, users can create bespoke deepfakes for a small charge in seconds (Deeptrace, 2019; Security Hero, 2023; McAfee), using mainstream payment methods such as Visa, Mastercard, and cryptocurrency. As per a 2024 study, the most prominent deepfake platform (MrDeepFakes) hosted 43,000 sexual deepfake videos depicting 3,800

individuals, with these videos watched over 1.5 billion times (Han et al., 2024).However, as of 6th May 2025, MrDeepFakes has ceased to operate after the support of a “critical service provider” was permanently lost (The Independent, 2025).

The scale of deepfake VAWG is challenging to calculate as many victims are unaware that their audio, image, or video has been used and manipulated without their consent unless they happen to see it or are notified by someone else. Even if victims are aware, the challenge to understand the scale is exacerbated by the fact that deepfake VAWG is not currently captured in mainstream crime recording. In April 2024, some police forces adopted the National Data Quality Improvement Service tool, which helps automatically flag offences such as domestic abuse and child sexual abuse, to identify offences that have an online element (ONS, 2024). However, it is unclear whether data on the online element will be collated to understand the scale of deepfake VAWG. The Crime Survey for England and Wales does not explicitly ask adults (aged 16+) or children (aged 10 - 15) whether they have experienced any form of deepfake victimisation. This lack of data collection further contributes to the challenge of understanding the scale of deepfakes.

Estimates suggest that as of February 2024, there are 276,149 non-consensual sexual/intimate deepfake videos, representing a 1,780% increase since 2019 (Compton & Hunt, 2024). Viewership has increased by 3,942% since 2019, estimated to exceed a total of 4.2 billion views (Compton & Hunt, 2024).

More broadly than deepfakes, there are pockets of information about the scale of intimate image abuse more generally (i.e. including real or synthetic images). Since 2015, Revenge Porn Helpline, which supports adults in the UK to remove content, has reported approximately 338,000 intimate images to platforms for removal – this includes both authentic and deepfake content. Additionally, a survey conducted by Refuge in the UK, which used a representative sample of adults (n = 2060) in England and Wales, found 1 in 14 adults (4.4 million people) had experienced threats to share their intimate images or videos (Refuge, 2020).

Legislative and regulatory frameworks

While the technology around deepfakes has rapidly developed, legislation has been slow to keep pace. However, recent legislative reforms have included:

- In 2023, the Online Safety Act made it an offence to share non-consensual intimate images, including deepfakes
- The Data (Use and Access) Act (2025) includes provisions to supplement the Online Safety Act (2023) by making it illegal to create an intimate image of another person without consent. However, at the time of writing, this section of the Data (Use and Access) Act is not yet in force.
- The Crime and Policing Bill (2025), currently at its second reading in the House of Lords, proposes to provide police with further powers to tackle intimate image abuse more broadly (and inclusive of deepfakes).
- From February to May 2025, Ofcom held a consultation to inform the development of guidance to support online services to better protect women and girls on their platforms (Ofcom, 2025).

There remain key gaps in deepfake legislation, highlighted in particular by advocates. This includes the need for legislation to be culturally sensitive, including material that is considered "culturally intimate" for the victim, such as a Muslim woman being pictured without her hijab if she habitually wears one in public (House of Commons, 2025).

Literature on the symbolic and practical role of legislation and regulation in preventing and reducing the harms of deepfake VAWG is explored in 'The role of legislation'.

The scale and nature of violence against women and girls (VAWG)

Summary

- Violence against women and girls (VAWG) has been described as an epidemic. The scale of VAWG is increasing and can have a detrimental impact on victims

- The rise of deepfake technology has introduced a new method of gender-based violence
- The vast majority of deepfake media are sexualised in nature, with women being the disproportionate target of this abuse
- There is limited evidence on perpetrators of deepfake VAWG, although studies on perpetrators of intimate image abuse more generally offer some insight.

Definition and scale of violence against women and girls (VAWG)

This evidence review adopts the National Policing Chief's Council's (NPCC) definition and understanding of VAWG. This means VAWG includes a range of offences, including domestic abuse, rape and serious sexual offences, tech-enabled VAWG such as online stalking and harassment, and child sexual abuse and exploitation. Whilst men and boys can also experience these harms, women and girls are disproportionately affected.

VAWG in England and Wales has been described as an epidemic, reflecting its scale, complexity and profound impact on victims (National Policing Statement for Violence Against Women and Girls, 2024). Between 2018/19 and 2022/23, police-recorded VAWG-related crimes increased by 37%, with an estimated 2 million women victimised each year, and 2.3 million perpetrators annually (National Policing Statement for Violence Against Women and Girls, 2024). These figures likely underestimate the true scale of harm due to underreporting. In recognition of the severity of this crisis, policing has adopted the 4Ps framework (Pursue, Prepare, Protect, and Prevent) to drive a consistent national response. This strategic approach prioritises relentless pursuit of high-harm perpetrators, enhanced protection for victims, and proactive prevention of VAWG at a systemic level (The National Framework for Delivery, 2024). The integration of this model underscores VAWG as a national policing threat, requiring the same level of coordination and urgency as serious organised crime and terrorism.

The impact of VAWG is severe and far-reaching, with both immediate and long-term consequences for victims' mental and physical health. In the year to March 2023, one in six homicides in England and Wales were domestic abuse related, and the Domestic Homicide Project found a year on year increase in suspected victim suicides, nearly doubling across a three-year dataset (National Policing Statement for Violence Against Women and Girls, 2024 / Vulnerability Knowledge and Practice

Programme, 2022). More than 50% of domestic abuse victims report mental health issues, while 75% of sexual assault survivors meet the criteria for post traumatic stress disorder within one month of the offence (National Policing Statement for Violence Against Women and Girls, 2024). The long term emotional toll of VAWG is profound: 63% of women who were raped or sexually assaulted since age 16 report lasting emotional or mental health issues, and 10% reported suicide attempted as a result (National Policing Statement for Violence Against Women and Girls, 2024 / ONS, 2021).

The rise of 'deepfake' VAWG

The rise of deepfake technology has introduced a new method of gender-based violence. Lucas (2022) summarises this issue, drawing on research by Dunn (2021) and Hao (2021): "The use of deepfake technology to create non-consensual sexual deepfakes is a violence-against-women issue and deepfakes provide a relatively new way to deploy gender-based violence." This aligns with broader research showing that over 90% of deepfake video content is sexualised, with women overwhelmingly the target (Security Hero, 2023). Non-consensual synthetic intimate images (NCII) have become a new frontier of VAWG, enabling perpetrators to commit AI-enhanced harassment, exploitation, cyber-mobbing and abuse (Dunn, 2021; Hao, 2021). Women in the public eye are frequent targets who suffer large-scale victimisation, though the majority of victims are ordinary women (Compton & Hunt, 2024).

Perpetrators of 'deepfake' VAWG

Understood by policing as an online VAWG offence, it may be that deepfake VAWG perpetrators have a similar profile to other VAWG offenders, who are overwhelmingly male, and are often slightly older than their victim (VAWG STRA, 2024). Though there is limited knowledge of deepfake VAWG perpetrators specifically, evidence exists around perpetrators of intimate image abuse more generally. Henry and Beard's (2024) review of 26 studies published between 2013-2023 on intimate image abuse perpetrators provides a useful summary. The review found that men, young adults and members of the LGBTIQ+ community were more likely to perpetrate this offence. Motivations were complex, although similar to other VAWG-related offences, and included social reward, power dynamics, sexual gratification, and retaliation. The review highlighted that some studies found a

positive relationship between perpetration and endorsing rape myths, as well as an association with Dark Tetrad traits. These traits refer to characteristics of sadism, psychopathy, narcissism, manipulation, and self-interest. This evidence places deepfake VAWG perpetration within wider patterns of offending, likely on the same victim. Other evidence suggests that some deepfake creators are prolific - one creator alone is known to have posted over 1,800 videos (Compton & Hunt, 2024). The limited evidence base on deepfake VAWG perpetrators makes it challenging to develop effective responses, and in particular prevention tactics.

II: Public awareness and attitudes

Summary

- Public attitudes surveys exploring awareness and perceptions of deepfakes have centred on political and fraudulent deepfakes, rather than deepfake VAWG
- There are a number of surveys that assess public awareness of deepfakes. However, the findings of research on public awareness vary
- The public's level of interaction with deepfakes is more clear; as many as half of adults in the UK have seen a deepfake, according to a 2024 Hiya survey. However, we have only identified two surveys that have specifically assessed the proportion of adults who have seen pornographic deepfakes
- Data on the experience of victimisation of deepfakes is focused on children and young people. Two surveys have found that around 1 in 10 young people have been victims or know someone who has been a victim of 'deepfake nudes'
- There is some research on the public's attitudes towards deepfakes. This evidence primarily demonstrates that the public is concerned about the harms of deepfakes
- Some studies assess which harms related to deepfakes hold the most concern for the public. They have found that the public are concerned about the emotional and psychological impacts on victims; the potential for reputational damage; the increasing quality of content which may lead viewers to think the content is authentic; increasing distrust in information; manipulation of public opinion; and increasing misogyny and online VAWG. Women appear more concerned about these risks than men

- There is a gap in the literature around attitudes to deepfake legislation

A number of public attitudes surveys have explored awareness and perceptions of deepfakes, including the harms associated with deepfakes. However, there are few reliable findings on public attitudes towards deepfake VAWG specifically; thus far, public attitudes surveys have centred on 'political' and fraudulent deepfakes, as opposed to pornographic (which are the majority of all deepfakes) (Maddock, 2020), and more of the public appear to have seen fraudulent, political, and satirical deepfakes than sexual (Ofcom / YouGov, 2024).

The need to better understand awareness and attitudes has been underscored by the Alan Turing Institute, which noted that "understanding people's perspectives on data and AI is a key component in ensuring these technologies are developed to align with societal values, and are able to address societal needs" (Alan Turing Institute, 2024)

Awareness of deepfakes

Findings vary on the extent to which the public is aware of deepfakes. There is currently limited understanding on how public awareness has changed over time and factors that have impacted levels of awareness. A 2024 survey of over 1,400 UK adults (18+) found that 82.7% of participants had heard of the term deepfakes (Sippya et al, 2024) suggesting a high level of general awareness. By contrast, a 2019 survey from iProov found that almost three-quarters (72%) of the UK public (n = 2000) had never heard of deepfake videos (iProov, 2019).

Globally, a GASA/Feedzai survey (n = 58,329) found that respondents were less aware that AI was able to generate a video (65%) when compared with mimicking a voice, generating a dialogue, and creating an image (GASA & Feedzai, 2024). Another study (conducted in Australia, Belgium, Denmark, France, Mexico, Netherlands, Poland, South Korea, Spain and the USA) found that men were more aware of deepfakes than women (Umbach et al., 2024).

Results from a survey of 1,200 young people aged 13 to 20 suggests that awareness of deepfakes amongst this cohort is more limited, finding that 41% had heard the term "deepfake nudes,"(Thorn & Burson, 2025). A 2024 UK survey found understanding to be limited, too - almost two-thirds of

children (n = 1000) and almost half of parents (n = 2000) said that they don't know or understand the term 'deepfake' (Internet Matters, 2024).

There have not yet been detailed studies on which population groups are most aware of deepfakes, and how they have received this information in the UK. In Germany, one study identified that variables such as youth, male gender, and educational attainment are positively related with deepfake knowledge, as well as digital skills and social media usage (Bitton et al., 2024).

Experience of deepfakes

There is more substantial data on the public's interactions with deepfakes. One Hiya survey (n = 2000) found that in the UK (Hiya, 2024):

- 39% had experienced an audio deepfake on a personal call
- 36% had experienced a deepfake video on Facebook
- 28% had experienced a deepfake video on Instagram
- 14% had experienced an audio deepfake on Facebook
- 14% had experienced an audio deepfake on YouTube
- 14% had experienced an audio deepfake on a work phone call
- 13% had experienced a deepfake video on a video call at their job.

The Alan Turing Institute and Oxford Internet Institute survey (n = 1400) found that: "50.2% of people have personally encountered deepfakes of public figures from the entertainment industry such as actors, social media influencers and/or musicians, while exposure to deepfakes of politicians was at 34.1%. While nearly half of all respondents (49.3%) had encountered deepfakes created for educational or entertainment purposes, 15.8% of all respondents were exposed to deepfakes that could be potentially harmful, including deepfakes that spread health or religious misinformation or propaganda. Furthermore, exposure to harmful deepfakes, including non-consensual deepfake pornographic images/videos (18.8%) and deepfakes that falsify identities for frauds/scams (9%) were relatively low, though still high enough in absolute terms to indicate a considerable proportion of the population encountering them." (Sippy et al, 2024). Another survey (n = 1081) identified that 14% of adults had seen a sexual deepfake, reaching 24% in the 18-24 demographic (Ofcom / YouGov, 2024).

Thorn & Burson's survey (n = 1200) found that 10% of teens (aged 13-20) reported personally knowing someone who had deepfake nude imagery created of them, and 6% disclosed having been a direct victim of this form of abuse (Thorn & Burson, 2025). Another study of children's experiences with deepfakes (n = 1000) found that teenage boys (18%) are twice as likely as teenage girls (9%) to report experience with nude deepfakes (Internet Matters, 2024). A 2024 ESET survey (n = 2004) revealed that 1 in 10 young people reported being a victim of deepfake pornography, knowing a victim, or both (ESET, 2024). Ofcom/YouGov's survey found that among those who have seen a sexual deepfake, 64% (n = 92) were of a celebrity, 42% were of a stranger (n = 60), and 15% were of someone they know (n = 21) (Ofcom / YouGov, 2024).

There is some data on the public's use of technology to create deepfakes. One survey (n = 1400) identified that usage of deepfake creation technologies is relatively low (8.1% of respondents) (Sippya et al, 2024). Thorn & Burson's 2025 survey (n = 1200) found that the limited sample (2%) of young people who admit to creating deepfake nudes of others described easy access to deepfake technologies.

Attitudes to harm

There is some research exploring public attitudes towards deepfakes and, in particular, perceptions of the harms caused by deepfakes. The Thorn & Burson 2025 survey of young people (n = 1200) found that 84% believe that deepfake nudes cause harm. The harms identified by respondents included emotional and psychological impacts on victims, the potential for reputational damage, and the increasing quality of the imagery which may lead its viewers to think that the content is authentic (Thorn & Burson, 2025).

The Alan Turing Institute and Oxford Internet Institute survey (n = 1400) found that 90.4% of respondents were either very concerned or somewhat concerned about the spread of deepfakes, primarily due to child sexual abuse concerns, followed by distrust in information, manipulating public opinion, and increasing misogyny and online VAWG (Sippy et al, 2024). iProov (2022, n = 2000) found that over 60% of respondents agree that deepfakes are dangerous. However, the primary

concerns of respondents did not relate to online VAWG, and more commonly respondents were concerned with misinformation.

There is little information on how demographics, behaviours, or attitudes to other concepts may impact an individual's attitudes to deepfakes. A 2024 survey (n=2004) revealed that 57% of under 18s are concerned about being a victim of deepfake pornography, compared with 61% of adult women and 45% of adult men (ESET, 2024). One German study identified that women had a much greater perceived risk of becoming a victim of a harmful deepfake than men, and more women than men felt that they had to protect their personal data more strictly. Men were more likely to agree that deepfake technology had the potential to be used to create pornographic content. This study also found that women were more concerned about the risk around deepfakes. Those with greater digital skills and internet application use were also more concerned about deepfake harms, while there was no significant association for age, education level, or income (Bitton et al., 2024). There is a need for UK-based research to determine the applicability of these findings.

There are few studies on public attitudes towards deepfake legislation. However, one study not based in the UK (conducted in Australia, Belgium, Denmark, France, Mexico, Netherlands, Poland, South Korea, Spain and the USA) found that non-consensually filming or photographing someone was seen as significantly more deserving of criminalization as compared to making deepfake pornography (Umbach, R. et al., 2024).

III: The social and psychological impacts of deepfake VAWG

This section looks at the impact of deepfake VAWG at an individual and societal level. It explores the evidence base on who is affected and the state of reporting, as well as the social, psychological, and practical impacts on victims, and the broader impact of deepfake VAWG on society.

Who are the victims of 'deepfake' VAWG?

Summary

- Evidence suggests that deepfakes disproportionately impact certain groups, and literature increasingly points to certain demographic characteristics that are associated with a greater risk of victimisation
- Women are more impacted by synthetic sexual content than men, while men may be more likely to be victims of other types of deepfakes
- Younger people and women with a public profile appear at greater risk of deepfake victimisation, while Black and minority ethnic women and sex workers may experience victimisation and reporting differently to other women
- Ultimately, many of these factors likely intersect, rendering some women especially vulnerable. However, the literature in this space is not yet robust enough to be conclusive.

There is some literature on who is most affected by deepfake VAWG. However, there is limited evidence on the reasons why some groups are disproportionately impacted, and further research to understand victimisation is necessary.

Gender

Evidence suggests that the type of deepfake bears substantial weight on the likely gender of its victim. Almost all victims with content on deepfake pornography sites are women (96%), however non-pornographic deepfakes on YouTube are majority male (61%) (Deeptrace, 2019).

Ethnicity

There is a limited amount of evidence on the ethnicity of deepfake VAWG victims. However, just as the disproportionate victimisation of women is aligned with wider misogyny in society and perpetrators, there is evidence that Black and minority ethnic women are disproportionately victims of gendered online abuse as well (Glitch 2023). More broadly, there have been studies into the representation of Black women in pornography that demonstrate that pornographic content involving Black women is more aggressive and violent than content involving White performers (McGlynn et al, 2024; Fritz et al., 2021), leading to harmful sexual objectification and racial stereotyping.

Age

There is evidence that deepfakes disproportionately impact younger people (Flynn et al., 2021). In Flynn et al's study which surveyed people across the UK, Australia and New Zealand, found 20.4% (n = 76) of respondents aged 16-19 years had experienced deepfake abuse of a sexual nature - including digitally altered nude and sexual imagery abuse victimisation - compared with 3.4% (n = 54) of respondents aged 50-64. However, the study notes that younger people may be more likely to self-report.

Nationality

There is limited research on the nationality of victims of deepfake VAWG. One Deeptrace study found that over 90% of deepfake videos on YouTube (including but not limited to sexual content) featured Western subjects. However, non-Western subjects featured in almost a third of videos on deepfake pornography websites, with South Korean K-pop singers making up a quarter of the subjects targeted (Deeptrace, 2019). Since this 2019 study, news sources indicate a further global spread of deepfake VAWG. The differences in type of deepfake victimisation and experiences of victimisation by nationality have not yet been studied.

Profile and profession

Most deepfakes hosted on mainstream sites are of public figures involved in the entertainment industry – in particular, female actresses and singers (Deeptrace, 2019). Increasingly, female politicians (Waterson, 2024) and journalists (Roberts, 2025) have been targeted in the UK, demonstrating the dangerous intersection between deepfake VAWG and the use of 'political' deepfakes to spread misinformation and to undermine the legitimacy of public institutions, and women in positions of power.

Sex workers are substantially impacted by deepfake VAWG. Traditional pornographic deepfakes involve at least two victims, one of whom is likely to be a sex worker and/or performer whose image and likeness has been unconsensually taken for the purposes of victimising another woman (Cole, 2024). Studies involving sex workers and their experiences of deepfake victimisation are scant, and there is a clear evidence gap in this area.

Victims' experience of reporting

Summary

- The known scale of deepfake VAWG victimisation is likely to be an underestimate, due to barriers in reporting
- Most victims will never know that they are victims of deepfake VAWG, therefore hindering their ability to report
- There are also some known barriers to victims' reporting. The evidence base is more substantial on barriers to reporting VAWG more generally, while there is some evidence on barriers to reporting deepfakes
- Barriers to reporting deepfakes include not knowing where to report, not having the legislative framework or police capability to pursue prosecution, and challenges identifying, removing, or retaining the deepfake content
- Addressing these barriers will improve our understanding of the scale and nature of deepfakes and the needs of victims.

The evidence on victims' experiences of reporting deepfake VAWG, and barriers to reporting, is limited. More broadly, evidence from a UK-based survey (n=2000) suggests that around half of people would tell others that they have been a victim of deepfakes (Hiya, 2024). However, it is not clear how this would translate into reporting to an authority, including the police.

There is a wealth of evidence on reporting behaviours of victims (more widely than VAWG victims), including barriers to reporting. The 2022 Victim Survey (n = 489) (the Victims' Commissioner, 2022) identified that lack of faith and trust in the criminal justice system (CJS) was a key reason why victims were not reporting to the police. Only 8% of victims surveyed were confident that they could receive

justice by reporting a crime and, of those who had reported a crime, 37% of women said they would not report a crime to the police again – this is higher than men (29%). Barriers to reporting identified in this survey included a fear that victims would not be believed or taken seriously, fear, dislike or distrust in the police and CJS, and poor prior experiences of the police and/or CJS.

Barriers to reporting VAWG crimes in particular were identified in a 2022 report by the British Criminology Society. Key barriers include:

- A lack of confidence that reporting would help (APPG-UNW, 2021; Solymosi et al, 2017)
- Fear of being blamed for one's victimisation or perceived as an underserving or precipitous victim (Brooks, 2014; Vera-Gray & Kelly, 2020);
- Doubting that an experience was serious enough to merit reporting (APPG-UNW, 2021; Solymosi et al, 2017);
- Ambiguity regarding what defines violence and/or harassment (Ball & Wesson, 2017);
- Delayed recognition that an experience constitutes sexual violence (Waterhouse et al, 2016).

One study specifically focused on victims' experiences of deepfake victimisation within a sexual context, including digitally altered nude images and sexual imagery abuse, identified barriers to reporting (Flynn et al., 2021), including:

- **System barriers,** including perceived gaps in the law and legislative framework for reporting, and jurisdictional challenges
- **Practical barriers,** such as the onus on victims to produce and procure evidence of deepfake abuse, and to have the content removed (or, in some cases, to keep the content up to support investigations) (Glamour, 2024) and limited police resources
- **Other barriers,** including victim-blaming attitudes.

Statutory agencies benefit from deepfake VAWG being reported, as this contributes to a clearer understanding of the threat landscape, influencing how intelligence is gathered and informing prevention and remediation efforts (Burton et al., 2025). We know that many victims choose to report to support services and helplines, as opposed to the police (see 'Specialist Support Services'). To date, there has been little analysis into how victims' experiences of reporting, and associated

outcomes, compare across different agencies and services.

Impacts on victims of 'deepfake' VAWG

Summary

- The psychological impacts of VAWG deepfakes are the most well-documented, and link strongly to the impacts of other forms of image-based sexual abuse, and sexual abuse more generally. These include PTSD, depression, anxiety, self-image, and feelings of humiliation and violation.
- Victims can face economic and professional harms as a result of deepfake VAWG, and victimisation may have material consequences on women's decision to be 'online' and their privacy.
- There is currently limited literature linking deepfake abuse with offline violence.
- Social impacts are frequently reputational, and may have an outsized impact on women from particular cultures where modesty is highly valued. Victims may withdraw from social life, and their relationships with family, friends, and employers may become strained.

There is an emerging body of literature around the impacts of deepfake VAWG on victims. Broadly, these impacts can be divided into the following categories: psychological, social/reputational, and physical.

Psychological impact

The psychological impacts of deepfake VAWG are best understood. Primary research suggests these frequently include high levels of stress, depression, anxiety, low self-esteem, insecurity, paranoia, obsessive behaviour, and suicidal thoughts (Huber, 2022). Victims have also reported feelings of humiliation, violation, fear, helplessness, and powerlessness (O'Malley et al., 2020). Similarly, victims of intimate image abuse report PTSD, depression, anxiety, suicidal ideation (Bates, 2017; Citron, 2019; Deeptrace, 2019; McGlynn et al., 2019,; Powell, et al., 2022), and damage to

their sense of self (Clevenger & Navarro, 2021; Jankowicz, 2021). Several women have taken their lives as a result of deepfake VAWG (Compton & Hunt, 2024).

These symptoms mirror the effects of sexual harassment and contact VAWG offences, like sexual assault and rape. For example, 'revenge porn' was found to have similar effects to sexual harassment on a victim's mental health, including high levels of stress, alcohol use, PTSD, clinical depression, and placing blame themselves for the assault (Bates, 2016). Image-based sexual abuse was found to cause a "feeling of vulnerability…distrust…and being unsafe" that "affects you potentially forever", which speaks to the lived experience of rape (McGlynn et al., 2020)

Indeed, some victims have described the impact of image-based sexual abuse as akin to physical sexual assault (Huber, 2022):

"*It made me feel violated, it actually made me feel raped almost.*"

"So, I felt helpless, I felt abused like, that's not the best word but I felt as if I'd physically been abused even though obviously, I hadn't been physically abused but it's just that sick feeling that I had. It was as if somebody had actually punched me in the stomach…it's just, it's complete devastation"

Though this evidence draws parallels with the psychological impact of other VAWG offences, some impacts of deepfake VAWG relate specifically to it being an (online) image-based offence. Namely, evidence suggests that revictimisation occurs when images are spread (Knipschild, 2024). Victims are not only impacted by the creation of the image, but also by the ongoing threat that the image will be spread and seen. These possibilities create an uncertainty beyond the victim's control, making it difficult for them to find closure (Nygard, 2024). This kind of revictimisation is more prevalent in victims who are concerned about the social ramifications of the offence (for example, fears that they will be treated differently by those who see the image (Nygard, 2024)).

Other impacts specific to deepfake VAWG include the perpetrator being "faceless", on the other side of a screen and often unknown to the victim:

"That an unknown perpetrator has used readily available technology to fantasise about forcing me into an array of sexual activities can only be described as a violation [...] It's haunting, not least because whoever has abused me in this way is out of reach, faceless and therefore far beyond the limited sanctions presently available." (The Times, 2024)

Scholars note that whilst it is validating to see the psychological impact of deepfake VAWG through a trauma lens, it is important to recognise that the experience of deepfake victimisation should not be confined to the medical, and is interconnected with a victim's wider social identity (McGlynn et al., 2020). Psychological impacts bleed into other spheres: the victim's paranoia regarding online public spaces can morph into paranoia in offline public spaces (Huber, 2022).

Social impact

Social and interpersonal impacts are also well-documented. Damage or fear of damage to reputation is frequently discussed in literature on deepfake VAWG. Where victims are afraid that family, friends, colleagues, and employers will see deepfake content of them – in particular, sexual deepfakes – they may withdraw from these relationships (NPCC, 2024). Victims can experience fears about who they can trust, who has seen the deepfake, and paranoia about how the image may spread (Compton & Hunt, 2024). There can also be material repercussions with regards to a victim's job and relationships when family, friends, colleagues, and employers become aware of the deepfake content (Paris & Donovan, 2019).

Another study has identified that image-based sexual abuse causes victims to censor themselves and withdraw from online spaces, which has substantial implications for victims' online citizenship and the digital divide between men and women (Rigotti et al., 2024).

Physical impact

Physical impacts tend to follow from the psychological and social impacts noted above, causing physical health problems for some victims. There have been instances where women have died by suicide because of deepfake abuse, and fear of the repercussions of this abuse (Jankowicz, 2021, Lucas, 2022). There is currently limited literature linking deepfake VAWG with offline VAWG.

Ultimately, most of these impacts occur when a victim becomes aware that they have been victimised. However, this is not the case for the majority of victims, who are unlikely to ever know that deepfakes of them exist (Burgess, 2020). This could either be because their deepfake has never entered the public domain, or because no one has informed the victim or reported it to the authorities.

Societal Impacts

Summary

- Evidence suggests society is negatively impacted by deepfakes on main two counts;
- Firstly, deepfakes are inseparable from and compound existing systemic gender inequalities. The technology introduces and deepens the power asymmetries which largely fall along gendered lines.
- Secondly, deepfakes also introduce new problems for society, calling into question the credibility of audio-visual information, incurring economic costs and hampering individual freedom of expression.
- Evidence gaps persist around the costs of VAWG deepfakes on society. This data would support the Government and police to respond to this new form of abuse.

There is some evidence on the societal challenges created by the emergence and proliferation of deepfake VAWG. This evidence is largely theoretical but can be substantiated by empirical evidence, particularly evidence of public attitudes towards the creation, consumption and criminalisation of deepfakes. Evidence on the societal harms of deepfake VAWG can be broadly grouped as:

1. Harms which are inseparable from, and compound existing, systemic gender inequalities
2. New challenges presented by deepfake technology for VAWG

Societal harms linked to existing systemic gender inequalities

At its root, scholars argue that deepfake VAWG is a product of systemic gender inequalities (Brown & Fleming, 2020; Jacobsen & Simpson, 2022; McGlynn & Toparlak, 2024). Deepfake VAWG is deeply rooted in systemic misogyny, rather than being traceable to the perpetrator alone (Ohman, 2020). Therefore, academic evidence suggests that deepfake VAWG is both a result of systemic misogyny, while also perpetuating gendered violence by objectifying women through a "new voyeurism" (McGlynn & Toparlak, 2024). Instances of deepfake VAWG are therefore regarded by scholars as "inseparable" from systemic misogyny, since they uphold the "cultural scaffolding" which normalises sexual violence against women (McGlynn & Toparlak, 2024, p11). Scholars thus argue that deepfakes impact wider society by perpetuating cycles of misogyny and gendered violence.

Much of the evidence which looks at the link between societal culture and deepfake VAWG examines how deepfakes perpetuate the misogynistic attitudes that have enabled this crime to proliferate. Firstly, some scholars suggest society insufficiently condemns the distribution and consumption of deepfake porn, normalising its circulation on large-scale social media platforms like Twitter (Maddocks, 2020). Algorithms also play a role in this normalisation. For example, algorithms can promote extreme misogynistic content to *"impressionable young men"*, allowing these ideologies to spread (Hornle, 2025).

Alongside the collective failure to condemn deepfake porn, scholars identify a lack of collective action. The lack of collective action around the threat of deepfakes "normalises, trivialises and even condones" (misogynistic) sexual violence (McGlynn & Toparlak, 2024). The narrow legislative definition of offences relating to non-consensual intimate images enables perpetrators to maintain power over victims through possessing (without distributing) these images (Hornle, 2025). By failing to act with urgency, legal and criminal justice scholars suggest that the systemic gender inequalities that exist in the offline world have been compounded by deepfake VAWG: for example, the sharing of deepfake sexual/intimate images can harm the reputation of victims (mostly female) whilst allowing the perpetrators (mostly male) to maintain some anonymity, if just temporarily (Huber, 2022). Mixed messaging and the collective failure to act with urgency leaves women and girls vulnerable to this "invisible threat", a fear that sexual digital forgeries might be created of them with little consequence to perpetrators (McGlynn & Toparlak, 2024).

The theoretical link between deepfake VAWG and gendered violence is substantiated by a small but growing body of empirical evidence. One source surveyed over 16,000 people (M_{age} 46.0 years) across 10 different countries and found gender to relate to both experiences of and attitudes towards deepfake pornography; men deemed deepfake pornography behaviours less harmful or less deserving of punishment than surveyed women (Umbach et al., 2024).

New challenges for society as a result of deepfake VAWG

Credibility of information

The societal impact of deepfakes on the credibility of digital (audio-visual) information is backed by a significant body of evidence.[1] This challenge owes to the fact that deepfakes are often hyper-realistic, making it difficult to disentangle real from synthetic images. The mother of one child victim of deepfake VAWG spoke to this challenge "If I didn't know what my daughter looked like naked, I would have thought the pictures were real" (Spiegel International, 2025) An experimental study examined how deepfakes may impact perceived credibility of audio-visual media, finding statistically significant but small effects which suggested that "people may no longer believe that audio-visuals in general represent reality, and people may no longer believe in their subjective ability to discern between real and fake visuals" (Weikmann et al, 2024). On a societal level, this can affect public trust in the credibility of digital (audio-visual) information.

However, alternative theories exist around the implications of deepfakes for public trust in the credibility of information. Viola & Voto (2023) argue that, while deepfake porn might decrease the credibility of audio-visual media, this might not be to society's detriment. By making the public more critical of the information they consume, the scholars suggest that the proliferation of deepfake technology may cause society to place less value on images and videos for information. In turn, they speculate that this could reduce the incentive for perpetrators to make deepfakes in the first place.

Economic cost

[1] This likely owes in part to the disproportionate focus on political deepfakes, and some of this evidence relates to deepfake technology rather than sexual/intimate deepfakes, though many researchers recognise that VAWG content is a substantial majority of all deepfaked videos.

There is very little evidence on the societal costs of deepfake VAWG. In 2021, the European Parliamentary Research Service quantified the cost of cyber-enabled VAWG to lie between €49.0 to €89.3 billion (EPRS, 2021). A later report suggests healthcare, social, policing, and legal costs are incurred by society, as well as productivity losses which damage GDP. Whilst this estimate is relatively robust, it includes all tech-enabled VAWG offences of which deepfakes are just a subset. The cost of deepfakes alone on productivity is thus likely to be substantially lower.

Other economic costs at the societal level exist in:

- Health-related expenses, which have not been estimated (Rigotti et al., 2024)
- Wider effects if sex workers are economically compromised by deepfake technology. This cost has not been estimated (Yavuz, 2025).
- PR and other crisis-management costs to mitigate reputational damage (The Times, 2025).

The cost of policing deepfakes has also not been estimated, though the role of deepfake tools in increasing the volume of tech-enabled VAWG offences is acknowledged (Hornle, 2025).

Freedom of expression

Some scholars discuss how deepfakes require a balancing of the right to freedom of expression and the right to respect for private and family life (Yavuz, 2025). The author suggests deepfakes will have an adverse effect on these rights, imploring lawmakers to consider the impact on the free expression of public bodies like watchdogs, politicians, and individuals.

Perspectives from victimology suggest wider society suffers when individuals are victimised (McGlynn et al., 2024). The individual dignity losses suffered by an individual can cause them to withdraw from online spaces, diminishing their freedom to exist in cyberspace. At the societal level, this can contribute to the polarising effects of social media, broadening the divide between women and men.

IV: Preventing and reducing harm

'The use of digital technology in perpetrating domestic violence creates significant issues for policymakers, justice system officials, employers, and domestic violence scholars and practitioners. Deepfakes can be combatted if they are addressed via legislation, organisational policies, interdisciplinary research, education, and training.' (Lucas, 2022)

The role of legislation

Summary

- The symbolic role of legislation around deepfake VAWG relates to the validation and legitimisation of victim experience, setting the boundaries of acceptable behaviour, and promoting public trust and confidence.
- The practical function of the legislation is somewhat challenged by the speed at which the deepfake landscape evolves.
- Legislation around deepfakes exists, but scholars and organisations have called for legislation on deepfake VAWG to go further.
- Ofcom plays an important role as the UK's communications regulator, and enforces the Online Safety Act (2023) by monitoring compliance, investigating violations, and imposing penalties. However, there are challenges to doing so.

Legislation has the opportunity to set the terms of acceptable behaviour online (including between genders) and holds symbolic and practical value - both of which support efforts to prevent and reduce the risk of deepfake VAWG. Guidance can support legislation by setting out practical steps by which it can be enforced. Particularly for emerging threats, police and Ofcom will benefit from clear guidance about how to enforce legislation around deepfake VAWG.

The symbolic role of legislation

Legislation that defines and illustrates the role of deepfakes and online abuse in the context of wider gender-based violence (Glitch blog, 2024) is symbolic to victims as it recognises their experience. This recognition in legislation can be empowering and prevent feelings of isolation or shame. Legislation and regulation on certain uses of technology can build public trust and confidence. From a survey of 24,673 people living in Great Britain aged 18 or over, 62% stated they would be more comfortable using AI if there were laws and regulations that 'prohibit certain uses of technologies and guide the use of all AI technologies' (Alan Turing Institute, 2023).

The practical role of legislation

The Online Safety Act (2023) made it an offence to share non-consensual intimate images, including deepfakes. The legislation includes a requirement for technology companies and platforms to remove non-consensual intimate image content, and take steps to prevent it from appearing online. Ofcom's codes of practice and guidance on tackling illegal harms issued under the Online Safety Act (2023) compels platforms to improve the testing of their algorithms to make illegal content harder to disseminate (Ofcom, 2025). Ofcom also holds enforcement powers for online platforms.

The Data (Use and Access) Act (2025) includes provisions to supplement the Online Safety Act (2023) by making it illegal to create an intimate image of another person without consent. However, at the time of writing, relevant sections of this Act are not yet in force. The Crime and Policing Bill (2025), currently at its second reading in the House of Lords, proposes to provide police with further powers to tackle intimate image abuse more broadly (and inclusive of deepfakes).

This legislation goes some way to criminalising and providing powers to police (enforcement), Ofcom (regulator), and technology companies to tackle deepfake VAWG. However, scholars and organisations have called for legislation on deepfake VAWG to further include:

- an Online Safety Commission to oversee non-consensual intimate images (NCII) registry (House of Commons, 2025)

- reducing perpetrator's ability to create deepfakes in the first instance (Kira, 2024; (McGlynn et al., 2024)) (House of Commons, 2025)
- protecting those at the greatest risk of harm (Karagianni & Doh, 2024)
- 'culturally intimate' imagery which may not fall under the legal definition of intimate, which predominately includes engaging in a sexual act or nudity (House of Commons, 2025)
- providing clear guidance for enforcement (Sippya et al., 2024).

Legislation and guidance in this space serve distinct roles. While the Online Safety Act (2023) mandates that platforms implement measures to protect users from illegal content, Ofcom has delivered guidance to provide practical advice on complying with these laws. For instance, Ofcom has issued enforcement guidance detailing how it will regulate platforms under the Online Safety Act (2023) (Ofcom, 2024). Ofcom plays an important role as the UK's communications regulator, and enforces the Online Safety Act (2023) by monitoring compliance, investigating violations, and imposing penalties. However, there are challenges to doing so – in particular, it is difficult to detect deepfakes due to the rapid development of technology used to create them. Moreover, platforms hosting deepfakes often operate across borders, and there are challenges associated with Ofcom's ability to ensure adherence beyond the UK's legal reach (Ofcom, 2024). The Independent Pornography Review (Bertin, 2025) found that many pornography platforms are based outside of the UK, with little transparency in how they operate. The Review identified that Ofcom has been unable to enact business disruption measures against non-UK platforms. The Review highlighted that payment providers are "the unofficial regulators" of the industry, and have historically used this power to cut ties with Pornhub after a New York Times investigation revealed a swathe of child sexual abuse material on this site. Despite the fact that deepfakes hosted on these sites are illegal, it remains extremely easy to access deepfake intimate image abuse (House of Commons, 2025), and Ofcom and payment providers have not yet been able to regulate the industry. As of June 2025, the British Board of Film Classification (BBFC) is in discussion with the government over extending its role to include the monitoring of online pornography, as per recommendations of The Independent Pornography Review (The Guardian, 2025).

Other guidance issued to support the regulation of deepfake VAWG includes Crown Prosecution Service (CPS) guidance to support decision-making for prosecutors. This year, the CPS made changes to its guidance to *"stop -perpetrators retaining these images and continuing to take*

gratification from their crimes" via earlier and increased use of deprivation orders, which stop perpetrators from keeping explicit photos of their victims (The Guardian, 2025).

The Online Safety Act (2023), the Data (Use and Access) Act (2025) and the Crime and Policing Bill (2025) have made significant steps to legislate against deepfake VAWG. However, the pace of AI legislation and regulation is viewed as too slow by most in the UK (Ipsos, 2024). Particularly to mandate action from tech companies, who are able to make swift and far-reaching change but "don't have the impetus" to proactively support these efforts, legislation plays an essential role (The Guardian, 2025). It is important to key stakeholders that legislation is responsive, and keeps pace with the next steps of AI-generated VAWG.

Lessons can be learned from how other countries have responded to deepfake VAWG. Landmark legislation was passed in the USA in May 2025 when the "Take it Down Act" was signed into law, prohibiting the publication of non-consensual sexual deepfakes, requiring host platforms to remove the images (The White House, 2025). The Danish government is also taking steps to amend copyright laws so that individuals have the right to their own body, facial features and voice.

Other action to supplement the legislative response to deepfake VAWG includes work by the e-Safety commissioner in Australia, including an image-based abuse removal scheme which enables eSafety to engage directly with platforms to report users and remove harmful content, supporting victims to regain a sense of agency over the situation (Burton et al., 2025). Demand for the service is high (reports to the scheme increased by 960% between 2018 and 2023) and has been implemented with good evidence of success (89.9% of reports in the above time period led to some or all material being removed) (e-Safety). Elsewhere, South Korea's ministry of education set up an emergency taskforce, and its National Police Agency has urged officers to "take the lead in completely eradicating deepfake sex crimes." (CNN, 2025).

Police response

Summary

- There is limited evidence on the policing response to deepfake VAWG. However, what is

known illustrates an often inadequate response.

- Evidence suggests that police are not currently supporting victims of deepfake VAWG adequately. The escalation in VAWG offending from lower-level to more serious offences, recognised by the Angiolini Inquiry, suggests police need to make a concerted effort to focus resources around this emerging threat.
- The police currently have limited guidance on how to respond to deepfake VAWG and could use the VAWG Framework for Delivery to build a response framework. Using this framework would also encourage police to work in partnership with technology companies to tackle deepfake VAWG.

There is limited evidence on the current policing response to deepfakes. However, the evidence that does exist characterises the police response as frequently inadequate, even causing harm to victims. Relatedly, guidance for policing on how to respond to deepfake VAWG is limited, including how police should respond when victims report.

Guidance on policing response

The National Police Chiefs' Council (NPCC) and College of Policing (CoP) have recognised that online VAWG is an emerging threat. This threat is highlighted in the 2021 VAWG National Framework for Delivery: Year 1, the VAWG Strategic Threat Risk Assessment 2023, and the 2024 VAWG National Policing Statement. In particular, the NPCC's VAWG Strategic Threat Risk Assessment 2024 recognises that '*Technology used to facilitate VAWG, specifically, to generate deepfakes and commit image based abuse, is highly likely to be the fastest growing threat in the next 12-24 months*'. In all documents, policing acknowledges their limited capacity and capability to respond to online VAWG and the necessity for clear guidance.

The House of Commons (2025) report identifies that police are receiving improved training on how to respond to no-contact sexual offences in response to the Angiolini Inquiry; this training includes some forms of image-based abuse. However, the report advocates for further improvements and recommends that '*the College of Policing, Ofcom, and the RPH* [Revenge Porn Helpline] *together should produce guidance to improve the police response to reports of NCII abuse*'. The NPCC and the CoP's VAWG National Framework for Delivery provides police forces with a 4P framework

(prepare, prevent, pursue, protect) to support a policing response to tackling VAWG. This framework could be used to develop guidance for policing to drive an improved response to deepfake VAWG.

Responding to victims

There is very little literature on the challenges of policing deepfakes, or deepfake VAWG. However, there is some literature more broadly on the difficulties of tackling online VAWG.

Limited technological capabilities

The 2023 VAWG STRA identifies challenges in the way that police engage with digital evidence and crime; in particular, the STRA found that policing has struggled to obtain the capability and capacity to appropriately deal with digital evidence, which has led to a backlog in digital forensics demand (NPCC, 2023). An unpublished Centre for Policing Women Online (CPWO) report highlights the '*limitations in police digital forensic capabilities and the impact of forensic backlogs on investigations*' (Bakina et al., 2025).

Notwithstanding the difficulties in digital evidence handling in general, policing faces unique challenges to protecting deepfake VAWG victims. Whilst some deepfake imagery exploits the likeness of a real victim, other hyper-real, synthetic images depict the face of someone who does not exist. Particularly with regards to CSAM, the proliferation of deepfake content complicates the task for policing to protect victims by introducing an additional step whereby policing must understand if CSAM is synthetic or real; and if it is synthetic, whether or not it depicts a real child who needs to be protected (BBC, 2025). Deepfake technology thus complicates the policing response to CSAM and other intimate image abuse, making it more difficult to identify and protect victims.

Lack of intelligence sharing between policing and technology sectors

The unpublished CPWO report found that there is no established mechanism for reporting and data sharing between online platforms and policing, which makes it difficult for policing to build an accurate picture of the scale of online VAWG (Bakina et al., 2025). This finding is echoed in the 2024

STRA, which found gaps in data and information sharing between agencies which limits policing's ability to track online offending behaviour (VKPP, 2024).

Lack of national coordination

An Open University and CWPO publication also highlights that there is a lack of a national operating model for online VAWG, meaning that efforts to tackle online VAWG are not consistent across forces, there are disparities in training, and inconsistencies in the degree of strategic prioritisation (Bakina et al., 2025). Police are limited in their ability to take action to protect victims because of the speed at which technology is advancing. It is critical that resources are focused on equipping police with the digital capabilities to tackle deepfake VAWG so that police can support victims who report these offences and prevent retraumatisation by the spread of deepfake VAWG (Knipschild, 2024).

The House of Commons report 'Tackling Non-consensual Intimate Image Abuse' (2025) also highlights barriers relating to the way that policing engages with victims of non-consensual intimate image abuse specifically. These included:

- Victim-blaming
- Retraumatisation when reporting to the police
- Victims/survivors get the sense their content cannot be removed
- Victims being asked to keep their deepfake imagery online during their court case
- Refused access to the Criminal Injuries Compensation Scheme
- Perpetrators are rarely prosecuted or convicted
- Perpetrators being returned devices with the deepfake VAWG remaining.

Current research indicates that victims feel that deepfake VAWG is taken less seriously by police than more traditional forms of VAWG (Harris, in Lucas & Maddocks, 2020). In the wider context of VAWG offending, evidence suggests that 'less serious' VAWG offences should not be neglected by police as this may allow for an escalation in offending, as evidenced by the findings of the Angiolini Inquiry (2024). This escalation in offending has been further documented; for example, evidence suggests a relationship between misogynistic views, widespread dissemination of such views on

social media, and lone-act terrorism (ISD, 2023; McCain Institute, 2021). An improved policing response to deepfake VAWG would build on the findings of the Angiolini Inquiry to recognise the escalation of risk and ensure that all VAWG-related offences are taken seriously to prevent further harm.

Working with technology companies

Technology companies have an important role to play in responding to deepfake VAWG, and research calls for greater accountability of companies that host deepfake VAWG (Glitch, 2024; Equality Now, 2024).

Evidence suggests some technology companies may be responsive to lobbying. A 2025 BBC investigation found 4 apps, available for download on Apple and Google, which encouraged users to make non-consensual sexual/intimate deepfakes but did not advertise these capabilities on their website or App Store pages. Alerted about their usage for non-consensual sexual/intimate deepfake creation by the BBC, Apple (but not Google) removed them from the App Store until their practices changed, going further and revoking developer rights for one app. Action against complicit technology companies can be successful, but investigative journalism alone lacks enforcement and regulatory power. Lessons from Australia's eSafety image-based abuse removal scheme demonstrate the efficacy of work partnered with tech companies: where image removal was unsuccessful, the scheme took steps to limit their discoverability by de-indexing results on search engines, most commonly (for 92.4% of cases), on Google (Burton et al., 2025). Government and policing in England and Wales have the opportunity to support such mechanisms of accountability.

There is also an opportunity for the Government to work with technology companies to develop effective safeguarding measures, reporting templates, and mechanisms to tackle deepfake VAWG (Equality Now, 2024; Thorn & Burson, 2025). Such 'safety by design' measures may include, for example, robust human moderation that reviews each individual piece of content uploaded onto a website before it is uploaded (Bertin, 2025). However, due to advancements in technology, the limitations of such measures are acknowledged in the literature emphasising the importance of a holistic approach including education and awareness raising (Jacobsen & Simpson, 2023; Umbach et al., 2024). The Home Office, NPCC, and the CoP are well-placed to work with technology

companies in this area. The VAWG National Framework for Delivery can assist as it encourages partnership working to tackle VAWG which can be applied to deepfake VAWG.

Education and awareness

Summary

- There appears to be a consensus among scholars that improving public awareness of deepfake VAWG will enable better identification of deepfake VAWG.
- However, it is recognised that the ability to identify deepfakes does necessarily translate to reporting. There is a call for more campaigns to raise awareness of how to spot, report, and seek support for deepfake VAWG.
- Children and young people should be a particular focus for education and awareness campaigns on deepfakes. There is an opportunity for early prevention and intervention programmes to readjust cultural norms and reject the notion of deepfake VAWG by understanding the harm it causes and building a trusting environment to report

There has been a unified call from academics, subject matter experts and specialist organisations for more education and awareness of deepfake VAWG, for the wider public and among children and young people in particular.

The public

The 'public awareness and attitudes' section of this report illustrates the limited knowledge that the general public has of deepfakes, including deepfake VAWG. There appears to be a consensus among scholars that improving public awareness of deepfake VAWG will enable better identification of deepfake VAWG. Baroness Bertin's Independent Review of Pornography identified '*significant gaps in public knowledge around Intimate Image Abuse (IIA) and deepfake IIA, as well as how to report this on platforms and where to seek support*'. She called for '*an appropriate public awareness campaign [...] to raise awareness of how to spot, report, and seek support for IIA*', to be funded by platforms that host pornography (Bertin, 2025, p.21).

However, there is mixed evidence on the benefits of an improved ability to identify deepfake content. Some literature calls for the public to improve digital literacy and be educated on how to identify a deepfake so they can '*distinguish truth from deception*' (Mustak et al., 2023). The thought is that this might discourage deepfake VAWG consumption (Umbach et al., 2024) and reduce harm by highlighting the artificial nature of the media (Viola & Voto, 2023), or encouraging the public to flag the content as deepfake and have it removed. However, other scholars argue that being able to identify an image or video as a deepfake has a minimal impact on the social and psychological effects of deepfake VAWG. Clark & Lewandowsky's (2024) work demonstrates that the harmful influence of deepfakes on people's beliefs and behaviours remains the same regardless of the authenticity of the content.

Literature suggests that public education could go some way to prevent future perpetration by informing people of the detrimental impact that deepfake VAWG can have and that it is non-consensual abuse (Henry & Beard, 2024; Umbach et. al., 2024). Furthermore, Umbach et al. (2024) suggest that, to prevent and reduce further harm from deepfake VAWG, the public needs to be educated that it is illegal and must be reported to the police.

Children and young people

Our rapid evidence review found widespread agreement across the literature that to increase education and awareness of deepfake VAWG there should be early prevention and intervention programmes targeted at children and young people. News reports increasingly demonstrate that children and young people are particularly vulnerable to this kind of victimisation, which happens frequently in schools and is reported on a global scale (The Times, 2025; Spiegel International, 2025). Identified as "the next big sexual violence issue that is going to impact schools" (The Independent, 2025), schools in particular need clear guidance on how to deal with deepfake VAWG. Scholars have also identified the need for early awareness raising and education among children and young people, suggesting this can help to '*dismantle harmful misconceptions and foster greater accountability*' (Thorn & Burson, 2025). This can provide children and young people with the tools to develop healthy and respectful behaviours and relationships (Henry & Beard, 2024). Through these tools, and early educational starting points, children and young people are given an opportunity to

readjust cultural norms and reject the notion of deepfake VAWG thus going someway to prevent and reduce further harms.

One key challenge for deepfakes in education sectors is sextortion, a form of blackmail where a perpetrator threatens to share nude or semi-nude images of an individual unless the individual pays them money. Sextortion is increasingly impacting children and young adults, and these images are sometimes deepfakes (NCA, 2024). The Internet Watchdog Foundation (2024) evidenced that in 2023, 91% of the reports they received of sextortion were of boys being targeted. In 2024, the National Crime Agency (NCA) warned education professionals of the global increase in sextortion and identified the largest proportion of cases are boys aged 14-18 (NCA, 2024). Creating an environment that promotes open discussion in a trusted setting can remove the power that can often be held by and associated with deepfake VAWG and can encourage children and young people to ask questions and seek help.

Specialist support services

Summary

- There are a number of specialist support services to support victims of deepfake VAWG, but more are needed
- The complexities of being a victim of deepfake VAWG mean that services require specialist knowledge and expertise to support victims to reach desired outcomes
- Specialist support for victims who may face additional barriers to reporting or accessing support, including from marginalised groups, is critical

Specialist support services provide tailored support for specific needs. Specialist support which is underpinned by expertise and knowledge provides victims with effective support. There are several UK-based organisations that offer specialist support to victims of deepfake VAWG. Organisations such as Revenge Porn Helpline, Childline, and National Ugly Mugs offer specialist support through caseworkers which can include supporting the victims to report to the police. The former two also

support victims with content removal. Revenge Porn Helpline (2023) successfully took down 90% of the 12,921 images which were reported by victims/survivors in 2023.

However, there is a need for more specialist deepfake VAWG support services (Glitch, 2024). Services should be victim-led and appropriately resourced (Glitch, 2024), trauma-informed (Lucas, 2022), and understand what remediation approaches are sought by victims (Umbach et al., 2024). As evidenced above, some marginalised groups are disproportionately impacted by deepfake VAWG. Specialist support services should proactively consider how to make themselves accessible to these groups. For example, providing free services, having staff members who reflect affected groups, and adopting a trauma-informed approach.

An assessment of studies in England between 2013 - 2023 which focused on image-based sexual abuse perpetrator behaviour identified the need for specialist support services due to the overlap between victims and perpetrators (Henry & Beard, 2024). Therefore, specialist support services could also prevent further harms.

Centring victims' experience

Summary

- Victims' first-hand experience is critical to understanding the impact of deepfake VAWG
- There is lots of good practice in centring victims' experience in the wider VAWG space, including policy and support services. There is an opportunity to replicate this approach when refining the response to deepfake VAWG.

Technological advances mean that the nature and impact of deepfake VAWG is constantly evolving. The first-hand experience of victims is central to understanding this impact, and desired outcomes to drive an improved response from agencies and service providers.

There has been a concerted effort to create space for victims to lead the conversation around VAWG, influence policy decisions, and inform and run specialist support services. For example, the Modern Slavery Policy and Evidence Centre (MSPEC) has established a Lived Experience Advisory Panel

(LEAP) which is victim-led and ensures MSPEC's research is informed and co-produced by victims. The guidance documents created by domestic abuse charities SafeLives and Women's Aid (2024) highlight best practices for working with victims and reflect the broader shift toward victim-led approaches in the VAWG space. Victim-led work can be very effective and beneficial but requires the victim, and those making a request to hear from victims, to acknowledge their positionality and consider the potential personal impact to the victim due to their closeness to the subject matter (SafeLives; Women's Aid, 2024).

The deepfake VAWG space needs to keep pace with other VAWG-related work to ensure that policies, education and prevention initiatives are informed by or led by victims to offer '*a more complete understanding of deepfake and digitally altered imagery abuse and what types of prevention measures are possible*' (Flynn et al., 2021.).

Conclusion

'*The needs of victims, the motivations of content creators, the role of viewers, and the responsibility of porn sites are all issues that require public and scholarly attention. It is only by understanding these perspectives that deep fakes can be addressed in all their emergent forms*' (Maddocks, 2020).

This evidence review has informed Crest's lines of enquiry for primary research, and the development of research tools including discussion guides and survey scripts. Where relevant, evidence from this review is included in our research report and triangulated with primary research data to build the evidence base on the impact of deepfake VAWG.