Deepfake pornography and its impact on sexual consent

A qualitative mix-method study on social perceptions

Student Name: Xenia Liebscher
Student Number: 539124

Supervisor: Dr Marlen Komorowski

Master Media Studies - Media & Creative Industries
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

Master's Thesis
June 2023.

Word Count: 19 070

DEEPFAKE PORNOGRAPHY AND ITS IMPACT ON SEXUAL CONSENT: A QUALITATIVE MIXED-METHOD STUDY ON SOCIAL PERCEPTION

ABSTRACT

Deepfakes are a new form of deep learning algorithm that is used to swap one person's face with another. This technology has been found in politics, media manipulation and 'fake news', but the most prominent use is in pornographic videos, with 96% of deepfakes being pornographic. When used in pornographic videos, deepfakes are most commonly done without the prior knowledge of the individuals depicted in the video, becoming a form of technology facilitated sexual-violence (TFSV). This lack of consent when creating and sharing deepfakes online is a complex issue on an individual and social level. This is due to the fact that consent functions differently online, and the ability to create fake content re-structures understanding of consent again. Furthermore, questions of ownership of content and legal aspects also play a vital role in deepfakes and consent. However, even considering the harm that pornographic deepfakes can cause, current research focuses mainly on the political and legal aspects of deepfakes. Therefore considering these premises, this research investigates the following research question: "What are the implications of deepfake pornography on perceptions of consent online?" The analysis is based on a qualitative mix-method approach using focus groups and expert interviews to find key social perceptions on deepfake pornography and consent. The main findings show that social anxieties around consent practices online increase when considering deepfakes. This is mainly due to the realisation that non-consensual deepfakes do not require any previous interaction between victim and perpetrator and, therefore, can expedite the TFSV happening online. In addition, participants noted that deepfakes are a form of TFSV that is difficult to remove from the online sphere due to digital infinity and the ownership dispute deepfakes can create. Deepfakes can be considered a form of content creation, similar to memes, and this can complicate the ownership of the video, impacting consent practices further. Lastly, another key aspect that was identified is the social and individual impacts deepfakes can create. On a social level, they perpetuate the confusion between truth and fiction, blurring the lines of consent. Whereas on an individual level, they can impact the victim in a similar psychological and physical way compared to an offline victim of sexual violence. Accordingly, the main results highlight that pornographic deepfakes do not only intensify non-consensual practices online, but they can have vast social and individual implications.

KEYWORDS: *Pornography, Deepfakes, Consent, Technology-facilitated sexual violence, Digital autonomy*

ACKNOWLEDGEMENTS

Six months ago, I would have never thought that I would get the opportunity to write and learn about such an interesting topic and find my passion in it. I had the pleasure of working with Dr Marlen Komorowski, who not only supported me but encouraged me to delve into this topic. I also owe great gratitude to my participants and Experts who gave me their valuable time and thoughts on the matter. Additionally, I want to thank my family, who always helped me see the bigger picture even when I could not see it myself. Last but certainly not least, I could not have undertaken this journey without the encouragement, laughter and love of my friends - you truly made this an unforgettable experience. I especially owe Nina my deepest gratitude, as she not only supported me emotionally but also helped shift my own thoughts into perspective with her valuable feedback.

Table of Contents

Abstract and keywords

[Preface]

1. Introduction 1

1.1 The Internet and the changing scene of Pornography 1

1.2 Deepfake pornography and consent 3

1.3 Scientific Relevance 4

1.4 Social Relevance 5

1.5 Chapter Overview 5

2. Theoretical Framework: Contextualising Deepfake Pornography and Consent Online 7

2.1 Contextualising Deepfake Pornography 7

2.1.1 Defining Pornography 7

2.1.2 The Digitalisation of Pornography 9

2.1.3 Deepfake Pornography 10

2.2 Consent Online 15

2.2.1 Placing consent in the digital space 15

2.2.2 Consent and digital (sexual) autonomy 16

2.2.3 Non-consensual sexual interaction online 19

2.3 Moving Towards a (new) integrated framework 21

3. Methods 23

3.1 Qualitative Mixed-Method Approach 23

3.2 Focus Groups 23

3.3 Expert Interviews 24

3.4 Sampling and Participant Demographic 24

3. 5 Data Collection and Operationalisation 27

3. 6 Data Analysis 29

3.7 Limitations and Ethics 30

4. Results 32

4.1 Digital aspects of (non)consensual deepfake pornography 33

4.1.1 Consent and the ability to create fake content 33

4.1.2 Digital infinity and deepfake consent....36
4.2 Social aspects of (non)consensual deepfake pornography....37
4.2.1 Societal awareness of consent online and deepfakes....37
4.2.2 Social Stigma and Deepfake Impacts....40
4.3 Ethical aspects of (non)consensual deepfake pornography....41
4.3.1 Ethical conflicts regarding deepfake pornography....42
4.3.2 Deepfakes compared to other nonconsensual sharing online....44
4.4 Legal aspects of (non)consensual deepfake pornography....47
4.4.1 Autonomy of content and deepfake consent....47
4.4.2 Regulation of deepfakes and consent....50
4.5 Connecting the themes....51
5. Conclusion....54
5.1 Deepfake impact on Consent....54
5.2 Theoretical and Social Implications....56
5.3 Limitations and Recommendations for future research....57
References....59
Appendix A - Consent form....66
Appendix B - Topic Guide....69
Appendix C - Coding tree....71

1. Introduction

1.1 The Internet and the changing scene of Pornography

In 2017 famous actress Gal Gadot found herself online in a pornographic video. However, interestingly enough, she had never filmed, seen, or distributed this film before. Instead, she became one of the first individuals who fell victim to a form of deep learning algorithm used to predominantly create fake sex videos of women. This technological development within the pornography industry emerged due to the rise of internet pornography which allows easy access, anonymity, and affordability for the consumer (Lim, Carotte & Hellard, 2016). The pornography industry is estimated to account for over 30% of world wide web traffic (Anthony, 2012) and has the highest viewing rates of any media sector. Moreover, digital devices, which are now intrinsically interconnected within our world, have allowed this form of pornography to become even more accessible through digital platforms such as social media (Paasonen, 2016).

Therefore, it is no surprise that these devices and digital platforms have become a reflection of the society in which we live (Powell & Henry, 2017). Leading to an increase in the pornification of content in which sexually explicit images have become common in popular culture (Drenten, Gurrieri & Tyler, 2019). Furthermore, through the widespread use of smartphones with cameras and societal participation in online platforms, the digital landscape has provided "a ready platform for the perpetration of sexual harassment, abuse and violence" (Powell & Henry, 2017, p. 2). Hence, it may be unsurprising that digital appliances can also expedite sexual abuse online (Powell & Henry, 2017).

This can be seen reflected in how technological development has led to changing forms of sexual interactions online, from things such as 'sexting' to forms of image-based sexual abuse such as 'sextortion' or 'non-consensual sharing of images' (Powell et al., 2019). The way sex and private images are used online has been reshaped completely, as consent practices are more overlooked and complex than ever (Marcotte & Hille, 2021). Research has shown that over 20% of users have no issue sharing content without the prior consent of the individual (Garcia-Gomez, 2017), a percentage that is likely to increase in the coming years.

This increased access to pornographic content and sexually explicit images has raised many issues and concerns in psychology, laws, and media studies. Main concerns surround themselves around the negative impacts on health and well-being this may create. These concerns include that the depiction of "violence against women lead to increased violence against women in real life" (Lim, Carrotte & Hellard, 2016, p. 3). Other anxieties are based on the worries of pornography addiction or that individuals view pornography as 'real' rather than fiction (Attwood, 2005). Often overlooked in academic research, this last concern has exponentially worsened through Artificial Intelligence (AI) technology such as deepfakes.

Deepfakes are one of the latest developments in AI, having come to the general public's attention in 2018, and are a form of 'deceptive media' (Paris & Donovan, 2019). They specifically encompass the production of hyper-real videos in which a form of deep learning algorithm analyses an individual's facial features and superimposes these over another individual (Öhman, 2019). Through the data collected, and the 'learned' angles and expressions, the algorithm can reproduce the first person's face in a way that it mimics the second individual's expression. Thus, similar to the human brain, deep learning algorithms learn from the information it gathers and generates their own combination of that input (Öhman, 2019). With this in mind, due to the fact that deep learning technology learns rather than takes data, this process does not involve any infringement of private or unauthorised access to personal information (Harris, 2019). Instead, this can be achieved by utilising publicly accessible images or videos.

Various open-source software already exists, such as DeepFaceLab and Faceswap, as well as other applications like ReFace, Zao, Deepfakes web β, and Wombo, among others. These platforms offer user-friendly interfaces for which no expertise is needed, that aid the manipulation of images or videos. Hence, any individual with these applications has the ability to create and disseminate deepfakes.

This form of technology has been found within politics, media manipulation and 'fake news' but most prominently are found within pornography. When used as a form of pornography, they place into the larger scope of image-based abuse through technology, also known as technology facilitated sexual-violence (TFSV) (Powell & Henry, 2017; Powell, Flynn, & Suguira, 2021). This is because these deepfake videos or images are mainly done without the consent or prior knowledge of the person being shared. In fact, according to Deeptrace Labs, which did a study on deepfake detection, "96% of

deepfake videos are non-consensual porn, mostly of women celebrities" (Van der Nagel, 2020, p. 424).

Therefore, this form of AI learning can pose a significant concern as they enable the creation of videos that can be highly persuasive and lead the viewer to believe that genuine recordings are being shown. Meaning this technology has the potential to generate extremely convincing pornographic content without the consent of the individuals involved (Kirchengast, 2020), creating potential social, financial, and psychological harm to the individual implicated.

1.2 Deepfake pornography and consent

Considering these changes in the pornography landscape online, concerns about consent practices being implicated arise (Van der Nagel, 2020). The combination of accessible technology and the removed need for sexually explicit images of the targeted individual allow for a potential increase in TFSV (Powell & Henry, 2017). This opens up a whole new form of sexual violence online, in which consent is implicated in various ways, exacerbating the creation of forms of revenge porn (Kirchengast, 2020).

Moreover, due to the infinite nature of the internet, consent practices online are already met with a new form of spatial and temporal framework (Marcotte & Hille, 2021) in which sexual interactions can happen from anywhere and continue past the act of giving (non)consent. This new framework for consent interaction is then also further impacted by the ability to create fake content that may be difficult to claim ownership over (Meskys et al., 2021). Deepfakes can be considered a form of content creation, as the person who uses the various images to teach the deep learning algorithm creates new content, similar to how memes are made. With this, the ownership of that created media arguably belongs to the creator. However, the individuals who are shown in that deepfake may not have given consent and potentially want to remove it, but as they are not the legal owners of the deepfake, this process becomes more difficult. Therefore, the legal problem of deepfake ownership can additionally complicate how consent is viewed and performed. The question of ownership and creating (fake) content also impacts how digital autonomy works. This is because the practice of using someone's identity and potentially owning that content removes the ability to have autonomy online.

Furthermore, the increased normalisation of pornographic content through the pornification of non-sexual images or videos (Goldfarb, 2015), has also led to an environment in which pornographic deepfakes can thrive under the pretence of being authentic sexual content. This is another key issue attributed to consent and deepfakes, the inability to discern true content from fake content (Paris & Donovan, 2019). This could not only implicate the individual in the video but can also have larger ramifications for the societal understanding of consent.

Therefore, considering these worrisome developments of consent online, this paper aims to combine these numerous problems and research the following Research Question: *What are the implications of deepfake pornography on perceptions of consent online?* In order to successfully answer the research question, three sub-questions are discussed in this study:

1) What factors impact how deepfake pornography and consent online are perceived?

2) To what extent does deepfake pornography impact the understanding of digital autonomy? And how does this impact consent online?

3) How does the context of deepfake pornography impact ethical understandings of consent online?

This research paper will analyse the responses from various focus groups and experts working in the fields of sexual violence and/or deepfakes to answer the questions proposed. In addition, this selection of participants will allow for the comparison of various levels of expertise on the topic of sexual violence and/or deepfakes, which will aid an extensive covering of deepfakes and consent.

1.3 Scientific Relevance

Academically, this research can add substantial insight into the phenomenon of deepfake pornography, especially in regard to its negative impact on consent perception online. This research will aim to close the literary gap in various aspects, considering that most deepfake studies place their focus on political propaganda and the threat of fake news (Gamage et al., 2020; Gosse & Burkel, 2020) instead of pornographic deepfakes. Additionally, the current body of academic literature concentrating on

deepfake pornography is quantitatively driven by “studies focusing on detection and regulation” (Godulla et al., 2021, p. 74). Therefore, through a qualitative approach, this research aims to add academic value regarding societal reactions and attitudes (Gamage et al., 2020; Godulla et al., 2021).

Furthermore, the aspect of consent regarding deepfake pornography has yet to be thoroughly researched, as current literature fails to acknowledge the issue extensively. Especially as a form of TFSV that can harm both individuals and society, deepfakes and non-consent are overlooked. Most prominent research on consent understandings online in regard to deepfake pornography has been in legal and philosophical terms (Öhman, 2019; Öhman, 2022), in which questions of identity recognition and copyright infringement were the main aspects researched. Hence, pornographic deepfakes and consent as social issues have yet to be addressed.

Therefore, by having a qualitative approach that investigates societal perceptions of deepfake pornography and its impact on consent, this research aims to shed light on how the changing digital environment can have severe societal implications.

1.4 Social Relevance

In terms of social relevance, this paper can aid social understanding of deepfakes and consent practices online. With the rise of sexual violence online and recent social movements such as #metoo, it is important to continue identifying violent practices in which consent is disregarded. The #metoo social movement managed to shed light on experiences of sexual abuse, harassment and what has been coined the ‘rape culture’: a term defined as a society or environment whose attitudes normalise or trivialise sexual assault and abuse (Rape Culture, n.d, Oxford Dictionary). The pornification of content and the use of technology such as deepfakes are key actors that reinforce this rape culture, as deepfakes tend to continue the trend of exploiting the sexual autonomy and privacy of women (Van der Nagel, 2020). Hence, there is a social need to understand and create awareness surrounding this technology and its primary use. Furthermore, through this awareness and the identification of the issues with privacy and data for deepfakes, this research can help create a foundation for legislators or policymakers to enforce ethical and safe uses of deepfakes, with the hope of reducing the victims created by the dissemination of deepfake pornography.

1.5 Chapter Overview

This section gives an overview of the thesis structure and contents of the chapter to adequately cover the topic of the study. The second chapter consists of the theoretical framework of the paper, which is an extensive review of the existing literature on deepfakes and online consent. This framework was used to guide the data collection and the analysis of the results. In the next chapter, a thorough description of the method applied is presented, outlining the qualitative mixed-method approach using focus groups and expert interviews, as well as providing a description of the participant demographic, data collection process and the data analysis methodology. The results are presented and discussed within the fourth chapter, using theory to highlight key findings. Lastly, the conclusion offers a critical assessment by placing this paper's findings with the results of previous academic debates regarding deepfake pornography. This section will further highlight the social and scientific implications while proposing potential areas for future research.

2. Theoretical Framework: Contextualising Deepfake Pornography and Consent Online

To effectively address the theoretical aspects of the research question, this chapter has been divided into two sections. These sections provide the basis for the framework that will be used to drive the research. The first section places the use of pornography into a digital context and thus allows for the contextualisation and conceptualisation of deepfake pornography. The second section defines consent in a digital sphere, highlighting key differences to offline consent that underline the complexities of deepfake pornography.

2.1 Contextualising Deepfake Pornography

2.1.1 Defining Pornography

For this research, the concept of pornography needs to be defined in order to place it in the larger context of the paper. Previous research has shown that how pornography is defined and contextualised may impact how the research is approached (Fisher & Kohut, 2020) and must be addressed. The difficulty in defining pornography lies mainly in the ambiguous divide between explicit images and pornography and what qualifies as what. As stated by Hawkins and Zimring (1988), “Even if all know it when they see it [...], how do we know that the agreement is not totally illusionary?” (p. 20). This illusionary idea is also reflected in the existing operational definitions of pornography. The current definitions range from ambiguous descriptions that use terms such as ‘sexually explicit’ (Fisher & Barak, 2001) to more concrete descriptions that define pornography as a concept that involves the “depiction of nudity, genitalia, or sexual activity” (Fisher et al., 2019, p. 200). Nevertheless, even with more conclusive definitions, no decisive framework indicates which type of content can be considered pornographic. Some suggest that any content that implies sexual behaviour can be considered pornographic (Flood, 2009). In contrast, others disagree and state that pornography involves unobstructed representation of genitalia or sexual behaviour (Hald & Malamuth, 2008).

Furthermore, the discussion on the correct definition also involves disagreements on where depictions of nudity are placed within pornography. It raises questions on whether nudity can only be considered pornographic when the image

exhibits sexual behaviour or if any image portraying nudity can be distinguished as porn (Lo et al., 1999). Therefore, in the existing research, one can find both those definitions in which images of nude individuals are excluded (Hald & Malamuth, 2008) and those in which these are explicitly included in the definition (Lo et al., 1999). In addition to the questions of nudity and sexual content, many definitions of pornography also commonly describe the explicit intention or the use of such materials. These definitions underline the sexual arousal or gratification derived from pornographic content but also remark on secondary functions that include "oppression, offence, marketing for profit, and artistic expression" (Fisher et al., 2019, p. 200). However, these definitional elements tend to be included as a secondary aspect of the definition of sexual content, as they are unsuited for scientific research due to their subjective component (Fisher et al., 2019).

Lastly, many feminist theories define pornography as sexual material that harms women (Lindgren, 1993), by which they want to distinguish it from obscenity. Obscenity, within this definition, is defined as sexual material that lacks value, whereas pornographic content is defined to degrade women. This definition builds upon the model proposed by MacKinnon (1983). Which was defined as a model with three key components: "graphic sexual explicitness, the subordination of women, and depictions of any one of a long list of specific sexual acts." (Lindgren, 1993, p. 1157). This understanding of pornography allows for a more specific definition, in which the subordination aspect is important to note. This is because the subordination requirement is very often found with pornographic content, as mentioned in the text by Lindgren (1993).

With these differing definitions in mind, and the recommendation by Fisher et al. (2019) to adopt a definition of pornography that includes components of nudity and sexual behaviour, this research will build upon the definition proposed by Dworkin and MacKinnon (1983). It will define pornography as graphic sexual explicitness, including nudity and explicit sexual behaviour, which can have elements of degradation against women. This last component, although potentially limiting, is necessary to analyse the concepts this research aims to investigate as it looks at consent within deepfake pornography, and previous research has identified that deepfakes are often used as a way to degrade women (Powell & Henry, 2017).

2.1.2 The Digitalisation of Pornography

Through the accessibility of digital devices and the internet, the digitalisation of pornographic content has evolved drastically over the past few decades, with around 68% of digital media users stating that they have been involved in sexual interaction in a digital context, including pornography (Döring & Mohseni, 2018). Previous researchers agree that the internet is a powerful tool that has "fundamentally influenced the production of knowledge, distribution of tasks, and activities society engages, shaping the perspectives of generations of individuals." (Makin & Morczek, 2015, p. 1). This is reiterated by the technological determinism theory, which suggests that technology is the primary driver for social change (McLuhan, 1964). This concept has been used as a lens in various digital pornography studies that argue how pornography has evolved within the digital sphere, and what ways the technologies found within pornography may have unforeseen social and cultural consequences (Elmer-Dewitt, 1995**;** Gernert, 2010). What these studies suggest is that technological advances are the main driving force for changes in how society views, consumes and interacts with the pornography industry.

However, the present study rejects this simplistic view on how technology is the only driving force for change, as this would disregard social and cultural influences (Bakker & Taalas, 2007) and how explicit content is potentially used to harm or disempower certain members of society (Marcotte & Hille, 2021).

These social and cultural influences are important to consider when discussing how pornography can be related to digitalisation, especially when it comes to the pornification of media content. Pornification is a concept coined to conceptualise the overall use of pornographic images or aesthetics in popular culture (Goldfarb, 2015). The study by Goldfarb (2015) identified a trend in marketing strategies increasingly using sexually suggestive content, contributing to the overall amount of sexually explicit content online. This trend of pornification has been expedited by digital platforms, where sexually explicit images are common ground and available in more varied forms and shapes than ever (Paasonen, 2016). Drenten, Gurrieri and Tyler (2019) reiterate this increase in sexual content online through their study, which analysed the mainstreaming of pornography on non-pornographic sites and networks, such as Instagram and Facebook. They noted that through the rise of digital technologies and social media platforms, pornography has become more normalised and even part of

many influencers' monetisation culture, becoming "a central node in the culture, politics and economics of digital technology" (Saunders, 2019, p. 4). This normalisation is mainly done by distributing pictures and videos that are overtly sexual without explicitly performing sexual acts. The normalisation of pornographic content through digitalisation is important to this research as it has aided the desensitisation of such images within the digital culture (Paasonen, 2016).

Furthermore, individuals do not only consume (more) pornographic content, but they are also found to produce and share their own pornographic material online, fitting the broader trend of “collaborative creation of content facilitated by the internet” (Patella-Rey, 2021, p. 92). Together with desensitisation, this has even led to occurrences where non-pornographic images that exist within the online sphere are edited or transformed by individuals into a pornographic context (Goldfarb, 2015). These trends have led to the propagation of user-generated pornography and have aided the creation of new forms of control and power within the industry, as individuals, not only companies, are now also involved in establishing gender dynamics (Gosse & Burkel, 2020). However, these emerging forms of power and control reinforce existing gender biases in which women are viewed as objects ‘to be looked at’ rather than subjects who look (Mulvey, 1975).

As found by Gosse and Burkell (2020), “women’s images are understood as consumable, malleable, and brought into being for the enjoyment and gratification of men” (p. 497). This perspective is especially prominent within the pornography industry, where research has identified that pornography is used to degrade women (Lindgren, 1993). This means that the increase of pornography in digital spaces aids the larger tendency to strip "women's agency and privacy by using images - real or fabricated - to humiliate and disempower them" (Gosse & Burkel, 2020, p. 498). Thus, a concerning trend is identified by the literature in which women are targeted through these pornographic practices and suffer the real-life consequences of them.

2.1.3 Deepfake Pornography

This aspect of humiliating women through images (Gosse and Burkel, 2020) and the mainstreaming of digital pornography (Goldfarb, 2015) have brought about a new form of disempowering women through Artificial Intelligence (AI) called deepfakes. Deepfakes are a newer phenomenon that can be placed into a history of 'audio-visual

manipulation' (Paris & Donovan, 2019), in which content is modified to create new content. When used in non-consensual terms, they can be considered harmful to society. This harmful form of digital pornography can be examined through two theoretical lenses. One addresses ‘media effects’, and the other aims to discuss ‘sexuality and power’. Both frameworks are vital to researching deepfake pornography, as each addresses critical concerns.

The concept of ‘media effects’ was first coined by Bandura (2001) and refers to how media content can influence the attitudes, behaviours, and beliefs of an individual. Within the context of deepfake pornography, the framework suggests that perceptions of consent, agency and privacy may be implicated by exposure to pornographic content generated by AI technology (Van der Nagel, 2020). Here the main concern is that the creation and existence of deepfake porn will further blur the lines of consent online and will target the identity and likeness of women by creating fake content of them (Gosse & Burkell, 2020). This would infringe on the digital agency and privacy of individuals, as well as make it harder for society to distinguish reality from fiction (Gamage et al., 2022).

To understand this implication on truth and fiction, it is best to link it to the post-structuralist concept of hyperrealism (Baudrillard, 1981). This term is best defined as “the generation by models of a real without origin or reality” (Baudrillard, 1981, p. 3). A definition that could just as easily be used to define the creation of deepfakes as deepfakes can be said to be the act of creating something new with no ties to reality. Within this concept, it is argued that reality, or notions of reality, have evolved so that there is now a cultural state of disorientation between the signs and symbols created to understand reality. Baudrillard (1981) believes that due to the perceptions created by culture and media, what is understood as fiction and what is considered real are starting to blend. For example, the use of digital devices actively changes how individuals view the world, as they now have access to various perspectives and angles of events that blend understandings of reality with the digital context. Deepfake pornography can accelerate this effect, as it blurs the ideas of reality and fiction even more. It has been argued that videos hold a ‘claim to truth’ (Paris & Donovan, 2019), which is now questioned by the creation of deepfakes.

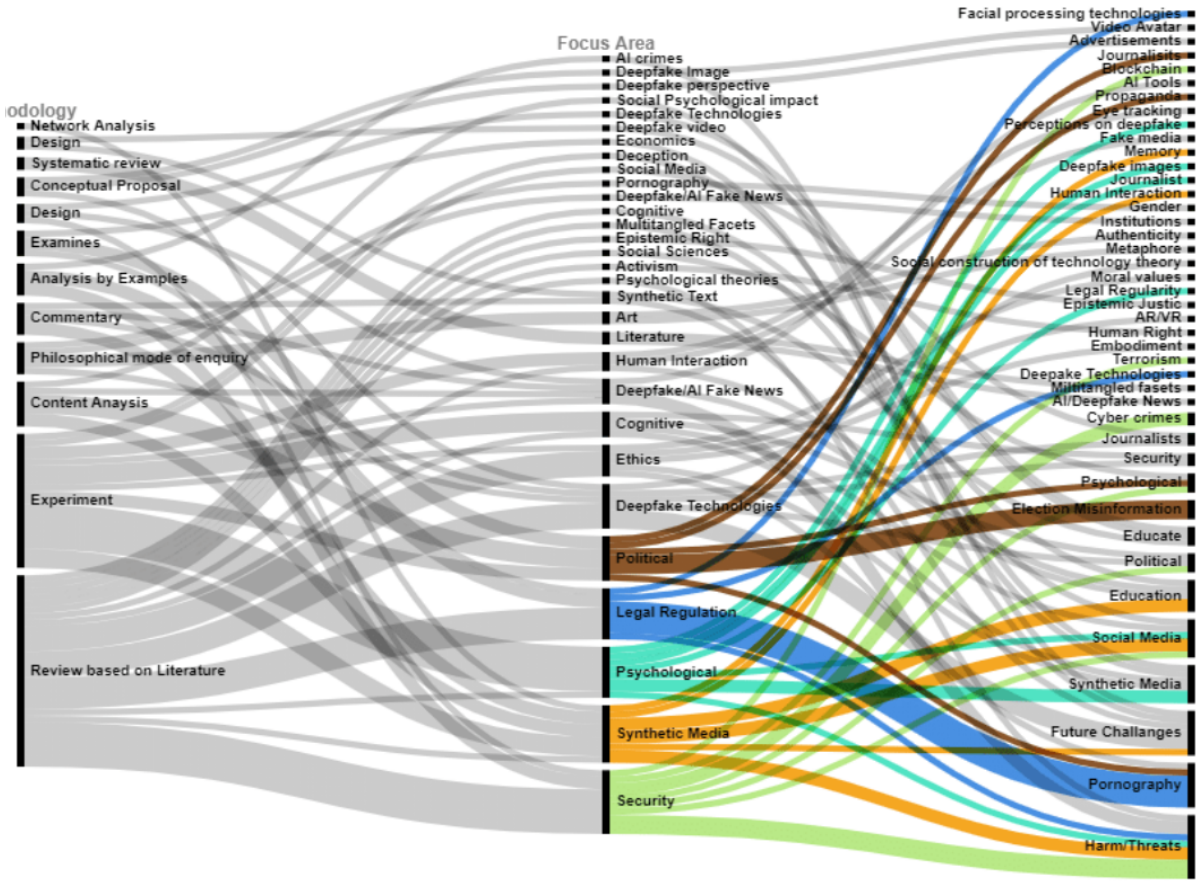
The second significant framework to study deepfake pornography is the framework of ‘sexuality and power’ (Foucault, 1998). This framework emphasises the

link between sexuality and power dynamics within society, such as gender, race, and class, stating that these dynamics impact how sexuality is experienced. Applying this framework to deepfake pornography, it has been argued that the use of deepfake technology underlines existing power imbalances. One way that deepfakes do this is by underlining the gender power imbalance in society, as deepfakes are often used to objective or commodify women's bodies through the distribution of fake sexual content (Wagner & Blewer, 2019). Van der Nagel (2020) further argues that "by treating women's faces as a digital resource to be edited onto sexual bodies by artificial intelligence, [deepfakes] reinforces the idea that women exist as sexual objects" (p. 426). This use of deepfakes accentuates existing gender power dynamics in the online sphere, as previous research has shown a significant gender disparity in who is implicated in deepfakes and who is not (Gosse & Burkell, 2020), with 96% of deepfakes being pornographic content of women.

Still, even considering all of the above aspects, deepfakes are most commonly researched through a political or legal lens, disregarding user or societal perceptions of the technology (Gamage et al., 2020). Therefore, even with the predominant use of deepfakes being within pornography, the academic body of literature does not reflect this (Gosse & Burkell, 2020; Godulla et al., 2021). As found in the text by Gamage et al. (2020) the five key research areas for deepfakes are security, synthetic media, psychological impacts, legal regulation and politics (see Figure 2.1).

Figure 2.1

Categorisation of research focus and methods of deepfake papers



Note: Figure taken from the text: "The Emergence of Deepfakes and its Societal Implications: A Systematic Review" by Gamage et al. (2020).

Furthermore, Gamage et al. (2020) also found that the body of literature that does focus on deepfake pornography is found to lack research on the social implications. Hence, to address the current state of the research field of deepfake pornography more clearly, the following table gives an overview to show in what ways deepfakes have been researched within the pornography field, highlighting the research gaps identified.

Table 2.1

Literature study on deepfake pornography research

Literature/ Approach	Focus in Research	Method	Key findings
Wagner & Blewer (2019)	Exploration on how the rise and dissemination of deepfake pornography aid the gender-based inequalities in visual media	Critical Literary review	1) Notes deepfake understanding as a gendered issue 2) Argues for societal need to learn media literacy
Öhman (2019)	The perverts dilemma: questioning ethics between sexual fantasies and deepfake creation	Philosophical mode of enquiry	1) Ethical understanding of deepfake use needs to be placed within the larger context of social dynamics 2) Considering these dynamics sexual deepfakes are considered unethical compared to other forms of deepfakes
Gosse & Burkell (2020)	Identifying discussions of deepfakes for porn and politics in media articles	Discourse analysis on news articles	News media discuss deepfakes in four ways: 1) Too easy to produce and disseminate 2) Aid creation of false beliefs 3) Undermine democracy 4) As non-consensual content
Van der Nagel (2020)	Addresses potential infringements on consent due to pornographic deepfakes	Case study	Proposes a verification system in order for women to be able to reclaim their own digital autonomy
Popova (2020)	Compares pornographic deepfakes to other sexualised audience engagement	Comparative content analysis of Reddit channels	1) Argues that audiences of deepfakes do not concern themselves with the intimate and private life of the celebrity 2) And that the deepfakes are created and disseminated within smaller communities
Maddocks (2020)	Investigates the relationship between political and pornographic deepfakes	Discourse analysis on twitter sourced material	1) Identifies that pornographic and political deepfakes are used in similar ways to silence or blackmail individuals 2) The technology is mainly used to target women
Wang & Kim (2022)	Focus on users perceptions and reactions to sexual deepfakes, with the key focus on K-pop idols	Online Survey	1) Crucial predictors: Previous perceptions about sexual harassment, experiences viewing pornographic content, and gender 2) Non crucial predictors: Media exposure and age 3) Highlights the need for collaboration between government, the pornography industry and education to address sexual deepfakes

2.2 Consent Online

2.2.1 Placing consent in the digital space

In scholarly discussions, previous studies have distinguished between the willingness to have sex and consenting to have sex (Marcotte & Hille, 2021), and it is important to consider these concepts when discussing the widespread development and use of digital tools for sexual interactions. A definition of consent that takes these two aspects, willingness and consent, into consideration is the one presented by Hickman and Muehlenhard (1999). They define consent as the act of "freely-given verbal or nonverbal communication of a feeling of willingness to engage in sexual activity" (Hickman & Muehlenhard, 1999, p. 259). Although clear and concise, this definition was built upon and made into a three-tiered conceptualisation of sexual consent by a review article done in 2016 by Muehlenhard et al. This definition described consent as 1) "an internal state of willingness", 2) "an act of explicitly agreeing to something", and 3) a "behaviour that someone else interprets as willingness" (Muehlenhard et al., 2016, p. 462). This model, to date, is the most vigorous framework in academic literature when discussing sexual consent (Marcotte & Hille, 2021), as it acknowledges the concepts of willingness, an act of communication and that consent is given freely.

However, to expand on this definition further, it needs to be noted that the definition does not consider the involvement of social power dynamics and how this may affect the ability to give or withhold consent. An important aspect to consider for this research. The 'networked feminist theory' (Balsamo, 1996) recognises this and gives a way to analyse how digital networks, online communication and social media impact feminist practices. This theory can be used as a base to analyse online consent, as it acknowledges how patriarchal power structures impact the digital domain and how this can create new power imbalances. It is important to understand and acknowledge these existing power dynamics, especially in regard to concepts such as consent, where power dynamics play a vital role in how consent functions.

In the text by Marcotte and Hille (2021), who used the 'networked feminist theory' (Balsamo, 1996) as their research lens, they found that communicating consent in digital spaces is a vital aspect to consider as consent practices online reflect the

gendered aspect of consent offline. There are clear gendered differences in how consent is communicated between men and women, and these have "different social outcomes based on gender" (Marcotte & Hille, 2021, p. 328). However, besides the research by Marcotte and Hille (2021), no extensive research has been done on how miscommunications and misinterpretations regarding online consent might develop on digital platforms.

The key difference when discussing consent online and offline is how consent and willingness are placed in the context of time. Burgin (2019) states that consent provided at one specific moment does not guarantee the same consent at another time, but this changes when discussing consent online. This has been researched by Marcotte and Hille (2021) through various digital consent case studies. They state that in the context of digital infinity and dissemination, consent interacts with a varying temporal frame. Meaning that consent in a digital space is not given in a singular time frame but rather has to be given in an unlimited one. This means it is important to understand digital consent as a divergence from offline consent, as consent given at a particular moment needs to be understood in the framework of digital infinity (Marcotte & Hille, 2021). Regardless, this notion of digital infinity for consent is still highly under-researched, as most online consent studies discuss the intricacies of image-based abuse rather than the impact of digital infinity.

Within this research paper, consent will be defined by the proposed model by Mulenhard et al. (2016) and the 'networked feminist theory' by Balsamo (1996). Also keeping in mind how Marcotte and Hille (2021) conceptualised consent as an ongoing and repetitive process that moves beyond the initial event of (non)consent giving. Connecting this conceptualisation of digital consent to deepfake pornography, it is important to note that deepfakes can further re-organise this understanding of consent (Van der Nagel, 2020). As they not only function within the digital sphere and, thus, the digital infinity that it provides, but they also remove the necessity to have access to sexually explicit images of the targeted individual in the first place.

2.2.2 Consent and digital (sexual) autonomy

The concept of digital infinity also plays into another aspect of consent online: consenting to general data being online. When discussing consent in the digital sphere, concepts such as autonomy of content also play a vital role in how consent is perceived

and how it can be given online (Gosse and Burkel, 2020). Whether inter-personal or with a digital platform, data has become a form of currency or commodity (Gosse & Burkell, 2020). In this situation, any data that has been shared is no longer owned by the sharing individual but rather by the platform on which it has been shared (Mayer & Lu, 2022). This adds another layer of complexity to consent online, in which content and the attributed consent no longer belong to a bi-directional relationship but instead include third parties.

This concern is reiterated by the digital self-determination theory (Belli, 2018), which is a theory that believes that individuals should have the ability to fully make choices about what happens with their personal data and digital interactions. Furthermore, it states that these online experiences should be done without interference or coercion from a secondary party. Belli (2018) also argues that the internet and what happens within its sphere should be structured and controlled by the users to ensure an equitable space that improves the lives of the individuals using it.

Now, the digital self-determination theory can also be applied to the intersection between data and sex, especially in regard to the relation this intersection has with consent. Sexual digital autonomy is impacted by the (non)consensual practices found online and how these are divided by gender. In the text by Sundén (2020), it is found that the way data, especially sexual data, is handled by third parties (such as companies like Meta) varies between male and female users, and the social reaction towards leaked sexual data also diverges drastically. More specifically, it was found that “the sexual agency and autonomy of men in terms of normative transgression was nobody else’s business” (Sundén, 2020, p. 5) but that the blame is on the technologies and those who exposed these data leaks. In contrast, when discussing the sexual autonomy or the right to digital sexual privacy of women online, the results were much different. Here it became apparent in the research by Súnden (2020) that the displacement of fear of digital technologies is placed on women's bodies, rather than the existing technologies. The reason given is that these instances are viewed through a social lens of slut-shaming, which does not happen for their male counterparts. Thus, the blame falls onto the women rather than the existing technologies (Sundén, 2020).

With this, it can be identified that how sexually explicit data is handled online is already gendered, having implications for both sexual privacy and consent online. This gendered aspect could possibly be further implicated by the creation of deepfakes,

where sexually explicit data does not need to exist for women's bodies to be blamed. However, an interesting point made by Sundén (2020) in her text is that the existence of digital sexual autonomy and the shame that comes with data leaks of women is not based on the technology that leaks such content, but rather its problematics lies in the reaction of the public. This is reiterated by Chun (2016), that also argues that instead of

> Pushing for privacy that is no privacy – a security that fosters insecurity – what would happen if we demanded more rigorous public rights? If we fought for the risk to be exposed – to take risks and to be in public – and not be attacked? (p. 11)

With that, Chun (2016) argues that digital sexual autonomy can only exist for women when the right to take risks and (sexually) exist in public is not met with shame and outrage but that this is rather redirected to the ones who did the sharing. As Albury (2017) questions, "What if being known as 'someone who gossips and shares sexual images without consent' was the more shameful identity and was presented to young people as such?" (Albury, 2017, p. 722). With this distinction, it is believed by the authors that women would regain their sexual digital autonomy and the right to digital privacy in the same way men already have (Chun, 2016; Albury, 2017).

Furthermore, studies that have researched deepfakes in the broader scope of ownership have noted the same pattern of infringement on the digital agency of women (Van der Nagel, 2020). Stating that the "harnessing [of] artificial intelligence to identify and replace faces in porn videos, the editing of images, and even the idea of owning one's own image, has a history entwined with the manipulation and control of women" (Van der Nagel, 2020, p. 428). Moreover, when it comes to deepfakes, Van der Nagel (2020) noted that the blurring lines of content ownership add another layer of complexity to consent practices but did not explicitly research this. Deepfakes, as a form of AI learning, can be considered detached from the original content, as it does not explicitly use said content (Meskys et al., 2021) but only uses it to learn to reproduce the image.

The existing research shows how challenging it is to have digital autonomy as data ownership does not solely lie with the individuals but also with third parties that can handle the data as they see fit (Mayer & Lu, 2022). In addition, it is clear that men

and women are not treated equally when it comes to the handling of data and the consequences of (sexual) data leaks (Súnden, 2020). The involvement of third parties distributes responsibilities of consent to multiple parties, and it then becomes harder for individuals to give consent as it moves beyond a bi-directional interaction (Mayer & Lu, 2022). This radically changes not only how consent is viewed and perceived online but also what legal measures can be taken if a non-consensual video is created and shared online.

2.2.3 Non-consensual sexual interaction online

Questions of sexual consent online also include the distribution of non-consensual images. Deepfakes can fall under this distribution and expedite it. In cases where recordings of consensual sexual acts in-person are used and dispersed without the knowledge or consent of the person in said material, such as nude photos or sex tapes for example, it highlights the "disconnects between consensual in-person sexual experiences and non-consensual digital sexual expression" (Marcotte & Hille, 2021, p. 324). These kinds of cases fall under what can be considered image-based sexual abuse (Henry & Powell, 2016**)** or 'revenge porn'. This includes a range of activities, such as the non-consensual creation, dissemination and use of images (Powell et al., 2019).

These kinds of online non-consensual activities discussed by previous studies illustrate how complex and nuanced the topic of digital consent can be (Bates, 2017; Marcotte & Hille, 2021). Distribution within the online sphere is not limited to individual connection and can therefore happen from anywhere. This means that the likelihood of this form of dissemination happening without the subject's knowledge increases, preventing the exchange of consent (Marcotte & Hille, 2021). This is reiterated by previous studies that have revealed through surveys that nearly one-fifth of individuals are sharing sexually explicit images without asking the individual for consent (Garcia-Gomez, 2017). The participants in this study were shown to have shared the image with approximately three or more friends, meaning the image's viewing grew exponentially, and this further increased when the image was shared through online platforms (Garcia-Gomez, 2017).

This problem is expedited by the possibility to create deepfakes, a form of pornographic content creation that requires no physical or in-person sexual expression

from the victim. In addition, creating non-consensual sexual deepfakes can be considered a form of technology facilitated sexual-violence (TFSV). The term TFSV was first coined by Powell and Henry (2017) that aims to conceptualise the various forms in which digital communication technologies aid or perpetuate "criminal, civil or otherwise harmfully sexually aggressive and harassing behaviours." (Powell and Henry, 2017. p. 5). Their research specifically aims to take a victim-centred approach as it intends to acknowledge that harm, specifically done within the online sphere, may not be universally considered harmful but may have had a significant impact on the affected individual.

Catherine MacKinnon's framework of "sex-based group defamation-as-discrimination" (MacKinnon, 1993, p. 99) is in line with Powell and Henry's (2017) discussion with an elaborated argumentation. This framework is used as a lens to highlight the nature of deepfake pornography and its ways of being used as not only a fantasised non-consensual sex video but rather as an intended attack on the sexual agency of the victim(s). Furthermore, it emphasises that deepfake pornography is as real as sexual assault and has similar psychological and physical impacts. As stated by Graber-Mitchell (2021),

> Just as rape survivors live through daily re-traumatization simply by existing in their bodies, victims of deepfake pornography feel exposed, vulnerable, and taken advantage of by the very act of being seen by others. The existence of an illusion of sex becomes as powerful as sex itself (p. 8).

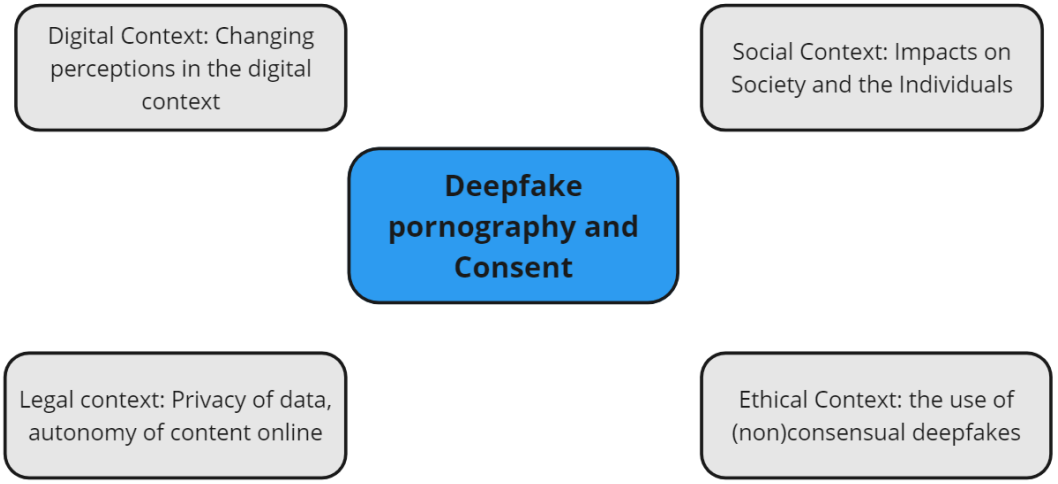
To see the implications that non-consensual content can have on victims' lives, it is highly concerning that these harmful consequences are not perceived as such by society (Marcotte & Hille, 2021). Moreover, the ease with which deepfakes are created exacerbates these harmful consequences, removing any sexual digital autonomy from the victim (Van der Nagel, 2020). This is further implicated by the fading between truth and fiction that deepfakes create, which can cause psychological and physical harm to individuals (Graber-Mitchell, 2021), and this can not only impact individuals but also, on a societal level, can impact understanding of consent by influencing how reality is viewed.

2.3 Moving Towards a (new) integrated framework

For this research, an intersection between data, sex and consent had to be made to place them into the discourse on pornographic deepfake creations. These intersections and distinctions also allow a deeper understanding of the potential implications of sexual deepfakes within pornography on consent online and the complexity of the issue. To successfully do this, a framework was established to aid the operationalisation of the research. The original framework was taken from the text by Quadara and El-Murr (2017), which researched the impact of pornographic content on children. They identified four principal research areas. These were then adapted using the existing research mentioned above in order to elaborate on the research question and answer it and the sub-questions. The adapted framework is as follows (Figure 2.2):

Figure 2.2

Situating deepfake pornography and its impacts on consent as an issue



Note: Framework adapted from the text: "The Effects of Pornography on Children and young people" by Quadara and El-Murr (2017).

When discussing the problematics of deepfake pornography and consent, it needs to be addressed as a multi-agent system (MAS) as described in the text by Öhman (2019) in his research on deepfakes and identity ethics. This is an assemblage of interconnected forces that, due to the distributed nature of the system, make it challenging to attribute direct responsibility for an issue to a singular entity. For

deepfake pornography and consent, this means that the MAS functions in a way that the various categories create, enforce, and reflect one another, oftentimes overlapping in conceptual and academic discussions. Hence, the framework proposed comprises four agents that are divided in theory but connected in practice. Nevertheless, the framework addresses the crucial aspects of deepfake pornography to answer the research questions and expand on previous literature.

3. Methods

3.1 Qualitative Mixed-Method Approach

This paper aims to advance the research done on consent and pornography by having a qualitative approach to the use of deepfakes within the pornography industry. The reasoning was due to the nature of the topic and research question, which involves societal perceptions of this phenomenon, and therefore needs a flexible method to analyse deeper meaning (Johnson, 2011). In addition, the research by Bancroft (1997) highlighted the importance of the cultural meaning and significance attributed to sexuality, a concept that is intrinsically intertwined with pornography as a research area and, therefore, valuable to this study.

Qualitative studies allow for personal and culturally different interpretations of a topic and are used to dive deeper into the meaning of a phenomenon with the aim of explaining the why and hows of a topic (Löfgren-Martenson & Mansson, 2009). A combination of focus groups and expert interviews was selected to gain insightful information, as these approaches allow for in-depth research on the subject (Powell & Single, 1996). The decision for a qualitative mixed-method approach was to gain access to the detailed and knowledgeable information from the expert interviews and the general experiences revealed within a group dynamic. Using both of these methods, the data collected could be compared, triangulated, and validated (Morgan, 1998).

3.2 Focus Groups

Focus groups are best defined as a method that brings together individuals to discuss and comment on a specific topic (Powell & Single, 1996). It is a technique that allows for “guided, interactional discussions as a means of generating the rich details of complex experiences and the reasoning behind actions, belief, perceptions and attitudes” (Powell & Single, 1996, pp. 499 - 500). It was decided as one of the methods for this paper, as it allowed for an understanding of the discussions on deepfake pornography, especially focusing on the values and perceptions of the use of TFSV and consent within the industry. This study is particularly interested in how individuals note this kind of trend and where the differences in gender perceptions could potentially be. Furthermore, according to Morgan (1998), due to the topic's sensitivity, this method is also suitable as the approach is deemed more respectful and sympathetic than the other

methods. Lastly, this method is especially insightful into group dynamics on this topic, which can be considered controversial, and thus should be looked at in-depth rather than on a larger scale (Plummer-D'Amato, 2008).

3.3 Expert Interviews

Expert interviews are best described as a method that aims to gain insights from "personal matters, such as experience, values and decisions, occupational ideology, cultural knowledge, or perspective" (Johnson, 2011, p. 3). They are often used to delve even deeper into knowledge in a space where the individual is not interrupted. Furthermore, due to the specialised knowledge, the expert could provide both validations of research findings from previous literature or the focus groups, as well as an exploration of the complexity of the topic at hand (Johnson, 2011). Especially considering the layered approach this paper has towards deepfake pornography, an approach that, to the best of the researcher's knowledge, is novel, it was interesting to analyse the expert opinions on this. In addition, the various expertise of the individuals covered the various aspects of this research; thus, they could provide useful and rich qualitative data. Moreover, the interviews were used to identify similarities or discrepancies between the perceptions of the focus groups and the experts, which revealed some interesting findings.

3.4 Sampling and Participant Demographic

The research had five focus groups with five people per group and three expert interviews, excluding pilot testing. All focus groups and interviews were held online, which facilitated the diversification of participants. However, due to this, one focus group ended up only having four participants because of last-minute technical difficulties on the participant's part. They were held between the 23rd of March 2023 and the 6th of May 2023. The discussion lasted between 45 minutes and 70 minutes for the focus groups and between 50 and 80 minutes for the expert interviews.

For the focus groups, it was chosen to have smaller groups due to the sensitivity of the topic and the fact that smaller groups would allow for deeper and more meaningful discussions. This is because smaller groups allow individuals to have the possibility to speak more, making the experience more interactive for everyone, as well as being more intimate, which helps people to discuss their opinions more freely.

These 24 people were sampled using voluntary response sampling in the hopes that it would diversify the participants and not limit them to the researcher's access. Furthermore, this type of sampling ensures that all participants were aware of the nature of the topic and willing to join the research and focus group, which is important for this type of research due to the sensitivity of the material. The criteria for the participants were that they were between 18 and 35 years of age and that they have no close relation to the researcher. These criteria ensured that participants did not feel restricted in their answers by their ties to the researcher. For example, if participants knew the researcher prior to the focus group, it could have potentially skewed their answers as they might not want to share information with someone that they know well. The secondary criterion was so that participants could relate to the other participants through their age. Being able to relate to other focus group members, at least through age, can help individuals to be more open to sharing their opinions.

To recruit the participants for the focus group, a message was posted in various social media groups and distributed on different university channels. This was specifically done to find participants that met the age requirement. The message stated the topic of the focus groups was the use of deepfakes in pornography and also stated that discussions about sexual violence may occur during the discussions. With this, the participants were cautioned that this topic may spark previous trauma or intense emotions as the topic dealt with certain types of sexual violence. Hence, it was ensured that each participant consented and knew that they had the option to withdraw at any point in case the discussion became too heavy. This was vital for the research, especially considering the topic, as no harm was intended with these focus groups.

In the end, the participants gathered were all between the required age, with the youngest being 18 and the oldest 32, and there were more women than men in the study, with the gender makeup being 17 women and seven men. Moreover, the sample group gathered was diverse in nationalities, with various different countries within Europe, the middle east and Asia being represented, which allowed for various insights from different countries and cultures that enriched the discussions. Table 3.1 expresses the demographic makeup of each focus group when they were held, and the pseudonyms that will be used throughout the paper.

Table 3.1

Overview of the focus group participants

Focus Group #	Pseudonym	Age	Gender	Nationality
Focus Group 1	Elena	20	Female	UAE
25/03/2023	Tamara	32	Female	Ukrainian
	Alina	21	Female	Dutch
	Arnold	23	Male	Sinhalese
	Siena	19	Female	Dutch
Focus Group 2	May	22	Female	Dutch
17/04/2023	Sonia	25	Female	Greek
	Carolina	18	Female	Spanish
	Cecilia	25	Female	Greek
	Jasmine	19	Female	Finnish
Focus Group 3	Samantha	24	Female	Dutch-Brazilian
22/04/2023	Henry	30	Male	Egyptian
	Ivan	23	Male	English
	Meredith	23	Female	Chilean
	Charlotte	19	Female	Brazilian
Focus Group 4	Felicity	22	Female	Danish
26/04/2023	Pablo	20	Male	Spanish
	Mike	19	Male	German
	Beatrice	18	Female	Spanish
Focus Group 5	Mary	21	Female	German
06/05/2023	Nora	23	Female	Polish
	Keith	23	Male	Polish
	Felix	22	Male	German-Spanish
	Amy	21	Female	French-Iranian

For the interviews, on the other hand, the recruitment and sampling process was based on purposive sampling. This was due to the fact that the interviews are meant to be with experts in the fields of (online) sexual violence, artificial intelligence, cyber security or deepfakes specifically. Therefore, the criteria were based on their expertise,

and the recruitment process was through researching companies or individuals in the field and reaching out to them. Furthermore, the already interviewed experts gave contacts for further possible interviewees. In the end, three experts were interviewed, and Table 3.2 expresses an overview of their expertise.

Table 3.2

Overview of expert interviews

Expert Interviewee	Gender	Expertise
Expert 1 23/03/2023	Female	CEO of an NGO helping sexual violence survivors, helping both victims of online and offline abuse
Expert 2 01/05/2023	Male	Expert in deepfakes, making them for business and an upcoming project in which deepfakes are used for trauma and exposure therapy
Expert 3 03/05/2023	Male	A lecturer in cyber security in Den Haag University, and a PhD candidate for research done on TFSV in Tilburg

To ensure consent from all parties, a consent form was sent to everyone (see appendix A), for which verbal consent was asked before each session, and consent to record and use the information was also given verbally. Thus, this study was conducted in agreement with four main ethical pillars: information, consent, confidentiality and the right to withdraw.

3. 5 Data Collection and Operationalisation

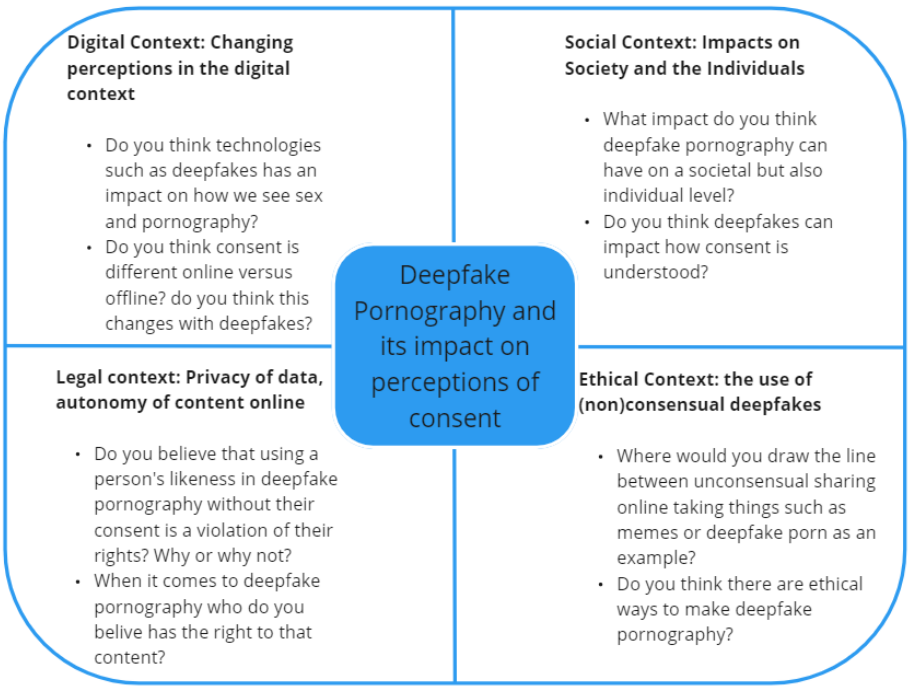
Whilst conducting the focus groups, there was a short moment for introduction for everyone as well as a moment for the participants to consent to being recorded. The topic was introduced, and then questions surrounding the topic were asked (See Appendix B). The questions were predefined using previous research that aided the identification of specific topics of conversation and, thus, the creation of the questions. Furthermore, this helped combat research bias (Johnson, 2011), as the questions were not based on personal ideologies. Lastly, the questions were open-ended, which ensured fruitful discussions and were neutrally phrased to avoid biased discussions. However, even though questions were predefined, the nature of the discussion would allow for

follow-up questions or spontaneous questions to allow for elaborations on new themes or topics.

The questions that were asked focused on the following four aspects that relate to the framework proposed prior; 1) the digital aspect, 2) the social aspect, 3) the legal aspect, and lastly, 4) the ethical aspect. Hence, all of these topics were used as a foundation for the focus groups and expert interviews. Figure 3.1 provides some examples of how the concepts were translated into questions. It has to be acknowledged that due to the intersection of topics and concepts across the aspects, there were instances where a clear categorisation of the questions was not possible. Furthermore, the questions varied for each group and interview in accordance with the expertise and flow of the conversation. Nevertheless, each aspect was touched upon by all participants.

Figure 3.1

Examples of questions touched upon during the focus groups and expert interviews



The same topic guide was used for the expert interviews, but the questions were more specific and technical, allowing for deeper insights from the expert. In addition,

the questions were tailored to the expertise of the interviewee, which allowed for the expertise of the individual to be explored.

3. 6 Data Analysis

The focus groups and expert interviews were transcribed verbatim and analysed. The transcription was done using a software called descript to aid the analysis. Nevertheless, the researcher listened to the recordings again and proofread the transcriptions to find any missing information or errors. The analysis was done using thematic analysis, which helped identify patterns within the discussions and interviews that allowed for an interpretation of the perceptions of using deepfakes in pornography and its implications for consent. This method of analysis is best described as a method “for identifying, analysing, and reporting patterns (themes) within data" (Braun & Clark, 2006, p. 79) by segmenting the data and reassembling it to transform it into findings (Boeije, 2010). According to Boeije (2010), there are three key components to coding within qualitative analysis: open, axial and selective coding. Thus, the first step that was taken was open coding, which was done after the transcription process, in which the researcher coded the text by highlighting aspects that are important for the topic (Boeije, 2010). In order to keep an overview of the codes, this process was done using the software Atlas.ti which aided the organisation of the open coding. Even though it is advised by Boeije (2010) to assign only one code to each text fragment, this was oftentimes difficult to apply due to the nature of the research and the complexity of the topic. With this, a coding scheme of 247 codes was achieved.

Next, for the axial coding, the researcher identified dominant codes to reorganise the dataset into categories. Here it is important to define the significant codes and reduce the data (Boeije, 2010). Hence, as suggested, codes that were similar or repeating were merged, and codes used only once were merged with codes that were used more often and were synonymous. This process was done manually, as it was easier for the researcher to visualise the connections that could be made. As a result, 19 axial codes remained.

Lastly, after axial coding, the researcher started selective coding, in which axial codes were merged once more, allowing for eight selective codes to emerge. These were as follows: “Social stigma and deepfake impacts”, “ethical perceptions of (non)consensual sharing online”, “deepfakes and ownership of content”, “digital infinity

and deepfake consent", "consent and the ability to create fake content", "regulations for deepfakes", "social awareness of consent" and "ethical conflicts regarding deepfake pornography", (See Appendix C). These were then further organised into four final themes that aim to answer the research question and reflect the theoretical framework: *The digital aspect, The social aspect, The ethical aspect,* and *The legal aspect*. In the upcoming chapters, these themes will be analysed, compared and contrasted by drawing upon the existing academic body that was discussed in the previous chapter.

3.7 Limitations and Ethics

Lastly, due to the sensitivity of the topic, a few ethical aspects and limitations need to be addressed. When it comes to studying pornography as a field in general, the text by Fisher and Kohut (2020) noted several key criteria that need to be kept in mind when researching pornography which is highlighted in the table below (Table 3.3).

Beyond the considerations mentioned in the table, it was important to further critically assess the possible limitations of this research. To achieve validity and credibility, the research approach was described thoroughly, and personal biases were kept in mind. This was done by aiming for an open attitude whilst also being aware that, as a woman, the researcher may have had potential biases regarding the topic. This bias was countered by embedding the research in previous findings. Furthermore, it must be noted that due to the small sample size and the nature of the research aims, it cannot be generalised to a larger population but adds value through its depth of content. Therefore, it is even more crucial to relate any findings back to previous research in order to be able to justify them critically.

Table 3.3

Table outlining criteria for pornography studies and how they were applied for this study

Methodological considerations for pornography research as presented in the text by Fisher and Kohut (2020)	Application of the considerations for this study, if applicable
Evaluate whether bias in research design has been introduced by harm-focused or other ideological narratives or is otherwise present.	Research is done with personal bias in mind, as the researcher is a woman and therefore may impact results. With this in mind, a self-reflective approach was undertaken by critically assessing every aspect of the research for personal bias.
Consider how self-selection and participant attrition may limit the generalisability of study findings.	Focus group participants were gathered through voluntary response sampling to avoid self-selection by the researcher.
Consider whether or not specific conceptual and/or operational definitions of "pornography" and "use" may influence interpretation of the study findings.	The theoretical framework critically discussed various operational definitions to apply one that encompassed most aspects of the term "pornography".
Determine the extent to which operational definitions of "pornography" and "use" in a given study may have empirical evidence of construct validity and reliability.	The definition of pornography applied within this research was defined in agreement with previous literature in order to impact results as little as possible and add to the validity and reliability.
Determine the extent to which the conceptual and/or operational definitions of "pornography" and "use" in a given study may limit or extend the comparability of the study findings to other related work.	The use of aspects of definitions by feminist theories could potentially impact and limit results. However, this definition was not shared with participants in order to avoid biases.
Consider whether specific social/environmental contexts of pornography use play a role or should play a role in the study findings.	The findings are based on social perceptions and understandings, and thus critically evaluate pornography in its cultural context.
Be sensitive to how violent pornography is defined, and consider the implications that specific definitions have for the study findings.	Aware that the study critically addresses pornography through a feminist lens, that does define it as violence against women which can impact findings.
If a causal models is assured, determine whether alternative direction and third-variable models have been discussed and empirically evaluated.	The research uses the multi-agent system approach (Öhman, 2019), which does not apply one agent as the sole cause of an issue, but rather a system that works together; therefore, no causal model was used.

Note: This criterion presented on the left-hand side is taken from the text: "Reading Pornography: Methodological Considerations in Evaluating Pornography Research" by Fisher and Kohut (2020).

Lastly, a potential limitation was the use of online meetings, as both the focus groups and the expert interviews were held online via the Microsoft Teams software. This, on the one hand, created a digital divide between the researcher and the

participants, but on the other hand, allowed for individuals to continue staying in their space, which ethically may have been positive as the conversations could be quite heavy. In addition, the use of online meetings allowed for a diverse group to participate in the study, adding to the richness of the data. However, this also created some unforeseen technical disturbances in which one participant had to drop out as the focus group was starting, meaning one focus group only had four participants. These are all things that were considered when analysing and discussing the data.

4. Results

To effectively present the analysis and answer the research question, four main categories were distinguished in the focus group discussions and expert interviews. These reflected the theoretical framework proposed and, thus, were as follows: The digital, the social, the ethical, and the legal aspects of (non)consensual deepfake pornography (see Table 4.1). The coding tree (see Appendix C) illustrates the coding mechanism under which the sub-themes were retrieved.

Nevertheless, it needs to be noted that, as discussed in the methods and the theoretical framework, the discussion around deepfake pornography and consent bases itself on a multi-agent system, allowing for the themes to interconnect on a conceptual and argumentative basis. This means arguments are nuanced within their themes and impacted, enforced, and strengthened by the other themes. The analysis of the various themes relates back to the three sub-questions that drive this research: *What factors impact how deepfake pornography and consent online are perceived? To what extent does deepfake pornography impact the understanding of digital autonomy? And how does this impact consent online? How does the context of deepfake pornography impact ethical understandings of consent online?* These are answered by the discussion of the result and are extensively answered by the interconnection of the themes at the end of this chapter.

Table 4.1

Overview of themes, sub-themes and key findings

Theme	Sub-Theme	Emerging Arguments
Digital Aspect	Consent and the ability to make fake content	- Fake content re-organises understanding of consent - Potential for TSFV increased through the ability to create fake content - Created unseen third perspective
	Digital infinity and deepfake consent	- Digital infinity undermines dynamic consent - Sexual violence becomes continuous online
Social Aspect	Societal awareness of consent online and deepfakes	- Understanding of consent online differs from consent offline - Lack of societal awareness of consent online - Hyperreality and accidental infringement on consent
	Social stigma and deepfake impacts	- Deepfake harms due to social shaming, stigma and context - Gendered social reactions
Ethical Aspect	Ethical conflicts regarding deepfake pornography	- Personalised deepfake pornography in relation to sexual fantasies - Ethical understandings are created due to the context - Paedophilia and ethics
	Deepfake compared to other non-consensual sharing online	- Identifying the ethical line between deepfake porn and other sexual content - Context impacts ethics: Humour - Deepfakes compared to Memes
Legal Aspect	Autonomy of content and deepfake consent	- Ownership through content creation - AI learning is similar to human learning - Ownership through the right to own likeness - The dilemma of identity leakage and ownership of content
	Deepfake regulations and consent	- Lack of legislative framework - Doubts in framework effectiveness - Possible legal frameworks

4.1 Digital aspects of (non)consensual deepfake pornography

The changing dynamic caused by digital devices in the understanding of consent and pornographic content is discussed within this theme. This theme specifically analyses the discussions on how the creation of fake content blurs lines of online content and what impacts this can have on non-consent online. As well as how digital infinity can implicate dynamic consent.

4.1.1 Consent and the ability to create fake content

Consent is most often defined as the act of "freely-given verbal or nonverbal communication of a feeling of willingness to engage in sexual activity" (Hickman & Muehlenhard, 199, p. 259). However, when exposing this definition to the ability to create fake sexual content of individuals, it loses the ability to encompass consent discussions fully. This was noted by the participants, who realised that fake content completely re-organises the understanding of consent that has been created. Some participants even argued that fake content removes the idea of consent completely, as no interaction between individuals even needs to take place for a deepfake to be made. In addition, the consensus across the focus groups was that explicit fake content that is not consensual infringes on the sexual autonomy that one has online. The following statement is an example of this:

> Anyone can be used for a deepfake and I think that creates like a very unsafe environment for people to explore their own sexual being, but as well as their being. I think to just have the feeling that you can be used at any time, at any moment, for something like a deepfakes [impacts autonomy]. (Siena, focus group 1)

This statement reflects arguments made in the text by Marcotte and Hille (2021), who discussed how nuanced and complex digital consent is. They highlighted that online consent is not bound by bi-directional interactions but instead requires no interaction at all. Non-consent can now happen from anywhere and be done by anyone, regardless of the relationship with the victim, and this is expedited by fake content.

All focus groups discussed the threat of deepfakes in regard to sexual violence and the potential to create content that can harm an individual's life but also bodily autonomy. Especially how non-consent now has a place in our homes as well, as it is not only “strange men attacking and raping you in the streets, but it's also like happening in your in your homes and now on the internet, it’s everywhere.” (Jasmine, focus group 2). This insight further means that the potential for non-consent online is increased, and the capacity to make sexual violence more accessible in the digital sphere as well (Marcotte & Hille, 2021). Expert 3 agreed with this notion and added that deepfakes not only potentially expedite sexual violence but can also aid the ability to create even more harmful content, as it “makes it easier to create more shameful content. For example, you could make a deepfake of me having sex with a child”. This can therefore simplify the process of sexual violence through the act of pornographic content creation (Goldfarb, 2015).

Hence, the data shows that deepfakes as a form of TFSV (Henry & Powell, 2017) can further perpetuate the sexually aggressive harassment reflected in offline society, placing it in the online sphere. The technology has allowed individuals to create content that can potentially harm and impose on someone’s sexual agency without having any prior relationship with that individual.

Lastly, Expert 2 pointed out that deepfakes as fake content have the ability to create an “unseen third perspective”. When asked to explain this idea, he stated;

> A lot of people see versions of themselves on the internet, and they start to believe that they are, they are it. And what we can do now is we can create versions of yourself that don't exist. And yeah, they, they start to form almost the identity that belongs, that that connects to it. [...] So if you look at yourself in a pornographic deepfake, and you are doing all those things that you have never done before. Some point, in your life or these images in your mind, could all almost become a form of memory. So you could almost see it as a memory, something you've done before.

This understanding of deepfake is closely related to the concept of 'media effects' coined by Bandura (2001), that believes that content created by AI technology will have an impact on how consent and agency are perceived (Van der Nagel, 2020), by blurring the lines between a physical and digital self (Gamage et al., 2022). However, having this ability to impact someone's self-understanding by creating fake content, at the very least, takes away any form of digital autonomy an individual has online.

This infringement of the digital self was reiterated by Expert 1, who elaborated on it by stating that a key phenomenon she identified is that individuals who have been impacted by these nonconsensual acts online show no difference in psychological and even physical impacts compared to offline sexual abuse. Some have even developed PTSD symptoms, stating that it is interesting "how powerful your brain is [...] it still has a physical reaction even if you're not being touched". This is reiterated by Graber-Mitchell (2021), who states that "the existence of an illusion of sex becomes as powerful as sex itself" (p. 8).

Considering this, it is vital to remember the impact deepfakes can have on victims of non-consensual sharing, a potential that does not go unnoticed by the participants. This fake content then also re-arranges the understanding of consent that has been created, and consent practices in regard to that content are met with scepticism.

4.1.2 Digital infinity and deepfake consent

The idea of the "digital footprint", which was brought up by Charlotte from focus group 3, can be linked to the concept of digital infinity, a concept that describes the existence of data online (Marcotte & Hille, 2021). It plays a vital role in the way consent online is understood, as it imposes the aspect of dynamic consent. As Burgin (2019) states, consent given at one point in time does not mean consent has been given at another time. A view that is echoed by Expert 3, who argued that even when "consent has been given, it can be taken as well. So consent is not like a static thing". Yet, when this is discussed in the online sphere, that notion is difficult to uphold due to digital infinity. This is due to the fact that once something is on the internet, it is near to impossible to remove it again (Marcotte & Hille, 2021). This was reiterated by the focus groups that agreed across the board that once something is online, they have the perception that it stays online, as you cannot "even consent fully to all our data on the internet, let alone send in kind of private image or anything else," (Ivan, focus group 3).

When placing this concern in the context of deepfakes, it becomes even more critical because, as stated before, an exchange of consent is not needed to create a deepfake and disseminate that content. Nonetheless, the content, and therefore, the potential harm to the individual, still exists within the context of digital infinity. Meaning that even though individuals may not even initially be aware that fake content exists online (Marcotte & Hille, 2021), the impact it may have on them will be long-term.

This was also touched upon by focus group 5, who discussed that a non-consensual deepfake pornographic video can not only impact you once but rather can do so infinitely. Amy explained,

> But it's not like a one-time thing. It's not like you got assaulted once, as long as that video is up every single new person who see that against her consent, not knowing that it's fake, believing that it was a choice that she made, it's gonna re-traumatise her.

This statement highlights the continuous effect deepfakes can have on consent practices online, as well as the fact that the implications of non-consensual deepfakes in an environment of digital infinity can be extremely harmful to an individual. Thus, in the same way rape survivors may be re-traumatised by touch or existing within their bodies, individuals implicated in deepfake pornography can be re-traumatised by the act of being seen by others (Graber-Mitchell, 2021).

4.2 Social aspects of (non)consensual deepfake pornography

This theme aims to analyse factors that impacted the understanding of consent, such as questions of how consent is understood online and how societal stigmas are considered the key aspects of non-consensual deepfakes. Furthermore, the social context of the pornographic deepfakes impacted how consent and the severity of non-consent were perceived by the participants.

4.2.1 Societal awareness of consent online and deepfakes

Consent online involves sharing data, whether explicit or not, and this data is not kept within that private relationship but includes third parties (Mayer & Lu, 2022). With

this in mind, the dynamics of consent are not bound to only the involved participants but are subject to the social dynamics created by the online world (Balsamo, 1996). This concern for changing consent dynamics online was reiterated by all of the focus group discussions and two of the experts. They highlighted the loss of control they experienced online that negated realities of consent. The quote from Ivan from focus group 3 is an example of this:

> With two people in a room together, it's very easy for two people to give consent. Whereas once you go on the internet, I think it's more, either it's a one-sided that if you post that image, you consent for that image to go out there. [...] But there's also a third party involved. There's whatever image you, whatever thing you share it on. So for example, us today, we are all consenting to be here, but like Microsoft and Teams is also a party in this. You know, they, they're also involved, and is that consent?

This statement is an example of the social anxiety that individuals have when it comes to how consent functions online. Furthermore, this fear is expedited by the threat of deepfakes and the possibility of losing further control online, and according to Samantha, “if you don't have control, there's no way of giving consent. And we have less and less control that goes on the internet. So I don't think consent really exists.” The participant further stated that the lack of control also stems from a lack of social unawareness and understanding of what consent online means, a thought reiterated by two of the experts.

Both Experts 1 and 3 discussed social awareness in relation to how consent is practised online. They both argue that our society has not yet conceptualised and regulated data sharing to such an extent that consent can exist, as data moves so freely on the internet, and the borders of a digital self have not yet been fully defined. As mentioned by Expert 3, “the question on consent itself [...] leaves open so many parts that are complicated or difficult or are open for interpretation”. Allowing for various interpretations of consent online means that the practice of consent is also tied to the perceptions of how media can be used online. These perceptions can then also vary

between the different parties involved. In general, the understanding of consent online by the participants was that it should not be different to offline consent, but that realistically, “the internet is just by design a different framework which allows for things like anonymity in certain spaces, which already confuses the person of consent to a further degree” (Mike, focus group 4).

A key difference that was noted in the data between some focus group participants and the experts working in sexual abuse was the understanding of deepfake impacts. The experts noted that when it comes to sexual abuse online and offline the reaction of the victim is the same psychologically, whereas the social reaction is completely different. Here Expert 1 states, “I think a lot of people don’t view this [pornographic deepfakes] as sexual violence at all.” She elaborated by arguing that people tend to disassociate themselves from the potential harm sexual violence can cause and that this is increased by TFSV such as deepfakes, where a lot of people are not aware of the potential implications at all.

This statement was proven by the reactions of some of the participants, notably from focus group 1, who believed that even if a deepfake was made of them that the “notion of that existing out there isn’t really going to make me really uncomfortable” (Alina, focus group 1). Many of the other participants agreed with that statement, arguing that they could tell their closest social circle, and that would decrease potential negative consequences they felt could happen, or that they wouldn’t care because “I don’t really care what people think of me in that way,” (Arnold, focus group 1). This highlights the difference in understanding of deepfake issues between the experts and focus group members.

Lastly, the participants further questioned the social understanding of digital consent regarding deepfakes through the concept of ‘reality’. Here they discussed that the impact of deepfakes on the understanding of reality critically influences the awareness of consent in society. This was most prominently discussed by focus group 5 but was also touched upon by Expert 3. Focus group 5 discussed how not knowing content may be fake and being a viewer of such content could implicate you in potentially non-consensual behaviour. This complicates the borders of consent, as interacting with something online could possibly be against the will of an individual. As Amy states:

> Maybe they just searched up her name or, like, I don't know. They were just like scrolling some website and they saw that video. They don't even know her. They weren't doing anything wrong cuz they didn't have any intent.

This discussion relates back to the concept of hyperreality (Baudrillard, 1981), a concept that argues that "the (mediated) truth becomes impossible to discern through the constant noise of synthetic content" (Öhman, 2022**,** p. 317). This cultural state of confusion then has implications on the awareness society has of interactions happening online, reflected in the anxieties discussed above. Furthermore, as mentioned by Alina from focus group 1, the effect of not everyone knowing something is not real would essentially make the deepfake real, and "if it looks real, it's gotta be real in its consequences"'.

These concerns voiced by the focus groups regarding deepfake and consent echo the Experts and their views on how deepfakes can threaten social understandings of consent. They argue that the implications are not clear yet, but the impact deepfakes will have on understanding consent online will be vast, especially once the technology has advanced so far that there is no distinguishing it from authentic content anymore.

4.2.2 Social stigma and deepfake impacts

As mentioned by both Chun (2016) and Albury (2017), an argument can be made that the problem with the existence of sexually explicit content is not the non-consensual sharing or its place in online culture but rather the way in how it is met by society. They discussed that the outrage of sexual content should not be directed at the individuals depicted but rather at the individuals who shared the content without consent. A similar argument was identified in the data, arguing that the harm of non-consent online is based on social shaming and the stigma surrounding sex and porn. Especially how nonconsensual deepfake pornography impacts life offline. The harms identified by the participants, such as job loss or victim-blaming, were often tied back to the social context in which such a deepfake was or could be exposed. The participants believed that this was because "the whole sexuality and sex and images and things are so stigmatised within the public that, you know, an image of you being made public like that can result in job losst" (Ivan, focus group 3). This was confirmed by the argument

made by Expert 3, who stated that in the end, “the picture is not the problem. The problem is the group of people around it pointing. So it’s the social, it’s the social element and the social shaming of it that makes it problematic”.

Three out of five focus groups reached a similar conclusion as Chun (2016) and Albury (2017) did in their text. They argued that if the reactions of the public would shift away from sex and pornography as a shameful act and destigmatise it, the harm of such leaks would be exponentially less. The following quote is an example of such a conclusion made by focus group 3:

> Until we get past the hurdle of actually dealing with the societal issues that we have with just that the sex and sexual assault and anything along that topic and actually destigmatise it, it's gonna be so hard for us to deal with something like [deepfakes] (Ivan).

If we could get past that hurdle, we, as a society, would move away from victim blaming and instead redirect the focus to the act of creating or infringing on personal data and sharing it without consent. However, Experts 1 and 3 believe that society is far from that.

Another social stigma that was discussed in all of the focus groups was the clear double standard people have when it comes to the gender of the victim. The argument was homogenous across all of them that women are held to a much higher level of sexual stigma than men. This was reiterated in the research done by Sundén (2020), who found that in regards to technology, slut shaming is the social fear of digital technologies that are displaced onto the bodies of women, which contrasts the way men are viewed in regards to sexually explicit images online. When it comes to men, the sharing of explicit images is usually met with outrage against the technology with which it was implicated rather than the content (Sundén, 2020). This was echoed by Expert 1, who argued that women are equally condemned for their actions as much as they are perceived as commodities to be used. An understanding of women that is reciprocated in the text by Van der Nagel (2020), who believes that women are understood as digital resources for men and that deepfakes emphasise this impression.

Therefore through the data, it can be argued that it is not the existence of the technology that is the principal issue but rather how it is used and disseminated. Expert

2 highlights this misconception between social and digital understanding of the problem by stating, “There’s a fear for the users. Yes. Not of this technology, so the technology did nothing wrong.” Therefore, the perceptions of consent might not be implicated by deepfakes as a technology but rather deepfakes as devices used by individuals in society. As well as how society meets these cases, creating offline implications that can exacerbate the issue.

4.3 Ethical aspects of (non)consensual deepfake pornography

This theme discusses ethical aspects to deepfake pornography in relation to other forms of non-consensual sharing online. Furthermore, the theme also touches upon the ethical placement of deepfake pornography as one of the factors that impact how consent is perceived.

4.3.1 Ethical conflicts regarding deepfake pornography

A critical theme that arose in the data was the ethical limitations of deepfake pornography, as many believe that there is no way to create deepfakes ethically. On the other hand, others believe that if all parties agree, then deepfake pornography can have a beneficial aspect to it as well. One of these ethical deepfake pornographies that was discussed is the potential for personalised pornographic content. This would expand on the already larger trend of pornographic creation, in which content is not only consumed but also made. This would allow people to view content tailored and made for themselves, similar to sexual fantasies. This argument was also made in previous literature, which argued that even though deepfakes are instinctively considered unethical due to the nature of their function, it is difficult to condemn them without at the same time condemning natural behaviours such as sexual fantasies (Öhman, 2019). According to Öhman (2019), AI learning algorithms needed to make deepfakes function much like the human brain, and in such a sense, on a conceptual level, there is limited difference between an imagined sexual fantasy and a created one. Hence, “it is not unlike an artificial, or at least an augmented, fantasy” (Öhman, 2019, p. 133).

However, Öhman (2019) also believes that the ethical understanding of deepfakes is not based on individual actions, such as creating a deepfake, but rather on the informational environment in which it is done, such as the social system of gender oppression. The participants echoed this, as they mentioned that “women are often the

target of that behaviour" (May, focus group 2) and that ethically deepfakes can be understood as a form of TFSV (Powell & Henry, 2017). The statement made by Elena is an example of this understanding of sexual violence:

> I think at the end of the day, using something like deepfakes is, I believe, for the most or all cases, its non-consensual. The person does not want to be portrayed like this or in this certain way, and [...] it's outside of their will. They're not willing to have this. So it's definitely a form of violence and especially a form of sexual violence when used in a pornographic context.

The link from pornography to the TFSV conceptualisation of deepfakes further brings in the aspect of gender. As previous studies have stated that when discussing pornography as a concept, you cannot disregard gender, considering that they are intertwined with each other and placed in the larger context of patriarchal structures (Lindgren, 1993; Öhman, 2019). With this in mind, the act of creating a deepfake that is pornographic in itself is highly gendered. This placement is arguably also why most participants struggle to identify ethical reasons for the use of deepfakes. As the macro lens of gender inequality creates an unethical understanding of deepfake creation (Öhman, 2019), making the use of it "morally and ethically completely unjustifiable" (Arnold, focus group 1).

Nevertheless, some participants argued for the potential positive uses of deepfakes. One that was identified and discussed by focus group 4 was the potential use of deepfakes for mental health illnesses such as paedophilia, where the ethical aspect of this was a key conflict point. Mike initially brought up the case with the following statement;

> I'm thinking about like really fringe cases and how those would be handled. Right now, for example, if we talk about things like paedophilia and the aspect of deep faking such things, which as I think of it legally poses no clear problem I can think of, but still ethically that doesn't feel right to me.

This sparked an immediate ethical and moral dilemma for the participants, with a few stating that for them, this was not a positive or even ethically correct use. Mainly because “if the child is not old enough to give consent for sex. The child is not old enough to give consent for their face being put in a deepfake” (Pablo, focus group 4). Whereas others saw the potential benefit of this, arguing that if AI uses enough data of various kids to create a new face, then “at that point it’s mush anyway. So I would argue that’s okay. But I agree, if it is a kid’s face, like out in the real world, absolutely not.” (Felicity, focus group 4).

The ‘mushing of faces’ links to the understanding of “identity leakage” (Mirsky & Lee, 2021, p.7), which suggests that deepfake technology can intermingle the facial or vocal characteristics of multiple individuals to the point in which it becomes challenging to identify one singular person (Öhman, 2022). Nevertheless, this argument and discussion made by focus group 4 highlights general ethical problems regarding deepfake pornography. The same argument that Felicity used to argue as a positive for deepfake pornography can also be used as a way to take digital autonomy and ownership away from individuals, as “it’s not like the AI is creating it out of nothing” (Pablo, focus group 4). Thus, In the end, the participants agreed that in ethical terms, it depends on personal morals, norms and values, “and in such dilemmas, I feel like there is no clear right or wrong or yes or no,” (Mike, focus group 4).

Regardless, this discussion underlines the intricacy of ethical uses of deepfake pornography and emphasises how context can change how ethics are viewed in regard to how deepfakes are used and created.

4.3.2 Deepfakes compared to other nonconsensual sharing online

In academic literature, the understanding and definitions of what pornography entails have been shown to be diverging (Hawkins & Zimring, 1988). Throughout the years, various definitions and frameworks for what can be considered pornographic content have emerged, but a general consensus has yet to be made. This lack of a definitional framework is also a problem that created some discussions on ethics between the participants. An ethical concern that the participants brought up was where the line for sexual content using someone's identity or name should be. Mike from focus group 4 explained:

> Because if I were to think about something like a fan fiction, which often also involves sexual interpretation of some celebrity, I'd never feel like that would be something that would warrant a ban of some kind. While when we talk about deepfakes, I'm actually, I'm inclined to say, okay, yeah, that should definitely be banned, that should be removed from the internet. And I feel like it's really hard to place, for me personally, where I feel like that border is from that's okay and that's not okay and I can't really put my finger on it.

This statement highlights the struggle in how to justify one infringement on sexual agency compared to another. Expert 2 agreed with this thought process stating that the line for consent in regards to an identity online is extremely hard to navigate, especially in regard to the normalisation of pornographic content online. As mentioned by Goldfarb (2015), the pornification of content online has made it significantly easier to normalise the existence of sexual content to the point that the line is hard to distinguish between humour, creativity and a violation of someone's sexual agency.

However, some participants disagreed with these arguments, and the comparison Mike made, saying that sexual content such as fan-fiction or skits is based on the knowledge that people know it's fake, whereas deepfakes are done with the intention to confound truth and fiction. Expert 3 repeated this concern, stating that "the problem with deepfake is that when somebody sees a comical drawing, they know it's a comical drawing". At the same time, Felicity stated that if the deepfake was made with that same approach as skits, for example, it would change her perception of pornographic deepfake use as it would "lean over into one, like, kind of that meme territory. And two, it's like warning, this is not Taylor Swift. Do you know?". This was reiterated by the discussion in focus group 3, where the participants agreed that if the deepfake is done with humorous intent, it would change the perception of the deepfake, even if it was done non-consensually.

This was an interesting ethical distinction that the participants made, as they believe that if it were clear that the video was fake, it would change how individuals, or themselves, would react towards the content. Further arguing that it would be less real, and thus, less real in its impact and ethical and moral infringement. However, Expert 2

linked this back to the concept of hyperreality, stating that he believes that even in an educational or funny context, where it is clear that the content is fake, people will still believe it is real. He argued that “we can't unsee things, that we actually take part of that and actually embedded it into our minds.”

The differentiation between deepfakes and memes mentioned by Felicity was another discussion point made by three out of the five focus groups and two of the experts. All agreed that disseminating memes or viral content is also most prominently done without the consent of the individuals. However, it is regarded in a much different light than compared to deepfake pornography. Memes that were made of celebrities, in particular, were often classified under public use, especially when the images were taken from public broadcasting, as these images do not belong to the individual. Pablo explained, “It’s not like an image that was taken from you because you were giving your image to television or like to a specific network” (focus group 4). This was further highlighted by Expert 1, who also agreed that most people “in general already feel like these people are commodity” and that they are not viewed as real people. Previous literature has also highlighted that “if the person featured in the video is a politician, for example, we may think that elements of humiliation and satire are more permissible than if they were a private individual” (Öhman, 2022, p. 318). Therefore, when it came to ‘normal’ people, the participants were more inclined to discuss the non-consent behind memes and viral content, stating that it can have a great impact on people’s lives. Nonetheless, they always considered these much more harmless compared to deepfakes. For example, Mary from focus group 5 said, “people turn into memes without being asked, of course, and then they go viral or something like that. So that’s also super unconsensual, and it’s still like in the harmless domain in a sense.” Here, the participants agreed that the defining factor for the harm of nonconsent is the intention with which it is done.

This thought process is further strengthened by looking at the framework suggested by MacKinnon (1993). This frames sexual deepfakes as a planned attack on the sexual agency of an individual. This means that the intended use of sexual deepfakes is (sexual) infringement on an individual, and memes, although non-consensual, do not fall under this framework. Furthermore, as stated previously, deepfakes fall under a larger context of patriarchal power that changes the environment in which it is disseminated (Öhman, 2019). Graber-Mitchell (2021) elaborates on this statement by

arguing that illusory sex videos not only are an individual attack but "constitute a gendered attack against women in general that manifests itself in our social reality" (p. 8). Through this variation of intent and social environment, the participants ethically understand (non)consent differently in the online sphere when comparing deepfakes to other online dissemination forms. Still, there were a few participants who did argue that regardless of intent or not, the non-consensual dissemination of personal content is still the same. That "in reality, like legally, like if you don't think about feelings or about the topic, you can feel that they're kind of the same thing" (Beatrice, focus Group 4). She continued to argue that the ethics behind non-consensual sharing online should not be based on social and moral classifications but that consent, in general, should be a standard practice online.

With this, it can be seen that how deepfake pornography, and other created content, is viewed regarding consent is still based on personal ethics and moral conflicts and that a consensus is difficult to achieve.

4.4 Legal aspects of (non)consensual deepfake pornography

This last theme looks at the legal aspect of deepfake pornography and how this may impact consent. It specifically looks at autonomy and ownership of content online and how that intersects with how consent can be understood in the context of deepfakes. In addition, this sub-theme also looks at how participants discuss ideas of regulation in regard to deepfakes.

4.4.1 Autonomy of content and deepfake consent

When discussing the intersection of consent, deepfakes and data, the question of the autonomy of content emerged as a theme, as the ownership of content impacts consent. Especially in regards to deepfakes, there was a conflict in who was considered the owner of the content, with various factors impacting the perceptions. Ivan, focus group 3, explained the core of the problem the best:

> So who in some ways is it, do we base it off? Who's the punishable party in this? Or do we base this off who's the most affected? And if you then put in the explicit, does that change the context around it? Do we then go, okay, because of the

> nature of this image, does it belong to the person most affected, or does it belong to the person who owns the most of the image or the video?

This statement shows the complexity of the issue of content ownership in regard to deepfakes. Nevertheless, most of the arguments made throughout the various focus groups were split into two diverging opinions. Some argued that from a legal perspective, the creator of a piece of media is considered the owner, whereas others firmly disagreed, stating that you have the right to your own likeness, and thus the content belongs to the person implicated in the deepfake.

Those arguments that fell into the first category most commonly looked at the problem from a legal perspective, arguing that the act of creating content means that you are the owner of it. Specifically, the creator is the individual who has the ability to,

> actually share it to different platforms and, created it and has like, not a physical copy, but like a copy of this video [...] on his laptop or her laptop. Um, so in that case, they can do whatever they want with it. So technically, they would be the owners (Nora, focus group 5).

This argument is expanded on by Expert 2, as he discussed the concept of 'free use of media' and the act of creative freedom online, especially in regard to AI learning. Arguing that AI learning and creativity online can be compared to human learning:

> You can almost compare it to a human creating a painting of a celebrity. Like, okay, I'm gonna train on many images of this celebrity, and then I'm gonna paint with my own hands a copy of this person. Am I then using the images on the internet to, uh, to falsify somebody's image? Or am I recreating from all these learned images something new?

This comparison illustrates the complexity of the issue of ownership, as Expert 2 adds another layer by arguing that the images found online are "somewhat of free use",

and when creating a deepfake, he is not actually using the images but instead training his data through the AI. He further stated that the end result, similar to human learning, is based on the 'training' the AI has been given through the selected data set and could be defined under creative expression.

Previous studies agree with this line of argument made by Expert 2, stating that if a creative product made by an AI can pass an adapted Turing test called the "Lovelace Test", it means the artwork cannot be distinguished from human creativity (Schönberger, 2018). Thus, copyright infringement should also be based on that understanding of creativity (Schönberger, 2018). This means that, through the lens of human creativity, the creator could claim "copyright authorship of the newly created deepfake" (Meskys et al., 2020, p. 10) on the basis of it being new content.

However, both Expert 2 and the participants all agreed that from a moral standpoint, when discussing non-consensual deepfakes, this perspective changed. Stating that even if they still believed that the owner was the person who had the original video, morally, that content should not belong to them if it harms another individual.

The second group argued from this moral perspective first, stating that if your image is being used, it automatically belongs to you, or if not, it violates your image rights (Meskyks et al., 2020). However, they did note that physically having ownership over that content is impossible. Nevertheless, Pablo, focus group 4, argues, "It's not that you have ownership [...] but you do have to be asked consent, like this picture cannot be used by someone that is not you, without consent." So regardless of the ownership of the content, consent should always be at the forefront of interactions online. Furthermore, these participants argued that "given the fact that everybody involved in the making is consenting, I'd say it belongs to everyone" (Mike, focus group 4). Thus, the second group is more critical in the way deepfakes are defined as created content in regard to ownership and instead place the focus back on the act of consent.

Lastly, there was also a third group, mainly the participants from focus group 2, who believed that the deepfake does not "belong to anybody because the content, it's not real" (Sonia, focus group 2). Therefore, because content has been taken from various sources to create something new, it cannot belong to just one singular identity and thus should not belong to anyone. These participants argued that because the person shown in the deepfake does not exist and comprises a mix of individuals, no one can

have the right to it. A similar argument was made by Mirsky and Lee (2021) through the previously mentioned “identity leakage” (p.7), in which deepfakes blend facial or vocal features to the extent that identification of one individual is near to impossible. With this leakage, it also becomes extremely difficult on a legal level to place ownership onto a singular identity. Still, in legal and academic terms, the most significant arguments that have been made are not in favour of these participants. Principally because most previous studies on deepfake ownership place in either of the two preceding arguments.

Nevertheless, both in academics and the discussions presented, it can be identified that the question of content ownership in deepfakes is complex in its definition and in practice. This complexity will also impact the understanding of consent concerning deepfakes, as the dispute of individual consent is challenged by the ownership of content.

4.4.2 Regulation of deepfakes and consent

Considering the complexity of ownership of content when discussing deepfakes, it is no surprise that this impacted how participants felt about regulations of deepfakes. Many recognised the current lacking legal framework for deepfake content and realised this could also impact consent online. This is because if deepfakes can not be regulated, then an infringement on consent is also not regulated. However, at the same time, many participants doubted if there was a possibility to regulate deepfakes prior to them being made or if legal actions can only happen once they have been shared. Expert 1 elaborated on this;

> I don't think that you can tackle it beforehand. I think what irks me the most about all these kinds of things is that people are always trying to, uh, make sure that these things don't happen, but the truth is we're not able to establish things. [...] I think we should look into, okay, when this happens, then we have the right legislations, we have the right support, we have the right people who know about how to deal with this kind of things to help.

This ideology was echoed by some of the focus group participants, like Alina from focus group 1, who also believes that the issue with the existing legal framework for laws such as the 'revenge porn law' is the lack of aid after the TFSV. In regard to deepfake pornography, she stated that "it would be really difficult to prosecute someone for doing this because in the sense that, um, they are not using real footage of you".

The text by Öhman (2022) further agrees with this statement, elaborating that to prosecute the abuse of an individual's likeness in synthetic media, it depends on how successful the replication of the individual was. Considering there needs to be some form of qualitative identity to establish a connection between the content created and the individual. The question then, in order to create a legal framework, is based on "how this identity can be established. How can we know whether the face or voice featured in a piece of synthetic content belongs to a person who makes claim to it?" (Öhman, 2022, p. 318). This would add another layer of complexity to constructing a functioning legal framework.

However, even with the doubts presented, some participants argued that there are possible legal frameworks that might allow for a way of regulating deepfakes in an ethically sound way. Three main identified possible solutions were 1) an amendment to the GDPR in which photos and media as a form of data should also be protected under that law, which was proposed by Expert 3 and focus group 1. 2) A disclaimer or authentication label that identifies content as fake; this was discussed by Expert 2, focus groups 4 and 5. And 3) A form of blockchain or code that allows for content to be tagged as misused or authentic content, respectively, which was also proposed by Expert 2. Especially the verification aspect, through blockchains or code, has also been mentioned in previous literature, where it is argued that verification can "lead to larger cultures of consent" (Van der Nagel, 2020, p. 428). Moreover, a verification code also highlights the importance of consent as a fundament aspect of online sexual interactions.

Yet, all of these solutions, as predicted by Expert 1, are one's that do not necessarily aim to stop the misuse of data online but rather focus on the ability to distinguish between real and synthetic media. With this in mind, the participants acknowledge that legally, there is little that can be done to stop non-consent from happening online, but that the harm of the disseminated content may be decreased by having a framework for authentic content.

4.5 Connecting the themes

As mentioned previously, the nature of this research is highly interconnected within its different parts, and to answer the three sub-question, it is best to discuss these intersections.

The first sub-question, *"What factors impact how deepfake pornography and consent online are perceived?"* is answered by all four themes. These themes identified the various factors that changed how participants viewed consent through deepfakes. The first factor that impacted this perception was the idea of reality, which was touched upon by the digital and social aspects. Within the digital aspect, the ability to be able to create hyperreal fake content opened up the space for an increase in TFSV that can continuously create harm. The TFSV is further affected by the social understanding and awareness of consent online that has been influenced by hyperreality. This understanding was a key factor that the participants identified, as it means that individuals cannot differentiate between what is real or not, complicating the borders of consent.

Furthermore, the participants noted that through the creation of deepfakes, the legal and ethical understandings of consent are also impacted, considering identity online is now even more difficult to distinguish and define. This difficulty in the distinction of identity is linked to the ownership of content, thus, impacting consent, as ownership over the content is harder to claim. The dispute in ownership of the content for deepfakes makes it difficult for consent to exist, as it can be argued that no data has been explicitly taken from an individual, and thus no consent is actually needed.

The second sub-question, *"to what extent does deepfake pornography impact the understanding of digital autonomy? And how does this impact consent online?"* is answered mainly through the digital and legal themes. Here the dispute over ownership is mentioned again, as the inability to claim ownership over your identity through a deepfake removes digital autonomy and consent. In addition, the combining of the different actors involved further takes away from that autonomy over one's own existence online. Even considering this, the most drastic impact on digital autonomy through deepfakes is its influence on the self-understanding of the physical self because deepfakes can create an 'unseen third perspective'. This perspective can construct false memories that can create physical and psychological reactions to the deepfake, similar

to offline sexual violence. Through this, the creation of deepfakes, on a legal and a digital level, can actively disrupt the idea of digital autonomy.

The last sub-question, *"How does the context of deepfake pornography impact ethical understandings of consent online?"* is answered through the ethical aspect in which memes were used as a comparison for deepfakes. Memes, which were also identified as a form of non-consensual sharing, were considered less impactful compared to deepfakes, mainly due to the context in which it is shared. Participants agreed that the gendered issue of deepfake dissemination was a key factor in why this non-consent can be considered more harmful than memes. This was also highlighted in the social theme, in which participants stated that women are much more targeted and stigmatised regarding deepfakes, underlining the larger macro lens of gender inequality. This macro-lens was also used to justify other uses of deepfakes, such as humorous content or deepfakes with a disclaimer, as this would change the ethical perception of consent. Considering this, the gendered context of the non-consent of deepfake pornography actively changes ethical perceptions.

5. Conclusion

The current research has investigated how deepfake pornography impacts consent perceptions by conducting and analysing the responses made by the focus group participants and expert interviews. Within this concluding chapter, an answer to the research question is formulated, considering the academic discourse surrounding deepfakes and consent and the framework proposed in the second chapter. In addition, theoretical and social implications are specified and discussed. Lastly, possibilities for future research on deepfake and consent are presented.

5.1 Deepfake impact on consent

Overall, in order to answer the research question *"What are the implications of deepfake pornography on perceptions of consent online?"* the two main findings need to be highlighted. These are derived from the sub-conclusions made in the previous chapter that aided the formulation of an overall conclusion.

The first main finding is that deepfakes, as fake content, change the perceptions of how consent functions online. As mentioned, the distinction between authentic and fake content impacted how participants understood consent. This is because the ability to create hyperreal content (Baudrillard, 1981) leads to a warped social understanding of truth (Bandura, 2001; Paris & Donovan, 2019), which changes how individuals interact with content online, thus changing consent practices.

Moreover, fake content also changes how individuals perceive their own sexual autonomy online, as it can impact the understanding of one's own identity. This separation of identity from content existing in the digital sphere is, therefore, one of the fundamental distortions deepfake pornography can cause. Like the participants identified, it can cause a disruption between your physical and digital self, even going as far as impacting how one views themselves - through the ability to create a third perspective. This would, therefore, not only remove online consent but also the ability to have digital autonomy.

This anxiety becomes significant when addressing the fact that deepfakes can also increase TFSV (Powell & Henry, 2017), as deepfakes do not need any previous interaction with the individual placed in the deepfake. Through this, participants worried that because deepfake pornography can be created of any individual without the need for a prior relationship, consent does not need to exist at all in relation to

deepfakes. Furthermore, in connection to the accessibility of digital devices, participants fear that TFSV can now happen from anywhere, at any time and in a continuous manner without their preceding knowledge. This fear alone is a central implication of consent perceptions and the ability to 'exist' online, as the loss of control changes the dynamics created.

Hence, the lack of ability to distinguish between legitimate content and fabricated content is one of the principal factors that influenced the perceptions of consent online and the perception of digital autonomy, as well as one of the key identified differences between deepfakes and consent practices offline and other online consent exchanges.

The second main finding is the complexity that deepfakes create when it comes to ownership of content and how this impacts consent. This was a critical aspect of the findings, as it highlights the integral difference deepfakes create for consent online. The participants underlined this through the lack of consensus identified in the data. They struggled to define their legal and ethical standpoints on the content ownership, as content creation in other contexts is considered synonymous with ownership. However, in the context of pornographic deepfakes, this was often not considered the case. Even in regard to other uses of an individual's sexual identity, it was difficult for the participants to place it in the larger context of ethical media use, as deepfakes were considered intuitively unethical (Öhman, 2019). Moreover, the fact that deepfakes are based on data learning rather than the 'taking' of images complicated this distinction further, as it was related back to the similarity of human learning and, therefore, the legal boundaries of human creativity (Schönberger, 2018). Thus, this brought forth the legal and ethical complexities of deepfake use, which not only impacted how participants viewed deepfakes but also how they understood consent within this complexity. Consent cannot be defined alongside deepfakes if ownership and the right to one's own likeness are not systematically assumed and implemented. In terms of digital autonomy, this further highlighted the blurring understandings of a digital self and identity, which has clear consequences on how consent is understood and practised online.

In conclusion, this study indicates that deepfake pornography has a dire impact on the perceptions of consent online, as it diverges understandings of authentic content, digital identity and content ownership. Whilst changing the temporal and spatial

framework for consent that creates overall social anxiety around the usage of deepfake pornography.

5.2 Theoretical and Social Implications

The results of this research offer some theoretical implications and contributions to the topic of deepfake pornography and consent. Therefore, this section will aim to re-evaluate existing theories based on the findings to be able to validate, challenge or expand on them. This research's main theoretical implication was to bring together concepts as well as shifting the perspectives from legal understandings to user perceptions of deepfakes, which had yet to be explored.

Especially the concept of deepfake pornography combined with online consent was a new perspective in academic literature. Here, the findings confirmed previous studies on the two concepts and elaborated on them by combining them. An intersection that was highlighted by the findings of the study and the theory presented is that both in consent discussions and deepfake studies, gender plays a vital role (Öhman, 2019; Van der Nagel, 2020; Marcotte & Hille, 2021). Moreover, this intersection further elaborated that how consent is framed online changes drastically (Marcotte & Hille, 2021). The findings revised that notion by linking it to deepfake pornography which impacted that consent framework further, through concepts such as hyperreality (Baudrillard, 1981) and digital infinity (Marcotte & Hille, 2021). Therefore, the intersection of the concepts of deepfakes and consent allows for deeper insights into pre-existing research.

Specifically, the elaboration of the digital consent definition through the concept of hyperreality (Baudrillard, 1981) could be considered a critical theoretical implication. As these concepts have not yet been integrated in such a manner that respondents have been given the opportunity to discuss them. The results highlighted this, as they showed a new perspective on how consent practices online are perceived. A perspective that can be used for future frameworks for online consent - that includes consent on an individual level but also through a social level. This is because deepfakes, as a nonconsensual act, go beyond the victim and can also impact social understandings of reality in accordance with the consent.

In terms of social implications, this research can aid the awareness of not only the deepfake phenomena but the larger implication it can have on consent and societal understandings of authentic content. As the experts stated, one of the key issues to

consent in practice is the lack of discourse and understanding of consent society has, especially in regard to the online sphere. Therefore, if individuals would be more aware of consent online, it may impact how deepfakes exist in the digital world. Especially on a social level, if consent were understood more thoroughly, then the discussion would potentially shift away from the shaming of an individual who was harmed by such content and instead place the focus on the individual who did the sharing (Chun, 2016; Albury, 2017).

Furthermore, by applying the findings of this study to ethical and legal discussions of identity ownership online, it can elaborate on these by highlighting the complexities and nuances of ownership of content. The results showed that these nuances could have a great impact on how consent is perceived and practised online, as a lack of regulatory framework also means that there is no way to aid victims of deepfake TFSV (Garcia-Gomez, 2017; Powell & Henry, 2017). Hence, an aspect that could potentially improve social relations to deepfake consent, if considered.

Thus this research could aid the creation of a legislative framework that could potentially help victims of deepfake pornography to claim ownership of their own identity and thus remove the content from the online sphere. In addition, if more research is done on the psychological and physical reactions deepfakes can create, as presented in this paper, more aid can be provided for victims of TFSV. This is because, currently, society does not yet fully understand the impacts of non-consensual deepfakes, which in the researcher's opinion, is due to the lack of discourse surrounding deepfakes. However, this paper can aid NGOs such as the one led by Expert 1, to highlight this discrepancy.

5.3 Limitations and Recommendations for future research

The main limitation of this study lies in the sampling and use of video calling for the focus groups and expert interviews. Although an attempt was made to have a diverse sample both in nationalities, which was achieved, but also in gender, which was not achieved, most participants in the focus groups were female. Especially in regard to the topic, which can and is often noted as a gendered issue, this may have led to potentially biased outcomes. Furthermore, consent discourse can be understood differently between genders (Marcotte & Hille, 2021), and hence, this needs to be considered for future

research, as it implies that findings may vary if the same study is done using a different gender balance.

Another limitation was the use of online meetings instead of physical ones; this was a conscious decision as it opened up the opportunity to have a more diverse make-up of participants, as well as facilitate the execution of the focus groups, as individuals could join for an hour and did not need to take more time out of their day. However, it did create some technical issues, with slow wifi sometimes disrupting participants and one individual even having to drop out minutes before a scheduled focus group. This individual had a technical issue with her computer and therefore did not manage to join. Nevertheless, considering the other four participants were already on the call, the decision was made to continue, and the discussion was still extremely fruitful and useful in answering the research question.

To the best of the researcher's knowledge, this study represents the first investigation of social perceptions surrounding deepfake pornography and consent, especially by using a qualitative mix-method approach. Considering this, it is important, in order to gain more insights into the phenomenon, to conduct further research on the same topic. The findings presented in this research can therefore be a beneficial starting point for future studies. Specifically, the complexities of consent in regard to content ownership is an aspect that should be explored further, as this has been identified as the key aspect that diverges deepfakes from other forms of (non)consensual interactions online. Moreover, the social perception and implications of deepfakes on the idea of authentic content being disrupted by deepfakes is another research area that requires further insights. Lastly, considering the target audience was quite young, it would be interesting to explore the perceptions of older generations to investigate if their reaction to the existence of deepfakes varies compared to Generation Z and Millenials. Especially considering that the younger generations tend to be more versed in new technology and so may be more inclined to regard deepfakes in a critically balanced manner, whereas the older generation may have stronger anxieties towards such technologies. In general, since deepfake pornography is a novel phenomenon that has not yet received much attention from a media and user perception standpoint, further research is highly encouraged.

References

Albury, K. (2017). Just because it's public doesn't mean it's any of your business: Adults' and children's sexual rights in digitally mediated spaces. *New Media & Society*, *19*(5), 713-725. 10.1177/1461444816686322

Anthony, S. (2012). Just how big are porn sites? *Extreme Tech*. Retrieved from http://www.extremetech.com/computing/123929-just-how-big-are-porn-sites

Attwood, F. (2005). What do people do with porn? Qualitative research into the comsumption, use, and experience of pornography and other sexually explicit media. *Sexuality and culture*, *9*(2), 65-86.

Bakker, P., & Taalas, S. (2007). The irresistible rise of porn: The untold story of a global industry. *Observatorio (OBS*)*, *1*(1), 99-118.

Balsamo, A. M. (1996). *Technologies of the gendered body: Reading cyborg women*. Duke University Press.

Bandura, A. (2001). Social cognitive theory of mass communication. *Media psychology*, *3*(3), 265-299. 10.1207/S1532785XMEP0303_03

Bancroft, J. (Ed.). (1997). *Researching sexual behavior: Methodological issues* (Vol. 5). Indiana University Press.

Bates, S. (2017). Revenge porn and mental health: A qualitative analysis of the mental health effects of revenge porn on female survivors. *Feminist Criminology*, *12*(1), 22-42. 10.1177/1557085116654565

Baudrillard, J. (2019). Simulacra and simulations (1981). In *Crime and Media* (pp. 69-85). Routledge.

Belli, L. (2018). Network self-determination: When building the Internet becomes a right. *IETF Journal*, 107.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77–101. 10.1191/1478088706qp063oa

Boeije, H. (2010). "Principles of qualitative analysis" (Chapter 5) and "Doing qualitative analysis" (Chapter 6). In *Analysis in qualitative research* (pp. 75–121).

Burgin, R. (2019). Persistent narratives of force and resistance: Affirmative consent as law reform. *The British Journal of Criminology*, *59*(2), 296-314. 10.1093/bjc/azy043

Chun, W. H. K. (2016). *Updating to remain the same: Habitual new media*. MIT press.

Cover, R. (2022). Deepfake culture: the emergence of audio-video deception as an object of social anxiety and regulation. *Continuum*, *36*(4), 609-621. 10.1080/10304312.2022.2084039

Drenten, J., Gurrieri, L., & Tyler, M. (2020). Sexualized labour in digital culture: Instagram influencers, porn chic and the monetization of attention. *Gender, Work & Organization*, *27*(1), 41-66. 10.1111/gwao.12354

Döring, N., & Mohseni, M. R. (2018). Are online sexual activities and sexting good for adults' sexual well-being? Results from a national online survey. *International Journal of Sexual Health*, *30*(3), 250-263.10.1080/19317611.2018.1491921

Elmer-Dewitt, P., & Bloch, H. (1995). On a screen near you: Cyberporn. *Time*, *146*(1), 38-45.

Fisher, W. A., & Barak, A. (2001). Internet pornography: A social psychological perspective on Internet sexuality. *Journal of Sex Research*, *38*(4), 312-323. 10.1080/00224490109552102

Fisher, W. A., & Kohut, T. (2020). Reading pornography: Methodological considerations in evaluating pornography research. *The Journal of Sexual Medicine*, *17*(2), 195-209. 10.1016/j.jsxm.2019.11.257

Fisher, W. A., Montgomery-Graham, S., & Kohut, T. (2019). Pornography problems due to moral incongruence. *Archives of Sexual Behavior*, *48*, 425-429. 10.1007/s10508-018-1291-7

Flood, M. (2009). The harms of pornography exposure among children and young people. *Child Abuse Review: Journal of the British Association for the Study and Prevention of Child Abuse and Neglect*, *18*(6), 384-400.10.1002/car.1092

Foucault, M. (1998). The will to knowledge: The history of sexuality vol. I.

García-Gómez, A. (2017). Teen girls and sexual agency: Exploring the intrapersonal and intergroup dimensions of sexting. *Media, Culture & Society*, *39*(3), 391-407. 10.1177/0163443716683789

Gernert, J. (2010). *Generation Porno*. Fackelträger-Verl..

Gill, R., & Orgad, S. (2018). The shifting terrain of sex and power: From the 'sexualization of culture'to# MeToo. *Sexualities*, *21*(8), 1313-1324.

Goldfarb, L. (2015). Pornification. *The International Encyclopedia of Human Sexuality*, 861-1042. 10.1002/9781118896877.wbiehs354

Gosse, C., & Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communication*, *37*(5), 497-511.

Graber-Mitchell, N. (2021). Artificial Illusions: Deepfakes as Speech.

Hald, G. M., & Malamuth, N. M. (2008). Self-perceived effects of pornography consumption. *Archives of sexual behavior*, *37*(4), 614-625. 10.1007/s10508-007-9212-1

Hald, G. M., Malamuth, N. N., & Lange, T. (2013). Pornography and sexist attitudes among heterosexuals. *Journal of Communication*, *63*(4), 638-660. 10.1111/jcom.12037

Hancock, J. T., & Bailenson, J. N. (2021). The social impact of deepfakes. *Cyberpsychology, behavior, and social networking*, *24*(3), 149-152.

Harris, K. R. (2021). Video on demand: what deepfakes do and how they harm. *Synthese*, *199*(5-6), 13373-13391.

Hawkins, G., & Zimring, F. E. (1988). *Pornography in a free society*. Cambridge University Press.

Henry, N., & Powell, A. (2016). Sexual violence in the digital age: The scope and limits of criminal law. *Social & legal studies*, *25*(4), 397-418. 10.1177/0964663915624273

Hickman, S. E., & Muehlenhard, C. L. (1999). "By the semi‐mystical appearance of a condom": How young women and men communicate sexual consent in heterosexual situations. *Journal of Sex Research*, *36*(3), 258-272. 10.1080/00224499909551996

Johnson, J. M. (2011). In-depth interviewing. In Jaber F. Gubrium & James A. Holstein (Eds.). Handbook of Interview Research (pp. 103–119). Thousand Oaks, CA: SAGE Publications, Inc. 10.4135/9781412973588

Johnson, N. L., & Johnson, D. M. (2021). An empirical exploration into the measurement of rape culture. *Journal of Interpersonal Violence*, *36*(1-2), NP70-NP95.

Kirchengast, T. (2020). Deepfakes and image manipulation: criminalisation and control. *Information & Communications Technology Law*, *29*(3), 308-323. 10.1080/13600834.2020.1794615

Lim, M. S., Carrotte, E. R., & Hellard, M. E. (2016). The impact of pornography on gender-based violence, sexual health and well-being: what do we know? *J Epidemiol Community Health*, *70*(1), 3-5.

Lindgren, J. (1993). Defining pornography. *University of Pennsylvania Law Review*, *141*(4), 1153-1275.

Lo, V. H., Neilan, E., Sun, M. P., & Chiang, S. I. (1999). Exposure of Taiwanese adolescents to pornographic media and its impact on sexual attitudes and behaviour. *Asian Journal of Communication*, *9*(1), 50-71. 10.1080/01292989909359614

Löfgren-Mårtenson, L., & Månsson, S. A. (2010). Lust, love, and life: A qualitative study of Swedish adolescents' perceptions and experiences with pornography. *Journal of sex research*, *47*(6), 568-579. 10.1080/00224490903151374

MacKinnon, C. A. (1983). Complicity: An Introduction to Andrea Dworkin, Abortion, Chapter 3, Right Wing Women-New York: Perigee, 1983. *Law & Ineq.*, *1*, 89.

MacKinnon, C. A. (1993). *Only words*. Harvard University Press.

Maddocks, S. (2020). 'A Deepfake Porn Plot Intended to Silence Me': exploring continuities between pornographic and 'political'deep fakes. *Porn Studies*, *7*(4), 415-423.

Makin, D. A., & Morczek, A. L. (2015). The dark side of internet searches: A macro level assessment of rape culture. *International Journal of Cyber Criminology*, *9*(1), 1.

Marcotte, A. S., & Hille, J. J. (2021). Sexual Violence and Consent in the Digital Age. *The Palgrave Handbook of Gendered Violence and Technology*, 319-332.

Mayer, M., & Lu, Y. (2022). Digital Autonomy? Measuring the Global Digital Dependence Structure.

McGlynn, C., Rackley, E., & Houghton, R. (2017). Beyond 'revenge porn': The continuum of image-based sexual abuse. *Feminist legal studies*, *25*, 25-46. 10.1007/s10691-017-9343-2

McLuhan, M. (1964). 1996 Understanding Media. *The Extensions of Man*.

McNair, B. (2014). Rethinking the effects paradigm in porn studies. *Porn studies*, *1*(1-2), 161-171.

Mendes, K., Ringrose, J., & Keller, J. (2018). # MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236-246.

Meskys, E., Kalpokiene, J., Jurcys, P., & Liaudanskas, A. (2020). Regulating deep fakes: legal and ethical considerations. *Journal of Intellectual Property Law & Practice*, *15*(1), 24-31.

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys (CSUR)*, *54*(1), 1-41. 10.1145/3425780

Morgan, D. L. (1998). *The focus group guidebook*. Thousand Oaks, CA: Sage.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative health research*, *8*(3), 362-376. 10.1177/104973239800800307

Muehlenhard, C. L., Humphreys, T. P., Jozkowski, K. N., & Peterson, Z. D. (2016). The complexities of sexual consent among college students: A conceptual and empirical review. *The Journal of Sex Research*, *53*(4-5), 457-487. 10.1080/00224499.2016.1146651

Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K. (2023). Deepfakes: Deceptions, mitigations, and opportunities. *Journal of Business Research*, *154*, 113368.

Öhman, C. (2019). Introducing the pervert's dilemma: a contribution to the critique of Deepfake Pornography. *Ethics and Information Technology*, *22*(2), 133-140. 10.1007/s10676-019-09522-1

Öhman, C. (2022). The identification game: deepfakes and the epistemic limits of identity. *Synthese*, *200*(4), 319. 10.1007/s11229-022-03798-5

Paasonen, S. (2016). Pornification and the Mainstreaming of Sex. In *Oxford Research Encyclopedia of Criminology and Criminal Justice*. 10.1093/acrefore/9780190264079.013.159

Paris, B., & Donovan, J. (2019). Deepfakes and cheap fakes.

Patella-Rey, P. J. (2021). *Cam Modeling: Labor, Intimacy, and Prosumption* (Doctoral dissertation, University of Maryland, College Park).

Plummer-D'Amato, P. (2008). Focus group methodology Part 1: Considerations for design. *International Journal of Therapy and Rehabilitation*, *15*(2), 69-73. 10.12968/ijtr.2008.15.2.28189

Popova, M. (2020). Reading out of context: pornographic deepfakes, celebrity and intimacy. Porn Studies, 7(4), 367-381.

Powell, A., Henry, N., Flynn, A., & Scott, A. J. (2019). Image-based sexual abuse: The extent, nature, and predictors of perpetration in a community sample of Australian residents. *Computers in Human Behavior*, *92*, 393-402. 10.1016/j.chb.2018.11.009

Powell, R. A., & Single, H. M. (1996). Focus groups. *International journal for quality in health care*, *8*(5), 499-504. 10.1093/intqhc/8.5.499

Powell, A., Flynn, A., & Sugiura, L. (Eds.). (2021). The palgrave handbook of gendered violence and technology. Palgrave Macmillan. 10.1007/978-3-030-83734-1

Quadara, A., El-Murr, A., & Latham, J. (2017). *The effects of pornography on children and young people: An evidence scan*. Australian Institute of Family Studies.

Rentschler, C. A. (2014). Rape culture and the feminist politics of social media. *Girlhood studies*, *7*(1), 65-82.

Saunders, R. (2019). Computer-generated pornography and convergence: Animation and algorithms as new digital desire. *Convergence*, *25*(2), 241-259. 10.1177/1354856519833591

Schönberger, D. (2018). Deep copyright: up-and downstream questions related to artificial intelligence (AI) and machine learning (ML). *SCHÖNBERGER Daniel, Deep Copyright: Up-and Downstream-Questions Related to Artificial Intelligence (AI) and Machine Learning (ML) in Droit d'auteur*, *4*, 145-173.

Schreier, M. (2013). Qualitative content analysis. The SAGE Handbook of Qualitative Data Analysis, pp. 170–183. 10.4135/9781446282243.n12

Sundén, J. (2020). Play, secrecy and consent: Theorizing privacy breaches and sensitive data in the world of networked sex toys. *Sexualities*. 10.1177/1363460720957578

Toffler, A. (1980). The third wave/alvin toffler. *New York: Morrow*, *544*.

van der Nagel, E. (2020). Verifying images: Deepfakes, control, and consent. *Porn Studies*, *7*(4), 424-429.

Wagner, T. L., & Blewer, A. (2019). “The word real is no longer real”: Deepfakes, gender, and the challenges of ai-altered video. *Open Information Science*, *3*(1), 32-46. 10.1515/opis-2019-0003

Appendix A - Consent form

Information and consent form

Deepfake Pornography and its Implications

Introduction

I am Xenia, and I am currently studying at Erasmus University Rotterdam. I am conducting some research on deepfake pornography and its implications. Deepfake pornography is a form of Artificial Intelligence where one's face is swapped with someone else. Due to the nature of the topic, some discussions surrounding consent, Sexual Violence and harassment may come, which is important for you to know prior to beginning this focus group.

I will explain the study below. If you have any questions, please ask me. While reading, you can mark parts of the text that are unclear to you.

If you want to participate in the study, you can indicate this at the end of this form.

What is the research about?

The Research is about the implications of deepfake pornography, an AI technology that swaps out one face for another. It will look at themes of consent, the autonomy of content and sexual violence online.

Why are we asking you to participate?

I ask you to participate because your opinion will help me learn about what and how people perceive this new form of technology.

What can you expect?

The study lasts approximately 1 hour.
If you participate in this study, you will take part in:

A focus group:

This is a discussion with 4 other people with the same experiences as you. You can decide what you want to share with us. The discussion will take place in Online on XXX. In this discussion the researcher will ask some questions and/or give some scenarios that have the purpose to spark a discussion within the group, you may participate at any point you wish. The topics will focus around deepfake pornography, and in that sense will discuss sex, consent, pornography, Artificial Intelligence, and Sexual Violence. After the discussion, you have the choice to stay longer for a debrief on the insights of the topic, and after the research has been conducted you can have access to a follow-up.

I will make an audio recording of the discussion.

You decide whether to participate

Participation in this study is completely voluntary. You can stop at any time and would not need to provide any explanation.

What are the potential risks and discomforts?

- During the interview, personal questions will be asked about potentially upsetting events you may have experienced. These may trigger unpleasant memories and emotions.
- If you need emotional support or a listening ear after the interview, you can contact Slachtofferhulp Nederland. They can also help you determine which further assistance may be needed and refer you on. You can contact them at 0900-0101 or via the chat on their website: www.slachtofferhulp.nl.

What data will I ask you to provide?

I will store your data so that I can be in contact with you. For the study, I will also need other data from you.

During focus group, the following personal data will be collected from you: Name, age, gender, audio recordings, sentiments about / feelings about / opinions about the topic,

In addition, it is also possible that you will talk about your political affiliation and religious or philosophical beliefs and those of others, as these may also relate to your opinion about Pornography.

I also need your email address, to send the results of the study to you by email.

Who can see your data?

- I store all your data securely.
- Only persons involved in the research can see (some of) the data. Only the principal investigator has access to your data such as your name, gender or focus group information.
- Recordings are transcribed. Your name is replaced with a number/made-up name.
- Data such as your name, age, and recordings will be stored separately from your answers/the transcription.
- I will write an article about the results of the study which will be published. The results will be accessible by anyone.
- We may use your specific answers in the article. If your answer can be traced to you or we would like to mention your name, we will ask your permission first.

Because you are participating in a group discussion, you should realize that the other participants will also hear your opinion about deepfake pornography. We ask all participants not to talk to people outside the group about what was said in the group.

How long will your personal data be stored?

Your data will be retained for a minimum of 10 years. We retain the data so that other researchers have the opportunity to verify that the research was conducted correctly.

Using your data for new research
(Part of) the data we collect may be useful in anonymized form, for example for educational purposes and future research, including in very different research areas. We will make the data publicly available after proper anonymization. We ensure that the data cannot be traced back to you/we do not disclose anything that identifies you.

What happens with the results of the study?

You may indicate if you would like to receive the results.

Do you have questions about the study?

If you have any questions about the study or your privacy rights, such as accessing, changing, deleting, or updating your data, please contact me.

Email: 539124xl@eur.student.nl

Do you have a complaint or concerns about your privacy? Please email the Data Protection Officer (fg@eur.nl) or visit www.autoriteitpersoonsgegevens.nl. (T: 088 - 1805250)

Do you regret your participation?

You may regret your participation. Even after participating, you can still stop. Please indicate this by contacting me. I will delete your data. Sometimes we need to keep your data so that, for example, the integrity of the study can be checked.

Declaration of Consent - This will be done verbally before the meeting officially starts

I have read the information letter. I understand what the study is about and what data will be collected from me. I was able to ask questions as well. My questions were adequately answered. I know that I am allowed to stop at any time.

By agreeing to this form, I

1. consent to participate in this research.
2. consent to the use of my personal data;
3. confirm that I am at least 18 years old.
4. understand that participating in this research is completely voluntary and that I can stop at any time;
5. understand that my data will be anonymised for publication, educational purposes and further research.
6. understand that the meeting will be voice recorded,
7. understand answers given in the focus group are stored and used for educational purposes and for future research, and also in other areas of research than this research.

This copy of the consent form is for you to keep

Appendix B - Topic Guide

These were the questions used to guide the focus group

General Introduction (done before starting to record)

Hi everyone, thank you for taking the time to join today, I very much appreciate your help. Before we begin I will quickly give an overview of how this goes and then we will get started. So the topic is about deepfakes which is a form of deep learning technology that superimposes one person's face over another person's body to make it seem like it is one singular person. It can be used in many different ways but is mostly found in pornography, which is also where our focus will be. But basically I will ask you some questions that you can discuss in the group, there are no right or wrong answers - because the discussion topics are based on your own feelings. Also do not feel like you need to ask for permission to talk or ask questions, or not answer at all, thats also okay. Basically join or discuss whenever you feel comfortable. Lastly, I would like for you all to introduce yourselves to the rest of the group and also state if you agree to the terms mentioned in the consent form - such as the audio recording of this discussion, which I would start now.

Warm up questions

- Before today have you ever heard of deepfakes? if yes in what way?
- What are you general thoughts on deepfakes?
- Have you ever interacted with sexually explicit images online? or distributed them yourselves?
- Have you ever received, or if you want to disclose, sent a nude or taken one, and if you have received one, what was your reaction? for example, did you like it?
- when you receive these kinds of pictures do you feel ownership over these pictures once they have been sent to you? why or why not?

Digital aspect

- Do you think technologies such as deepfakes has an impact on how we see sex and pornography?
- Do you think consent is different online versus offline? do you think this changes with deepfakes?
- What do you think has led to the technology we now see in pornography? such as deepfakes?
- Considering ownership over content, do you think that this changes online? or for example, let's say on a private Instagram account or a public Instagram account?
- How do you think deepfakes impact understandings of consent, privacy and autonomy of content?

Social aspect

- What impact do you think deepfake pornography can have on a societal but also individual level?
- Do you think deepfakes can impact how consent is understood?
- Where would you place, deepfakes into the conversation of sexual violence? Like, is deep fake a form of sexual violence?
- How would you feel if pornographic deepfake content was made of you?
- How do you think the lack of consent in the creation and distribution of deepfake pornography affects the victim's mental health and wellbeing?
- How can society change its attitudes towards pornography to better prioritize consent and prevent the creation and spread of non-consensual deepfake pornography?

Ethical aspect

- What ethical implications do you think deepfakes have?
- Where would you draw the line between unconsensual sharing online taking things such as memes or deepfake porn as an example?
- Do you think there are ethical ways to make deepfake pornography?
- if it is done consensually, do you think there are positive sides too?
- what are some ethical concerns that you belive are created by deepfakes?
- How can context change the ethics behind deepfake pornography?

Legal aspect

- Do you believe that using a person's likeness in deepfake pornography without their consent is a violation of their rights? Why or why not?
- When it comes to deepfake pornography who do you belive has the right to that content? How do you think this could impact consent?
- What legal implications do you think deepfakes have?
- What implications do you think deepfake technology has on other non-consensual purposes, such as political propaganda or revenge porn?
- What type of legal measures do you think would be appropriate for people who create and publish deepfakes?
- What steps can be taken to ensure that the creation and distribution of deepfake pornography is done with the explicit consent of all involved parties?

Any other comments/ questions?

Appendix C - Coding tree

Theme	Selective	Axial	Open
DIGITAL ASPECT	Digital Infinity and deepfake consent	Lack of bi-directional consent online	One-sided consent, no control, bi-directional consent, digital infinity,
		Continous sexual violence	content online difficult to remove, re-traumatisation, ability to be seen by millions, long-term impact, continuous sexual violence
	Consent and the ability to create fake content	No digital self	data concerns, infringement on data, digital identity impacted, loss of identity control online
		Sexual Violence more accessible	easier to do TFSV, no need for real content, can happen to anyone, fear for worse TFSV, ability to create harmful content
		Impact on self-understanding	Hyperreality, unseen third perspective, memory creating, changing understanding of yourself, self-shame, depicition of false reality, psychological impact

SOCIAL ASPECT	Awareness of deepfake consent	Changing social understanding of consent online	Consent online vs offline, social definition of consent, third-parties involved in consent, different framework for consent, blurring lines for consent
		Lack of ability to discern true versus fake	authentic content, truth versus fiction, media effect, blurring lines for content,
		Deepfake and social consent	Exposure to fake content, implications on social consent, impact on interaction online, unwilling interaction with non-consent, non-consent online
	Social stigma and deepfake impacts	Social aspect of sexual violence	Social shaming, social context, harm through social aspect, stigma, negative view on porn, association to porn, harm on reputation
		Gendered social reaction	Women impacted more, gender inequality, commodification of women, sexualisation of women, social stigma against women
		User is the problem	Deepfake tech not the issue, larger social issue, users the problem

ETHICAL ASPECT	Ethical conflicts regarding deepfakes	Context impacts ethics	ethical dilemmas, ethical context, ethical viewpoint
		Fringecase: Peadophilia	peadophilia and deepfakes, children in deepfakes, AI mushing, ethical concern: peadophilia
	Other forms of non-consensual sharing online	Humour changes ethical perception	Humorous context, clear connection to reality, memes, use of sexual identity
		Gendered ethical issue	unethical because of gender inequality, deepfake as gender issue,
LEGAL ASPECT	Deepfakes and ownership of content	Ownership through content creation	Content creation equals ownership, ownership and consent, copyright of sex worker
		Ownership and identity	depiction of identity, right to own identity, identity infrignment,
	Regulations for deepfake pornography	Lack of legislative framework	Lack of legislation, difficult to regulate deepfakes, legal aspect of gender issue
		Concerns for possible frameworks	Possible soultions for deepfakes, GDPR, deepfake disclaimers, concern for legislations