"Violation of my body:" Perceptions of AI-generated non-consensual (intimate) imagery

Grace Brigham
University of Washington

Miranda Wei
University of Washington

Tadayoshi Kohno
University of Washington

Elisa M. Redmiles
Georgetown University

Abstract

AI technology has enabled the creation of deepfakes: hyper-realistic synthetic media. We surveyed 315 individuals in the U.S. on their views regarding the hypothetical non-consensual creation of deepfakes depicting them, including deepfakes portraying sexual acts. Respondents indicated strong opposition to creating and, even more so, sharing non-consensually created synthetic content, especially if that content depicts a sexual act. However, seeking out such content appeared more acceptable to some respondents. Attitudes around acceptability varied further based on the hypothetical creator's relationship to the participant, the respondent's gender and their attitudes towards sexual consent. This study provides initial insight into public perspectives of a growing threat and highlights the need for further research to inform social norms as well as ongoing policy conversations and technical developments in generative AI.

1 Introduction

Technological advancements in artificial intelligence (AI) have enabled the creation of hyper-realistic synthetic media known as "deepfakes." This term, a portmanteau of "deep learning" and "fake," refers to synthetic image, audio, or video representations of individuals that has been automatically generated using machine learning [31, 49, 88]. Deepfakes encompass many forms of media synthesis, including voice-swapping, text-to-speech, face-swapping, face-morphing, full-body puppetry, and lip syncing [49]. Moreover, recent progress in generative AI has enabled the creation of deepfakes using only text prompts, rather than requiring a data set of training images depicting the target individual [53, 72, 92]. While deepfake technology has potentially benevolent applications in creativity, accessibility, and entertainment [12, 19, 30, 31, 88], it has also been used to spread disinformation, commit fraud (e.g., phishing), and non-consensually generate intimate imagery [2, 14, 20].[1] The latter has commonly been termed "deepfake pornography," but following evolving terminology around image-based sexual abuse [58], we refer to it in this paper AI-generated non-consensual intimate imagery (AIG-NCII).[2]

Current technical research around deepfakes has predominantly focused on developing generative AI systems capable of synthesizing such content, including face-swapping [64,96] and text-to-video systems [43, 78, 95], detection methods [11, 22, 97], as well as strategies to disrupt their generation [76]. However, research on attitudes of the general public towards deepfakes is far more nascent. A large body of literature and theory in information systems and HCI has underscored the importance technology acceptance — by individuals and by society — on technology use (and misuse) [46, 54]. Thus, this research seeks to bridge the gap between the technically possible (e.g., the academic research cited above) and the public acceptance of different uses of the technology. As computer security and privacy researchers, we are particularly interested in adversarial contexts, e.g., the generation of AIG-NCII. Hence, we ask: **What are people's attitudes toward the hypothetical non-consensual creation, sharing, and/or seeking out of deepfakes depicting them?** Decomposing this question, we ask specifically:

RQ1: How do attitudes differ depending on what is depicted: AIG-NCII vs. non-consensually created content depicting *non-sexual* acts?

USENIX Symposium on Usable Privacy and Security (SOUPS) 2024.
August 11–13, 2024, Philadelphia, PA, United States.

[1] Intimate imagery refers to "images and videos of people who are naked, showing their genitals, engaging in sexual activity or poses, or wearing underwear in compromising positions" [80].

[2] AIG-NCII is our preferred term because it emphasizes the non-consensual nature of the images and is more widely applicable to the range of technologies that can be used to create such images.

RQ2: How do these attitudes differ depending on contextual factors: who is creating the media and for what purpose?

RQ3: How do attitudes related to sexual (a) consent and (b) content influence these attitudes?

RQ4: How does gender influence these attitudes?

To answer these questions, we conducted a vignette survey of 315 individuals to assess attitudes towards different situations involving non-consensual synthetic media. This research elucidates contextual and individual factors that shape public acceptance of generative AI technology being used to construct deepfakes in addition to broader trends in attitudes and rationales. Through this work, we aim to inform future discourse regarding deepfakes, specifically AIG-NCII, in public, technical, legal, and policy spheres.

2 Background & Related Work

In 2017, a user named "deepfakes" posted synthetic videos of celebrities in sexual acts to Reddit [31, 49, 51]. Over 90,000 users subsequently joined an r/deepfake subreddit for creating and sharing similar content, drawing significant public attention before being banned by Reddit as "involuntary pornography" [73]. Online communities catalyzed the popular use of the term "deepfake" [31, 49, 51], and despite bans on mainstream social media platforms, AIG-NCII continues to be produced and circulated on dedicated forums [2, 84].

Image-Based Sexual Abuse (IBSA). AIG-NCII is one form of IBSA: the non-consensual creation, distribution, or threats made with intimate images [56, 57, 77]. Victim-survivors of IBSA often experience severe health consequences, such as post-traumatic stress disorder, anxiety, depression, and greater somatic burdens [7, 25, 44, 77]. IBSA harms are also social, e.g., isolation, lowered self-esteem, trust issues, and unhealthy coping mechanisms [7, 56]. Victim-blaming attitudes are prevalent when seeking support or justice after IBSA [33], and obstruct help-seeking [59, 66].

IBSA falls under a broader umbrella of technology-facilitated gender-based violence [23]. As with other gender-based violence, victim-survivors of AIG-NCII are predominantly, though not exclusively, women [2, 24].

IBSA and AIG-NCII are growing global issues [34]. Policy on IBSA is sparse in most countries [1, 91]; in the US specifically, legal scholars have called for legislation to sufficiently address its harms [16, 21, 37]. Understanding public attitudes about synthetic media, specifically AIG-NCII, can inform better policies on this emergent form of IBSA.

Public attitudes about AIG-NCII. Early research on non-sexual deepfakes found significant public concern about its creation and dissemination [39], but less if deepfakes are created for entertainment, humor, or with consent and traceability [52, 63]. Those who are more concerned about deepfakes are also less likely to feel individual responsibility [63].

Regarding AIG-NCII, i.e., sexual deepfakes, prior work has primarily focused on attitudes around criminality and perceived harm to victim-survivors [32, 51, 85]. Kugler and Pace found that individuals in the UK perceived significant harms from and strongly favored criminalization of sexual and non-sexual deepfakes [51]. Further, videos being labeled as fake did reduce the perceived harm of non-sexual deepfakes, but did not for AIG-NCII [51]. Fido et al. study AIG-NCII while varying the identity of the target, finding that deepfakes of celebrities were perceived as less criminal and less harmful, especially for celebrities who are men [32]. This work also found that creation of deepfakes for personal sexual gratification was viewed as less harmful and criminal than sharing. Finally, in Umbach et al.'s study across ten countries, awareness of AIG-NCII was low overall, but surveyed individuals believed victims had a right to be upset [85]. Men in this study also reported more perpetration and victimization.

We combine elements from prior work on non-sexual deepfakes and AIG-NCII to systematically study *acceptance* (vs. criminality or harm) of the use of generative AI technology to create different types of deepfakes. Specifically, we extend [51] to compare AIG-NCII with not-exclusively-harmful deepfake actions (RQ1): saying something – which is ambiguous regarding sexuality or harmfulness – and playing a sport – ostensibly, a neutral action. We make these comparisons across five disambiguated actions involving deepfakes: creating, private sharing, public sharing, resharing, and seeking out. Additionally, we explore the role of contextual factors (RQ2) such as *intent* of the creator; a factor not explored in prior work on AIG-NCII despite the fact that intent is a factor in existing laws that can be applied to deepfakes and image-based sexual abuse [17] and the fact that prior work on non-sexual deepfakes finds that intent affects the general public's attitudes toward acceptability [52, 63]. As a second contextual factor, we further explore the relationship between the creator and subject; we explore the role of intimate partnership while prior work explored, and found relevant, celebrity status [32]. We further explore the impact of individual factors on these attitudes. We select individual factors found relevant in prior work on offline sexual abuse such as sexual consent attitudes [45] but which have been unexplored in the context of deepfakes and AIG-NCII (RQ3); as well as individual factors found relevant in prior work on AIG-NCII criminality perceptions such as gender [85] (RQ4).

Finally, as noted by Fido et al. [32], prior work lacks qualitative exploration of *why* respondents held particular opinions. In our work, we collect and rigorously analyze qualitative data on attitudes toward the acceptability of creating AIG-NCII and other synthetic media.

Deepfake community attitudes. Research has also examined attitudes toward deepfake tech, revealing pro-deepfake

views among Reddit users [36] and on MrDeepFakes, the self-proclaimed largest deepfake platform [84]. Another study of a deepfake tool's open-source community found positive attitudes but also heightened misuse concerns [90].

3 Methodology

We conducted a survey of 315 U.S. Prolific respondents (full survey instrument provided in Appendix A). Our Institutional Review Board (IRB) found our study to be exempt and we followed the ethical considerations as described in Section 3.4.

3.1 Survey structure

Consent. The survey began with a description of generative AI and its capacity to generate realistic-looking but fake images. We chose to avoid using "deepfake" given potential priming effects (e.g., about political disinformation). Respondents then were told survey structure and asked to consent.

Vignettes. We used vignettes—short descriptions of hypothetical scenarios—to solicit respondents' attitudes about AIG-NCII. Vignettes are common in security and privacy studies to elicit reactions [28, 55, 62] and can approximate real-world behaviors [41]. Each vignette described generative AI being used to create a video of the respondent without their knowledge, varying three factors:

(1) action varies sexual explicitness, from unambiguously sexual behavior ('performing a sexual act') to non-sexual ('playing a sport') to ambiguous ('saying something'). This factor corresponds to RQ1.
(2) creator varies the relationship between the media maker and participant, either 'an intimate partner' or 'a stranger.' This corresponds to RQ2 and complements prior work [32, 51] exploring other relationships (e.g., of a celebrity).
(3) intent varies the creator's motivation, representing motivations reported by prior work [31, 88]: 'harming you,' 'entertainment,' and 'sexual pleasure,' also corresponding to RQ2.

We employed a 2 (creator) × 3 (action) × 3 (intent) full-factorial design to construct 18 vignettes (see Table 1). Each respondent was randomly assigned three vignettes to mitigate survey fatigue [68]. For each vignette, respondents rated the acceptability on a 5-point Likert scale from "Totally unacceptable" to "Totally acceptable"; for ratings other than "Neutral", they also wrote a short open-ended rationale about their choice.

Prior work found initial evidence [32, 51, 85] or hypothesized [81, 99] that acceptability may vary across behaviors. Thus, we assess acceptability for five AIG-NCII behaviors:

(1) creation of the video
(2) private_sharing by the creator, e.g., in a group chat
(3) public_sharing by the creator, e.g., posting it on Reddit
(4) resharing, publicly, by someone who received the video from the creator
(5) seeking_out by someone with whom it was not shared, e.g., searching online by a description of the video

ID	creator	action	intent
V1	an *intimate partner*	performing a *sexual act*	*entertainment*
V2	an *intimate partner*	performing a *sexual act*	*harm*ing you
V3	an *intimate partner*	performing a *sexual act*	*sexual pleasure*
V4	an *intimate partner*	playing a *sport*	*entertainment*
V5	an *intimate partner*	playing a *sport*	*harm*ing you
V6	an *intimate partner*	playing a *sport*	*sexual pleasure*
V7	an *intimate partner*	*saying something*	*entertainment*
V8	an *intimate partner*	*saying something*	*harm*ing you
V9	an *intimate partner*	*saying something*	*sexual pleasure*
V10	a *stranger*	performing a *sexual act*	*entertainment*
V11	a *stranger*	performing a *sexual act*	*harm*ing you
V12	a *stranger*	performing a *sexual act*	*sexual pleasure*
V13	a *stranger*	playing a *sport*	*entertainment*
V14	a *stranger*	playing a *sport*	*harm*ing you
V15	a *stranger*	playing a *sport*	*sexual pleasure*
V16	a *stranger*	*saying something*	*entertainment*
V17	a *stranger*	*saying something*	*harm*ing you
V18	a *stranger*	*saying something*	*sexual pleasure*

Table 1: The ID and contextual details of creator, action, and intent of each vignette. The italicized portions of the contextual details are the shorthand descriptions of the vignettes used in the paper text, e.g., V1 - intimate partner/sexual act/entertainment. The highlighted vignettes are AIG-NCII.

Sexual Consent Scale-Revised. To answer RQ3a about how attitudes towards sexual consent shape attitudes towards synthetic media, we use two validated subscales from the Sexual Consent Scale-Revised (SCS-R) [45] (Appendix A): SCS-R2 and SCS-R4. SCS-R2 measures respondents' attitudes toward establishing consent, with scores ranging from 1 (low) to 7 (high). SCS-R4 measures agreement with sexual consent norms based on relationship status and sexual activity.

Genuine Intimate Imagery (GII) and NDII Attitudes. To answer (RQ3b) about attutides towards sexual content, we assessed respondent attitudes on creation of intimate media in intimate relationships. Paralleling the vignettes, we also asked about four non-AI scenarios involving nonconsensual distribution of intimate images (NDII): (1) private sharing and (2) public sharing by the intended recipient, as well as (3) public sharing and (4) seeking out by someone who was *not* the intended recipient.

Demographics. The survey concluded with demographic questions, including gender (RQ4).

3.2 Respondents

We used power analysis to determine the required number of respondents for constructing our regression models with the ability to observe small-to-medium effects. We recruited 315 Prolific respondents who were over 18, lived in the US, and had over 95% approval on Prolific. Respondents who did not pass a Pew attention check question [13] or provided incoherent open-ended responses were excluded. The survey took an average of 15 minutes to complete. We compensated respondents $3, which we calculated based on our average pilot test length (12 minutes) and a rate of $15/ hour. 156 respondents were women, 150 were men, 6 were non-binary, 2 were agender, and 1 preferred not to say. Further demographic information is presented in Appendix B.

3.3 Data analysis

Quantitative analysis. Respondents' attitudes towards non-consensual synthetic imagery were analyzed using cumulative link mixed models (CLMMs). We built five CLMMs, one for each of the dependent variables concerning the synthetic video described in the vignettes, listed above. Each model included the same six independent variables. The first three were the vignette factors (creator, action, intent) (RQ1 & RQ2). For RQ3a, we included participant scores on the two SCS-R subscales. To evaluate potential co-linearity between variables, we tested the correlation between scores on the SCS-R subscales. Finding only a weak Spearman's correlation coefficient of -0.3 [3], we proceeded with including both subscales as distinct dependent variables.

Additionally, each model included one context-relevant independent variable capturing attitudes towards similar situations involving GII and NDII (RQ3b). For example, the model for creation included attitudes towards the creation of GII within an intimate partnership as an independent variable and the model for private_sharing included attitudes towards the indented recipient of GII sharing it privately outside the relationship, without consent. During initial analysis, we decided to bucket these attitude items into "unacceptable" and "not unacceptable" to increase our statistical power. Lastly participant gender (bucketed into men and minoritized genders, see below) was included to address RQ4.

AIG-NCII is a form of image-based sexual abuse and tech-facilitated gender-based violence, which is predominantly, though not exclusively, perpetrated by cisgender men targeting cisgender women, transgender people, and/or non-binary people [23, 24, 56, 57, 93]. While research continues to investigate gendered proportions of perpetration and victimization—one report finds that most online AIG-NCII targeted women [2], another report finds that men were more likely to report AIG-NCII victimization than women [85]—attitudes are nevertheless informed by the broader dynamics of gender-based violence. Thus, mens' attitudes of AIG-NCII may differ from the attitudes of people who are not men. In order to increase statistical power, we grouped people who were not men together, i.e., women, agender, or non-binary individuals and refer to this group as "marginalized genders."[3] Further, we only had 8 respondents who self-identified as agender, or non-binary; we bucketed them with women to include their responses in our quantitative analyses, rather than dropping the responses entirely.

To further examine the contextual factors' effect on acceptability (RQ2), another CLMM was built by adding interactions terms between intent and action as well as intent and creator to the original model for creation. To examine the effect of participant gender on attitudes towards synthetic imagery (RQ4), five additional models were built by expanding the original models to include interactions terms between gender and each vignette factor. Of the expanded models, only the creation model showed statistically significant interaction effects ($p < 0.05$) and thus was selected for further analysis. To compare acceptability across the actions of creation, private_sharing, public_sharing, resharing, and seeking_out, another model was built with acceptability rating as the dependent variable and these actions as the independent variable.

Qualitative analysis. We analyzed respondents' open-text rationale for their acceptability rating for the creation of the synthetic video using a coding reliability approach [9]. The dataset was divided into two subsets, justifications for and against acceptability. Two researchers familiarized themselves with all rationales and generated an initial set of codes. The researchers compared and discussed codes to establish a final codebook (Appendix D). In line with qualitative research perspectives on the limitations of multiple coders [4, 60], a single researcher performed the entire coding process for consistency and to preserve interpretive nuance [26]. A second researcher reviewed the codebook as well as 50 random responses from each subset in order to balance researcher subjectivity with thoroughness [89].

3.4 Other considerations

Ethical considerations. This study was deemed exempt by our IRB. However, ethical considerations extend beyond regulatory compliance [8]. As some vignettes describe the non-consensual creation and sharing of synthetic intimate imagery, we were concerned about the potential harm caused by placing respondents into hypothetical victimization scenarios, especially for those who have experienced image-based sexual abuse or other forms of sexual violence.

Consulting subject-area experts with training in clinical psychology and sexual trauma, we took the following steps for harm reduction: (1) surfacing in the consent form that the vignettes described synthetic media being created of the

[3] In our survey, we did not ask whether respondents were transgender, so our sample of men includes transgender and cisgender men.

respondent, (2) asking for re-consent after defining generative AI, (3) including 'prefer not to answer' option for all questions about intimate images, and (4) including contact information for IBSA support organizations at the end of the survey. We also provided support resources for members of the research team who analyzed open-ended survey responses.

Positionality statement. Recognizing the inherent subjectivity in research, we acknowledge that our positionality as researchers shapes our approach to this work [6, 10, 42]. We bring varied perspectives informed by our distinct social, cultural, disciplinary, and ideological contexts. Our research team consists of three cisgender women and one cisgender man who are all researchers in security and privacy. As our team composition does not fully reflect the diversity of identities among our study respondents, there may be limitations in our thematic analysis and interpretation of the collected data.

Limitations. While surveys offer valuable insights, there are inherent limitations to using them. We prioritized reducing survey fatigue by pre-testing and piloting our survey. To minimize social desirability bias, we emphasized that each response about acceptability was based solely on the respondent's personal opinions. Our data is limited to the attitudes and justifications respondents were willing to report.

Crowdworking platforms offer access to large and diverse populations and are frequently used to elicit security and privacy attitudes [28, 75, 87]; we chose Prolific for its higher data quality compared to other platforms [67, 70]. Anticipating that attitudes towards AIG-NCII vary by country, we chose to recruit solely in the US, which likely limits generalizability.

4 Results

To quantitatively analyze the 315 survey responses, we built eight CLMMs (see Section 3.3). The complete regression results for five, including the odds ratio (OR), confidence interval, and p-value range for each independent variable, are in Table 2 (see Appendix E, Figure 4 for visualization).

Additionally, we conducted thematic analysis of the 861 open-response explanations of why participants found the creation of synthetic media in each vignette either acceptable or unacceptable. Aligned with qualitative methods, our analysis aimed to surface general themes about participants' attitudes, rather than quantify their prevalence. Accordingly, we report the appearance of themes using the following terminology: a few (less than 25%), some (25-45%), about half (45-55%), most (55-75%), and almost all (75-100%). When providing participant quotes, we refer to each participant with the letter 'P' followed by their unique participant number and specify the vignette they were responding to. Figure 7 and Figure 8 in Appendices E visualize distributions of codes over vignettes and actions. In some figures and this section, vignettes are referenced by their ID (e.g., V5) and the factor description creator/action/intent (see Table 1).

In our results, we use *synthetic* media to refer to media that is AI-generated, e.g., deepfakes, and *AIG-NCII* to refer to synthetic media that are specifically intimate imagery.

4.1 General Attitudes

People generally found the creation of synthetic media unacceptable, with a median percentage of somewhat or totally unacceptable ratings across all scenarios of 89.54%. They perceived any sharing of these media as even more unacceptable: 94.39% for private_sharing, 94.44% for public_sharing, 94.22% for resharing. Attitudes were more mixed regarding seeking_out such media, however (52.78%). The results of the regression examining the acceptability rating as the dependent variable with these actions as the independent variable, support these results statistically (see Table 6 in Appendix C for full results): Across scenarios and controlling for within-subject variation we observe that private_sharing (OR $= 0.47, p < 0.001$), public_sharing (OR $= 0.26, p < 0.001$), and resharing (OR $= 0.42, p < 0.001$) are significantly less acceptable than creation (the reference level). seeking_out (OR $= 5.43, p < 0.001$) is significantly more acceptable than creation.

Figure 1 illustrates these results visually, depicting perceived acceptability across creation, private_sharing, public_sharing, resharing, and seeking_out for all vignettes. The rightmost column (seeking_out) exhibits far more variance in attitudes than the columns to the left, although these variances differ depending on the depicted action, as we investigate next.

AIG-NCII perceived as less acceptable than other synthetic media not depicting sexual acts. While people broadly found creation and any form of sharing of synthetic media unacceptable, this was particularly true for AIG-NCII (RQ2). Across creation, private_sharing, public_sharing, and resharing contexts, scenarios in which the action was playing a sport or saying something, as opposed to performing a sexual act, were rated as more acceptable by participants (OR $> 7, p < 0.001$ for all models in Table 2).

Turning again to Figure 1, we observe this effect clearly. With regards to creation, the most accepted scenario depicting a sexual act (V3) – an intimate partner non-consensually creating synthetic media of the participant engaged in a sexual act for their sexual pleasure – was considered (somewhat or totally) unacceptable by more than 3/4 of participants (82%). The most accepted scenario depicting the participant saying something (V7) – an intimate partner non-consensually creating synthetic media of the participant saying something for entertainment – was considered unacceptable by about half of participants (45.1%). The most acceptable scenario in our entire survey (V4), which depicted an intimate partner non-consensually creating synthetic media of the participant playing a sport was considered unacceptable by just a third (32%) of participants.

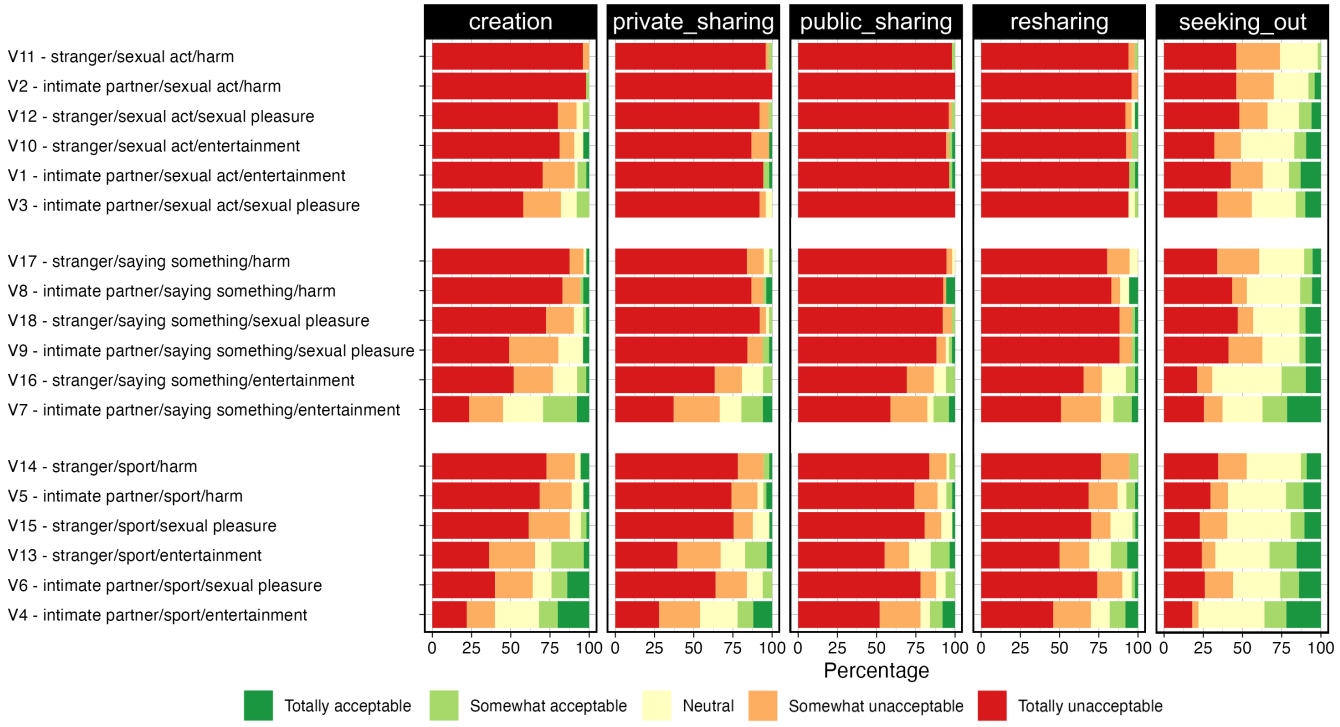


Figure 1: Respondents' perceptions of acceptability across all vignettes; each vignette is defined by the creator / action / intent.

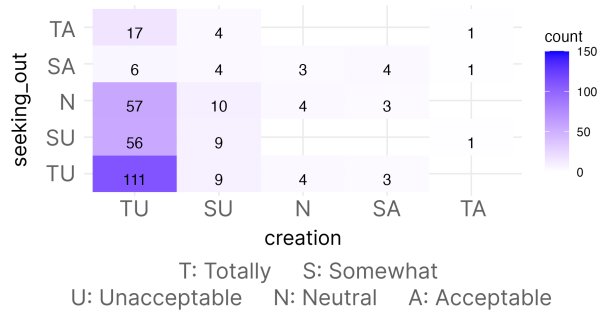


Figure 2: Heatmap of acceptability for creation and seeking_out when the action is performing a sexual act.

seeking_out AIG-NCII was also viewed as less acceptable than seeking_out other forms of synthetic content (OR $> 3, p < 0.001$; Table 2). However, when comparing seeking_out AIG-NCII to creating it, it is still more acceptable than creation as illustrated by Figure 2.

Portrayed action relates to perceived harm. When explaining their perception of a scenario, some participants remarked on potential harm to their reputation or lack thereof to explain why they viewed creation as acceptable or unacceptable. Lack of harm was the most common reason for finding synthetic media creation acceptable, typically when that media depicted the subject playing a sport. For example:

> *There is nothing sexual. . . that i woul[dn'}t want the public to know/see (P50, V13 - stranger/sport/entertainment).*
>
> *It is not harmful to me to have a video of me playing a sport, but it is a little weird (P74, V13 - stranger/sport/entertainment).*

On the other hand, when discussing AIG-NCII or depictions of them saying something they did not, some participants remarked on the potential harms of that content:

> *Sexual act will tarnish my image in the society (P193, V10 - stranger/sexual act/entertainment).*
>
> *AI can seem realistic. Whatever they have me saying could be used against me in a variety of situations (P32, V16 - stranger/saying something/entertainment).*

Further, when the action was performing a sexual act, a few participants also observed that the creation of AIG-NCII wrong because — even if synthetic — the images violated the sanctity of their bodies, e.g.:

> *It's a violation of my body and it is disrespectful (P49, V10 - stranger/sexual act/entertainment).*
>
> *I feel it's unacceptable to manipulate my image in such a way - my body and how it looks belongs to me (P195, V1 - intimate partner/sexual act/entertainment).*

Finally, while we only asked respondents to explain their judgements of (un)acceptability relating to media creation

		creation	private_sharing	public_sharing	resharing	seeking_out
Intercepts	Totally unacceptable | Somewhat unacceptable	5.13 [0.42, 62.94]	1.19 [0.29, 115.33]	29.21 [0.86, 987.2]	1.02 [0.04, 24.7]	0.03* [0, 0.8]
	Somewhat unacceptable | Neutral	29.38** [2.35, 367.77]	5.82 [0.29, 115.33]	122.87** [3.5, 4318.76]	4.72 [0.19, 114.86]	0.16 [0.01, 4.42]
	Neutral | Somewhat acceptable	99.64*** [7.82, 1269.53]	18.57 [0.93, 370.82]	289.97** [8.06, 10433.52]	15.28 [0.62, 374.82]	5.29 [0.19, 145.78]
	Somewhat acceptable | Totally acceptable	375.02*** [28.62, 4913.68]	83.63** [4.08, 1713.2]	1481.83*** [38.75, 56663.78]	70.09* [2.78, 1767.14]	25.60 [0.92, 710.87]
Controlled IVs	creator (Intimate partner)	3.24*** [2.23, 4.71]	1.69* [1.13, 2.55]	1.47 [0.9, 2.4]	1.00 [0.65, 1.53]	1.11 [0.8, 1.53]
	action (Sport)	13.39*** [7.96, 22.52]	34.72*** [16.76, 71.92]	66.61*** [22.75, 19504]	32.36*** [15.12, 69.25]	7.26*** [4.73, 11.15]
	action (Saying something)	5.44*** [3.27, 9.05]	11.01*** [5.45, 22.23]	19.49*** [6.91, 54.94]	12.47*** [5.92, 26.29]	3.40*** [2.21, 5.22]
	intent (Entertainment)	18.92*** [11.03, 32.46]	11.49*** [6.59, 20.05]	10.57*** [5.39, 20.73]	5.51*** [3.18, 9.56]	4.94*** [3.25, 7.49]
	intent (Sexual pleasure)	7.42*** [4.42, 12.47]	1.35 [0.77, 2.37]	1.15 [0.58, 2.28]	0.92 [0.52, 1.63]	1.37 [0.92, 2.04]
Uncontrolled IVs	Gender (Man)	2.45*** [1.45, 4.15]	2.12** [1.21, 3.7]	1.77 [0.88, 3.55]	1.41 [0.75, 2.66]	1.51 [0.76, 2.99]
	GII & NDII attitudes (Unacceptable)	0.21* [0.05, 0.84]	0.08* [0.01, 0.4]	0.09** [0.02, 0.41]	0.01*** [0, 0.05]	0.01*** [0.01, 0.03]
	SCS-R2	0.53*** [0.39, 0.72]	0.55*** [0.4, 0.77]	0.64* [0.42, 0.96]	0.76 [0.52, 1.1]	0.73 [0.48, 1.1]
	SCS-R4	1.06 [0.82, 1.36]	1.10 [0.84, 1.44]	1.30 [0.92, 1.82]	1.27 [0.93, 1.72]	1.14 [0.82, 1.59]

Table 2: Results from a single regression exploring the relationship between scenario acceptability (first row, intercepts), contextual factors (second row, controlled IVs), and personal factors (third row, uncontrolled IVs). Each column represents the output of one regression model. Numeric cells list the odds ratio (OR) and the 95% confidence interval. Reference levels: creator (stranger), action (sexual act), intent (harm), gender (marginalized genders), GII & NDII attitudes (acceptable). Significance of OR: $p < 0.05$ = * , $p < 0.01$ = ** , and $p < 0.001$ = *** .

(Section 3.1), some mentioned the stage of media production (e.g., creation vs. any form of sharing) influenced the likelihood of harm and thus their perception of acceptability:

It's not harming me or blackmailing me or anything. As long as it doesn't get shared I think it's ok (P163, V3 - intimate partner/sexual act/sexual pleasure).

Some respondents call on morality, legality, and privacy to explain the unacceptability of synthetic media. A few participants justified the creation of synthetic media depicting them as unacceptable because it was amoral or unethical to create fake content without the subject's consent, e.g.,

This is a false representation of me and highly unethical (P204, V16 - stranger/saying something/entertainment).

I don't think it is right to use a person[']s identity to say things that they didn't say (P302, V16 - stranger/saying something/entertainment).

While not specifically speaking to amorality, a few expressed sentiments of disgust often associated in psychological literature with intuitive responses to moral violations [38]: that the creation of the content was 'gross' (P50), 'creepy' (P24), 'weird' (P74), or 'nasty' (P112). Such feelings were especially prevalent when the content was created by a stranger or the action depicted was incongruous with the intent (e.g., a stranger creating a video of someone playing a sport for sexual pleasure). We explore these variations based on contextual factors further in Section 4.2.

In a few other cases, participants referred to the creation of the media as illegal or compared it to a crime, despite the fact that no federal legal protections currently exist on AIG-NCII [91]. Across all actions, participants called the act of creation slanderous, like P268 in response to V14 (stranger/sport/harm):

They are using faked info to harm me. This is slander.

When the action was saying something, the creation was often compared to libel or fraud, e.g.,

It seems like the equivalent of slander and fraud. If this were done in election ads, it would be disallowed/illegal (P253, V17 - stranger/saying something/harm).

[I]t is never acceptable to lie. I would sue for libel (P259, V7 - intimate partner/saying something/entertainment).

Specific to AIG-NCII, participants mentioned crimes of sexual violence,

This scenario is harmful and akin to some form of sexual ha[r]assment or assault, especially done without knowledge (P212, V2 - intimate partner/sexual act/harm).

Finally, a few respondents called the creation of synthetic media of them a privacy violation, e.g.:

This completely violates my sense of privacy (P10, V2 - intimate partner/sexual act/harm).

Creating an image of a person without their knowledge is a violation of privacy (P170, V6 - intimate partner/sport/sexual pleasure).

This attitude appeared relatively evenly and similarly in rationales across all actions.

4.2 Role of contextual factors (RQ2)

It is more acceptable for intimate partners to create synthetic media than strangers, but only if they do not intend harm. We observe from Table 2 that across all scenarios, when the content creator was an intimate partner as opposed to a stranger, participants were more likely to find the creation (OR $= 3.24, p < 0.001$; Table 2) as well as the private_sharing (OR $= 1.69, p = 0.01$; Table 2) of the synthetic imagery more acceptable (RQ2). However, when we consider interactions with the intent of the synthetic media (Table 3), we observe that there is no longer a significant relationship between creator and acceptability of creation and that there are three significant interactions between: (1) creator being an intimate partner and intent being entertainment (OR $= 2.83, p = 0.036$; Table 3), (2) creator being an intimate partner and intent being sexual pleasure (OR $= 3.76, p = 0.009$; Table 3), as well as between (3) action being playing a sport and intent being sexual pleasure (OR $= 0.08, p = 0.002$; Table 3), which we address later in this section. Thus, our interaction model demonstrates a more nuanced answer to RQ2. The main effect we observed in our original modeling for creation (without interactions) – that intimate partners creating synthetic media is more acceptable – was driven by attitudes that intimate partners creating synthetic media for non-harmful purposes is more acceptable. That is, if the creator is an intimate partner and the intent is entertainment (OR $= 2.83, p = 0.036$; Table 3) or sexual pleasure (OR $= 3.76, p = 0.009$; Table 3) the media creation is more acceptable. However, intimate partners creating media for the intent to harm is no more acceptable than a stranger doing so (see visualization in Appendix E, Figure 6).

Intimate partner trust related to explanations of (un)acceptability. Some explanations for acceptability, like P211's response to V1 (intimate partner/sexual act/entertainment), reflected trust in a partner enabling acceptable creation:

I feel if we are intimate, we're already engaging in similar acts. It's all in good sexual fun, as long as they don't distribute it or show anyone else.

This exhibits a belief that an intimate relationship permits intimate media creation within it, whereas no such trust exists in relationships with strangers, increasing feelings of violation:

The idea of somebody I don't know generating porn of me is insanely creepy (P24, V12 - stranger/sexual act/sexual pleasure)

On the other hand, some explanations for unacceptability

		OR; Confidence Interval
Intercepts	Totally unacceptable | Somewhat unacceptable	7.51; [0.42, 134.86]
	Somewhat unacceptable | Neutral	47.58; [2.61, 867.1]**
	Neutral | Somewhat acceptable	171.35; [9.26, 3169.4]***
	Somewhat acceptable | Totally acceptable	665.83; [35.15, 12613]***
Controlled IVs	creator (Intimate partner)	1.38; [0.62, 3.04]
	action (Sport)	48.94; [11.43, 209.59]***
	action (Saying something)	9.9; [2.24, 43.77]**
	intent (Entertainment)	13.14; [2.86, 60.3]***
	intent (Sexual pleasure)	20.82; [4.53, 95.72]***
Uncontrolled IVs	Gender (man)	2.64; [1.53, 4.57]***
	GII & NDII attitudes (Unacceptable)	0.19; [0.04, 0.8]*
	SCS-R2	0.51; [0.37, 0.71]***
	SCS-R4	1.08; [0.83, 1.4]
Interaction Terms	creator (Intimate partner) & intent (Entertainment)	2.83; [1.07, 7.5]*
	creator (Intimate partner) & intent (Sexual pleasure)	3.76; [1.38, 10.2]**
	action (Sport) & intent (Entertainment)	0.72; [0.15, 3.56]
	action (Saying something) & intent (Entertainment)	1.68; [0.33, 8.66]
	action (Sport) & intent (Sexual pleasure)	0.08; [0.02, 0.4]**
	action (Saying something) & intent (Sexual pleasure)	0.19; [0.04, 1.01]

Table 3: Results from a single regression exploring the relationship between the acceptability of creation (first row, intercepts), contextual factors (second row, controlled IVs), personal factors (third row, uncontrolled IVs), and interactions between intent and creator or action (fourth row, interaction terms). Reference levels: creator (stranger), action (sexual act), intent (harm), gender (marginalized genders), GII & NDII attitudes (acceptable). Significance of OR: $p < 0.05$ = *, $p < 0.01$ = **, and $p < 0.001$ = ***.

stated that the creation *violated* intimate partner trust rather than being acceptable because of it, e.g.,

> *I think this is just as worse because there is supposed to be a trust between people who are intimate and they completely broke that trust (P142, V3 - intimate partner/sexual act/sexual pleasure).*

About half of the rationales exhibiting this attitude were in response to the creation of synthetic media of sexual acts.

A few were flattered by the creation of material for sexual fantasy within an intimate partnership. In scenarios where synthetically generated media was created for sexual gratification by an intimate partner, a few participants reported feelings of being flattered by its production, e.g.,

> *The content she generated sounds cool and indicates she's attracted to me (P65, V6 - intimate partner/sport/sexual pleasure).*
>
> *I don't care what my intimate partners choose to do. I would be flattered (P65, V9 - intimate partner/saying something/sexual pleasure).*

A few noted that they couldn't control the sexual fantasies of others, regardless of whether they were in a relationship:

> *I don't particularly like that and I would prefer they don't do it, but I can't stop them from fantasizing about me in their own head. I can't stop them from writing down their fantasies on paper or drawing a picture or generating a synthetic video (P188, V12 - stranger/sexual act/sexual pleasure).*

While others expressed that, in the context of an intimate relationship, they would prefer to engage in their partner's fantasy by generating GII instead:

> *It's a bit bizarre and strange. I'd rather I actually perform this act instead of a fake AI version of me doing so (P165, V1 - intimate partner/sexual act/entertainment).*

Intent impacts acceptability ratings differently depending on stage in the media pipeline. We observe from Table 2 that regardless of the creator of the media, respondents rated as more acceptable those scenarios where synthetic videos were created, shared, and sought out for entertainment vs. with intent to harm (OR $> 4, p < 0.001$; Table 2). Respondents also found creation of synthetic videos with the intent of bringing the creator sexual pleasure more acceptable than creation with the intent to harm the subject. However, respondents did not rate the acceptability of any form of sharing or seeking_out synthetic videos created with the intent of sexual pleasure differently from the acceptability of sharing or seeking_out synthetic videos created with the intent to harm.

Incongruent actions and intentions increase unacceptability. Considering our interaction model, we find that these results hold but observe a further effect: incongruence between the action and the intent – even for actions and intents viewed as generally more acceptable – reduce attitudes of acceptability. For example, while creating media depicting the subject playing a sport was overall more acceptable than depicting them engaged in a sexual act and depictions of any action for sexual pleasure were more acceptable than depictions for harm, depicting someone playing a sport with the intent of sexual pleasure was less acceptable than depicting a more congruous action (saying something, a sexual act) with the same intent. A few participants shared explanations for the (un)acceptability of synthetic media creation that support this finding, for example, P25 when discussing V15 (stranger/sport/sexual pleasure) remarks:

> *That's really creepy! It just grosses me out, even if it's just sports.*

4.3 Role of sexual consent & content attitudes

Attitudes toward establishing sexual consent offline relate to attitudes toward AI media generation and sharing. We used the second subscale from the SCS-R to measure attitudes towards establishing sexual consent [45] and answer RQ3a. Those who scored higher on SCS-R2, indicating more positive attitudes toward establishing sexual consent, were less likely to rate non-consensual creation, private_sharing or public_sharing of synthetic content as acceptable (OR $< 0.7, p < 0.005$ for these models; in Table 2).

The most common explanation for finding synthetic media creation unacceptable is lack of consent. For example, P19 remarked in response to V3 (intimate partner/sexual act/sexual pleasure) that:

> *No content should be made in someone else's likeness without their consent.*

The fourth SCS-R subscale measures attitudes towards consent norms specifically in the context of relationships and sexual activity [45]. Scores on this subscale did not significantly affect any models.

Attitudes toward consensually-created genuine intimate imagery as well as NDII correlate with acceptance of synthetic videos including AIG-NCII. In addressing RQ3b, we sought to understand whether and how attitudes toward genuine, consensually-created intimate imagery related to attitudes toward synthetic, non-consensually created media.

Those who found consensual creation of genuine intimate imagery (GII) in an intimate relationship (somewhat or completely) unacceptable were also less likely to find non-consensual, synthetic creation of media depicting them acceptable, regardless of the act depicted (OR $= 0.21, p = 0.028$; Table 2). Those who found further sharing of GII without the original sender's consent – i.e., non-consensual distribution of

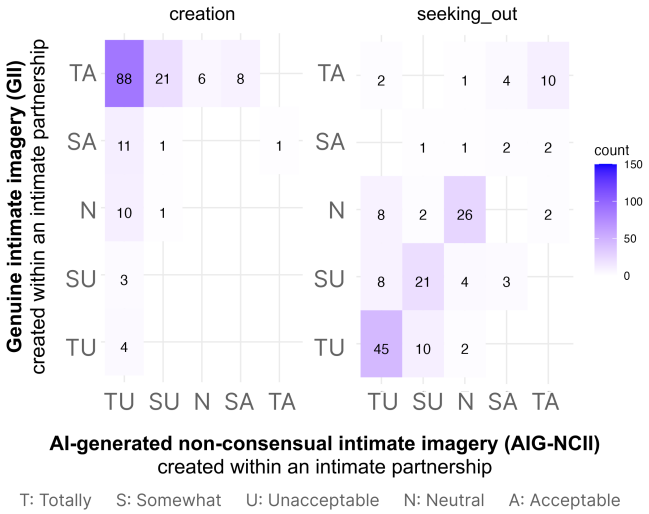


Figure 3: Heatmaps comparing acceptability of creation and seeking_out for AIG-NCII to similar actions for GII also created in an intimate relationship. See Appendix E, Figure 5 for heatmaps including all forms of sharing.

intimate imagery or NDII – unacceptable were also less likely to find sharing of synthetic videos depicting them acceptable ($OR < 0.1, p < 0.05$ for private_sharing, public_sharing, and resharing; Table 2). Finally, those who considered seeking out NDII unacceptable were less likely to find seeking_out synthetic videos acceptable ($OR = 0.01, p < 0.001$; Table 2).

Comparing AIG-NCII to GII in the context of an intimate relationship, we observe in Figure 3 that a little over half of respondents found consensual GII creation within an intimate partnership totally acceptable (88/154), while none viewed non-consensual synthetic intimate media creation within an intimate partnership as totally acceptable. A key difference is that the GII creation scenario implies awareness and consent, while the synthetic media vignettes explicitly do not involve awareness or consent. Considering non-consensual sharing, a majority of respondents viewed private_sharing (140/153[4]), public_sharing (145/151), and resharing (139/153) as totally unacceptable for both media types.

There was less consensus on seeking_out non-consensually publicized synthetic and non-synthetic imagery, with only some (45/154) finding it totally unacceptable for both.

4.4 Role of gender

For quantitative analysis, we binned respondents by gender into men and marginalized genders, as discussed in Section 3.3. Across scenarios, men were more likely to rate the creation ($OR = 2.45, p < 0.001$; Table 2) and private_sharing ($OR = 2.12, p = 0.009$; Table 2) more acceptable than people with a marginalized gender.

[4]Denominators vary because some participants preferred not to answer certain questions about synthetic and/or authentic intimate imagery.

		OR; Confidence Interval
Intercepts	Totally unacceptable | Somewhat unacceptable	4.77; [0.33, 69.72]
	Somewhat unacceptable | Neutral	29.29; [1.97, 435.77]*
	Neutral | Somewhat acceptable	104.04; [6.89, 1570.59]***
	Somewhat acceptable | Totally acceptable	410.88; [26.53, 6363.7]***
Controlled IVs	creator (Intimate partner)	1.75; [1.04, 2.96]*
	action (Sport)	17.58;[8.06, 38.33]***
	action (Saying something)	10.68; [4.82, 23.65]***
	intent (Entertainment)	20.05; [9.39, 42.85]***
	intent (Sexual pleasure)	4.90; [2.32, 10.33]***
Uncontrolled IVs	Gender (man)	1.54; [0.43, 5.61]
	GII & NDII attitudes (Unacceptable)	0.2; [0.05, 0.84]*
	SCS-R2	0.52; [0.38, 0.72]***
	SCS-R4	1.08; [0.83, 1.41]
Interaction Terms	action (Sport) & Gender (Man)	0.77; [0.29,2.02]
	action (Saying something) & Gender (Man)	0.32; [0.12, 0.88]*
	intent (Entertainment) & Gender (Man)	1; [0.38, 2.61]
	intent (Sexual pleasure) & Gender (Man)	2.36; [0.87, 6.43]
	creator (Intimate partner) & Gender (Man)	3.59; [1.71, 7.5]***

Table 4: Results from a single regression exploring the relationship between scenario acceptability for creation (first row, intercepts), contextual factors (second row, controlled IVs), personal factors (third row, uncontrolled IVs), and interactions between gender and contextual factors (third row, interaction terms). Reference levels: creator (stranger), action (sexual act), intent (harm), gender (marginalized genders), GII & NDII attitudes (acceptable). Significance of OR: $p < 0.05$ = *, $p < 0.01$ = **, and $p < 0.001$ = ***.

Men view synthetic media depicting them engaged in a sexual act more acceptable than others. To further examine the role of gender identity in shaping attitudes towards non-consensual synthetic imagery creation, we performed an addition regression that included interaction terms between participant gender and each vignette factor (Table 4). We observe that the main effect of gender is no longer significant, instead finding two significant interactions with gender. The first shows that, while participants viewed creation of synthetic videos of them saying something as more acceptable than a sexual act, people of marginalized genders were more likely to do so than men ($OR = 10.71$ for men vs. $OR = 3.42$ for marginalized genders, $p = 0.027$). This may align with a long history in defamation law of a double

standard between what harms a man's reputation compared to a woman's [5, 71] wherein laws for a long period specifically differentiated defamation as applying to claims about a woman's sexual behavior but not a man's [79].

Participants who are men are more accepting of intimate partners creating synthetic videos depicting them. Secondly, we observe that, holding all other factors constant, men were more likely to rate the creation of synthetic media by an intimate partner more acceptable (OR = 1.77 for men vs. OR = 6.27 for marginalized genders, $p < 0.001$). Again, this may align with defamation-related legal discussions about perception that claims of sexual activity could enhance a men's reputation but harm a woman's [5, 71]. From a critical perspective [74], in a relationship in which one person is a man, the man may have greater power; that greater power could contribute to them viewing the threat of synthetic videos as less likely or less harmful and, if such a video is created, the person with more power may be more confident in their ability to respond. Additionally, most participants who described the creation of AIG-NCII in an intimate partnership as being acceptable because it was a compliment or part of their partner's fantasy (as discussed in Section 4.2) were men.

5 Discussion

Overall, we find that creating, sharing, or seeking AIG-NCII is considered far less acceptable than creating, sharing, or seeking other forms of non-consensually-created synthetic media (RQ1: Section 4.1). Respondents were more accepting of intimate partners creating synthetic media of them than strangers, including AIG-NCII, but only when their intent in doing so was not to cause harm (RQ2: Section 4.2). Lack of consent was the most common reason respondents provided for why non-consensual creation of synthetic media, including AIG-NCII, was unacceptable. Our statistical models support this finding: positive attitudes toward sexual consent were inversely correlated with acceptance of non-consensual creation, sharing, or seeking_out of synthetic media of any kind (RQ3: Section 4.3). The second most common reason respondents gave for why creation was unacceptable was potential for harm, either reputational damage or bodily violation; conversely, the lack of potential for such harm was the most common reason among those who found creation acceptable. Men in particular were more accepting of synthetic media creation (RQ4: Section 4.4), especially by intimate partners. We hypothesize based on prior literature on perceptions of sexual reputation in the context of defamation law [5, 71, 79] and participants' open-text responses that this is likely due to differences in perception regarding reputation damage and creation as a form of compliment as well as, from a critical perspective [74], that men may be more accepting of such images if they have more power in a relationship. Respondents also expressed attitudes of unacceptability due to moral violations [38], including feelings of disgust, and privacy violation.

We focus the remainder of our discussion on implications for addressing the most unacceptable use of AI generative capabilities we find in our study, AIG-NCII, although we note that the implications are relevant to other synthetic media.

Distributed responsibility and individual deterrence. We believe it is important to understand the gap between the unacceptability of creation and sharing and the relative acceptability of searching for, and subsequently viewing, of AIG-NCII. Based on our results, we hypothesize that one contributing factor to the continued ubiquity of AIG-NCII is the broad acceptance of or neutrality toward searching for such content. The finding that searching for and viewing AIG-NCII is perceived as so acceptable suggests the harms entailed in AIG-NCII are not fully appreciated by many people. Yet as studies of the experiences of image-based sexual abuse victim-survivors and even legal cases note, viewing is a primary mechanism of harm for NCII: "there [is] a fresh intrusion of privacy when each additional viewer sees the photograph" [48].

Past works, although not written in the context of AIG-NCII, can provide possible explanations for this gap, which we encourage future research to explore in depth. As media scholar Lilie Chouliaraki concludes in her analysis of the viewing of violent imagery in television and online, "technology closes the moral distance between spectators and sufferers and ... yet, at the same time, it fictionalizes suffering and leads spectators to indifference" [15]. Media scholar Charles Ess [29], in his foundational work *Digital Media Ethics*, argues that such indifferent online behavior in new media networks is due to "distributed responsibility," which refers to the idea that ethical responsibility for an act is distributed across an interconnected, online networks of actors, rather than being attached solely to a single individual [29, 86]. Ess contrasts this collective responsibility with the traditional western understanding of ethical responsibility as matter of individual agency. For example, an individual might never steal an album from a physical record store but may illegally download of music from the Internet. In this and many cases, he argues, individuals consider themselves part of an anonymous, undetectable online collective without fear of punishment.

Thus, a key question for future work is how to combat indifference towards the harm of viewing AIG-NCII. Deterrence messaging, such as keyword-based warnings in search engines or advertisements that inform the viewer about the harms of consuming AIG-NCII, could be used to target individuals' sense of ethical immunity. Emphasizing personal accountability within the collective space could disrupt feelings of distributed responsibly related to AIG-NCII. Such messaging is currently effectively used to deter viewing of child sexual abuse material [69] but further research is necessary to find effective approaches to deter AIG-NCII consumption.

Harms vs. rights When analyzing our data, we observed different classes of arguments for (and against) the unacceptability of AIG-NCII. At the highest level, we saw arguments focused on harms and arguments focused on rights. For example, some argued that creating AIG-NCII was acceptable as long as no harms manifested, e.g., "It's not harming me or blackmailing me ... [a]s long as it doesn't get shared I think it's ok" (Section 4.1): a harms-based analysis. On the other hand, some argued that creating AIG-NCII was unacceptable, even if never shared, because it was a "violation of my body" (Section 4.1): a rights-based evaluation.

While prior work on AIG-NCII has primarily focused on harm perceptions [32,51], these two categories of arguments — harms-based and rights-based — align with the vast literature in philosophy and psychology on how different people may center different values in moral decision making, e.g., see [50] for a survey aimed at the security and privacy community. Using the terminology from philosophy, those who consider AIG-NCII unacceptable because it can lead to harms are centering a utilitarianistic (consequentialist) perspective on ethics; those who consider AIG-NCII unacceptable because it violates an individual's rights even if no harms manifest are centering a deontological perspective.

While our findings surfaced a breadth of rights that participants believe are impacted by the creation and possible sharing of AIG-NCII, we focus on two below: the right to consent, which is baked into the definition of AIG-NCII, and, given the SOUPS community, the right to privacy.

AIG-NCII as a consent violation. To our knowledge, ours is the first work to surface qualitative perspectives on consent for AIG-NCII. Our findings (Section 4.3) suggest connections between understandings and norms around consent in different contexts which have meaningful implications. First, shaping or enforcing norms around sexual consent, or consent in general, could influence norms and behaviors related to non-consensual synthetic media. Consent education, which involves setting and modeling behavioral norms like asking for consent before interacting with another person's body or space, is one approach to establishing and enforcing norms around consent for all ages in both sexual and non-sexual contexts [35, 83]. Second, centering consent as a priority in policies and technical developments around deepfakes is warranted. A growing body of work provides useful frameworks for operationalizing consent in sociotechnical systems [47, 82, 98].

AIG-NCII as a privacy violation. Like consent, privacy is a fundamental right. While our survey instrument did not mention privacy at any point, some participants stated that the creation of the synthetic media would violate their privacy.

The fact that contextual factors such as who created the content and for what purpose influence perceptions of AIG-NCII acceptability in our study aligns with existing technology privacy theory on contextual norms [94] and integrity [65], which find that experiences of privacy violation are dependent on contextual factors including what information is being shared, which actors are involved, and the purpose of the information sharing. This contextual focus is supported by our finding that the creator, action, and intent significantly affected whether respondents viewed AIG-NCII as acceptable (Section 4.2). Thus, frameworks of privacy as contextual norms may be one useful component of future policies about AIG-NCII.

At the same time, existing frameworks and technological conceptions of privacy often focus on *data* privacy. Yet, as technological capabilities continue to develop, technologists must increasingly contemplate how to measure and protect a more nebulous privacy right: to representational privacy. In the context of AIG-NCII the personal data involved is not necessarily sensitive. Rather, what is sensitive is a technologically-produced representation of the self made possible using a small amount of personal data (e.g., a photograph of the subject) and a large amount of other people's data (used to train the model that generated the AIG-NCII).

Legal scholars have already begun to wrestle with this issue, highlighting that existing regulation on privacy may not be wholly sufficient to protect sexual autonomy [18]. Citron proposes the recognition of sexual privacy — "the behaviors, expectations, and choices that manage access to and information about the human body, sex, sexuality, gender, and intimate activities" [18] — to provide more holistic protections for subjects of intimate images. What would a similar reformulation from data privacy to representational privacy mean for the technical security and privacy community? Answering this question will require translating notions of self-representation and consent into technical constraints that can govern systems.

6 Conclusion

Public familiarity with AIG-NCII is still low [85]. As more of it is produced [34] and it becomes easier to produce (e.g., through commercial text-to-video products or "nudify" apps [27, 61]), technological acceptance may increase and attitudes may change [40]. Continued work is needed to track and understand the development of technology for creating and sharing AIG-NCII as well as the attitudes around it. Our study contributes towards the understanding of attitudes towards non-consensual deepfakes across contexts, including AIG-NCII, providing insight into the rationales behind people's attitudes as well as the connections between gender, consent, genuine intimate imagery and these attitudes. Addressing AIG-NCII media requires a multifaceted response blending social science work on norms, legal scholarship, and socio-technical research to detect and prevent creation, sharing and viewing of harmful synthetic media.

Acknowledgments

We thank Samuel Dooley for his guidance and feedback on our statistical analysis. We are also grateful to Rosanna Bellini and Sharon Wang for their feedback regarding ethical survey design. Additionally, we appreciate the members of the Security and Privacy Lab at the University of Washington for their insights and brainstorming contributions.

References

[1] Executive Order No. 14110 - Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 2023.

[2] Henry Ajder, Giorgio Patrini, Francesco Cavalli, and Laurence Cullen. The State of Deepfakes: Landscape, Threats, and Impact. Technical report, Deeptrace Labs, 2019.

[3] H. Akoglu. User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3):91–93, 2018.

[4] David Armstrong, Ann Gosling, John Weinman, and Theresa Marteau. The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. *Journal of Sociology*, 31(2):597–606, August 1997.

[5] Roy Baker. *Defamation law and social attitudes: Ordinary unreasonable people*. Edward Elgar Publishing, 2011.

[6] Shaowen Bardzell and Jeffrey Bardzell. Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. In *Proc. CHI*, 2011.

[7] Samantha Bates. 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, January 2017.

[8] Rasika Bhalerao, Vaughn Hamilton, Allison McDonald, Elissa M Redmiles, and Angelika Strohmayer. Ethical practices for security research with at-risk populations. In *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 546–553. IEEE, 2022.

[9] Virginia Braun and Victoria Clarke. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, 21(1):37–47, 2021.

[10] Virginia Braun and Victoria Clarke. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3):328–352, 2021.

[11] Roberto Caldelli, Leonardo Galteri, Irene Amerini, and Alberto Del Bimbo. Optical Flow based CNN for detection of unlearnt deepfake manipulations. *Pattern Recognition Letters*, 146:31–37, 2021.

[12] Nicholas Caporusso. Deepfakes for the Good: A Beneficial Application of Contentious Artificial Intelligence Technology. In Tareq Ahram, editor, *Advances in Artificial Intelligence, Software and Systems Engineering*, pages 235–241, Cham, 2021. Springer International Publishing.

[13] Pew Research Center. Assessing the Risks to Online Polls from Bogus Respondents. Technical report, Pew Research Center, February 18 2020.

[14] Robert Chesney and Danielle Citron. Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics, January/February 2019. Accessed 29 January 2024.

[15] Lilie Chouliaraki. The spectatorship of suffering. *The Spectatorship of Suffering*, pages 1–240, 2006.

[16] Danielle K. Citron and Robert Chesney. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107:1753, 2019.

[17] Danielle Keats Citron. Sexual privacy. *Yale LJ*, 128:1870, 2018.

[18] Danielle Keats Citron. Sexual Privacy. *Yale Law Journal*, 128(7):1792–2121, May 2019.

[19] Valdemar Danry, Joanne Leong, Pat Pataranutaporn, Pulkit Tandon, Yimeng Liu, Roy Shilkrot, Parinya Punpongsanon, Tsachy Weissman, Pattie Maes, and Misha Sra. AI-Generated Characters: putting Deepfakes to Good Use. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–5, 2022.

[20] A. de Rancourt-Raymond and N. Smaili. The Unethical Use of Deepfakes. *Journal of Financial Crime*, 30(4):1066–1077, 2023.

[21] Rebecca A. Delfino. Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act. *Fordham Law Review*, 88:887, 2019.

[22] Brian Dolhansky, Joanna Bitton, Ben Pflaum, Jikuo Lu, Russ Howes, Menglin Wang, and Cristian Canton-Ferrer. The DeepFake Detection Challenge Dataset. *CoRR*, abs/2006.07397, 2020.

[23] Suzie Dunn. Technology-Facilitated Gender-Based Violence: An Overview. Technical report, Centre for International Governance Innovation, 2020.

[24] Suzie Dunn. Women, Not Politicians, Are Targeted Most Often by Deepfake Videos. Technical report, Centre for International Governance Innovation, 2021.

[25] Asia Eaton, Holly Jacobs, and Yanet Ruvalcaba. Nationwide Online Study of Nonconsensual Porn Victimization and Perpetration. Technical report, Cyber Civil Rights Initiative, 2017.

[26] Victoria Elliott. Thinking about the Coding Process in Qualitative Data Analysis. *Qualitative Report*, 23:2850–2861, 11 2018.

[27] Kim Elsesser. Apps That Undress Women's Photos Surge In Popularity. What Could Go Wrong?, December 2023. Accessed 10 February 2024.

[28] Pardis Emami-Naeini, Joseph Breda, Wei Dai, Tadayoshi Kohno, Kim Laine, Shwetak Patel, and Franziska Roesner. Understanding People's Concerns and Attitudes Toward Smart Cities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.

[29] Charles Ess. *Digital Media Ethics*. Polity Press, Cambridge, 2014.

[30] Hubert Etienne. The Future of Online Trust (and Why Deepfake Is Advancing It). *AI and Ethics*, 1(4):553–562, November 1 2021. ID: Etienne2021.

[31] Hany Farid. Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety*, 1(4), 2022.

[32] Dean Fido, Jaya Rao, and Craig A. Harper. Celebrity status, sex, and variation in psychopathy predicts judgements of and proclivity to generate and distribute deepfake pornography. *Computers in Human Behavior*, 129:107141, 2022.

[33] Asher Flynn, Elena Cama, Anastasia Powell, and Adrian J Scott. Victim-blaming and image-based sexual abuse. *Journal of Criminology*, 56(1):7–25, 2023.

[34] Asher Flynn, Anastasia Powell, Adrian J Scott, and Elena Cama. Deepfakes and Digitally Altered Imagery Abuse: A Cross-Country Exploration of an Emerging form of Image-Based Sexual Abuse. *The British Journal of Criminology*, 62(6):1341–1358, 12 2021.

[35] National Coalition for Sexual Freedom. Consent Counts. Accessed 9 February 2024.

[36] Dilrukshi Gamage, Piyush Ghasiya, Vamshi Bonagiri, Mark E. Whiting, and Kazutoshi Sasahara. Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 103:1–103:19, New York, NY, USA, 2022. Association for Computing Machinery.

[37] Anne Pechenik Gieseke. "The New Weapon of Choice": Law's Current Inability to Properly Address Deepfake Pornography. *Vanderbilt Law Review*, 73:1479, 2020.

[38] Roger Giner-Sorolla, Tom Kupfer, and John Sabo. What makes moral disgust special? an integrative functional review. In *Advances in experimental social psychology*, volume 57, pages 223–289. Elsevier, 2018.

[39] Jeffrey Gottfried. About three-quarters of Americans favor steps to restrict altered videos and images. Report, 2019. (2019).

[40] Andrina Granic. Technology Acceptance Model: a Literature Review from 1986 to 2013. *Universal Access in the Information Society*, 13(1):149–160, 2014.

[41] Jens Hainmueller, Dominik Hangartner, and Teppei Yamamoto. Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8):2395–2400, 2015.

[42] Andrew Gary Darwin Holmes. Researcher Positionality - A Consideration of Its Influence and Place in Qualitative Research - A New Researcher Guide. *International Journal of Education*, 8(4):1–10, 2020.

[43] Wenyi Hong, Ming Ding, Wendi Zheng, Xinghan Liu, and Jie Tang. CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers, 2022.

[44] Antoinette Huber. 'A shadow of me old self': The impact of -based sexual abuse in a digital society. *International Review of Victimology*, 29(2):199–216, 2023.

[45] Terry P. Humphreys and Mélanie M. Brousseau. The Sexual Consent Scale–Revised: Development, Reliability, and Preliminary Validity. *The Journal of Sex Research*, 47(5):420–428, 2010. PMID: 19685367.

[46] Lilly Irani, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E Grinter. Postcolonial computing: a lens on design and development. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1311–1320, 2010.

[47] Farzaneh Karegar. *Towards Improving Transparency, Intervenability, and Consent in HCI*. PhD thesis, Karlstad University Press, 2018.

[48] Tsachi Keren-Paz. *10: The Power of Property: Strict Liability for Viewing NCII*, chapter 10, pages 175–193. Bristol University Press, 2023.

[49] Jan Kietzmann, Linda W. Lee, Ian P. McCarthy, and Tim C. Kietzmann. Deepfakes: Trick or treat? *Business Horizons*, 63(2):135–146, 2020. ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING.

[50] Tadayoshi Kohno, Yasemin Acar, and Wulf Loh. Ethical frameworks and computer security trolley problems: Foundations for conversations. In *USENIX Security*, 2023.

[51] Matthew B. Kugler and Carly Pace. Deepfake Privacy: Attitudes and Regulation. *Nw. UL Rev.*, 116:611, 2021.

[52] Minghui Li and Yan Wan. Norms or fun? The influence of ethical concerns and perceived enjoyment on the regulation of deepfake information. *Internet Research: Electronic Networking Applications and Policy*, 33(5):1750–1773, 2023.

[53] Emanuel Maiberg and Samantha Cole. AI-Generated Taylor Swift Porn Went Viral on Twitter. Here's How It Got There, January 2024. Accessed 29 January 2024.

[54] Nikola Marangunić and Andrina Granić. Technology acceptance model: a literature review from 1986 to 2013. *Universal access in the information society*, 14:81–95, 2015.

[55] Kirsten E. Martin. Diminished or Just Different? A Factorial Vignette Study of Privacy as a Social Contract. *Journal of Business Ethics*, 111(4):519–539, 2012.

[56] Clare McGlynn, Kelly Johnson, Erika Rackley, Nicola Henry, Nicola Gavey, Asher Flynn, and Anastasia Powell. 'It's Torture for the Soul': The Harms of Image-Based Sexual Abuse. *Social & Legal Studies*, 30(4):541–562, 2021.

[57] Clare McGlynn and Erika Rackley. Image-Based Sexual Abuse. *Oxford Journal of Legal Studies*, 37(3):534–561, 2017.

[58] Clare McGlynn, Erika Rackley, and Ruth Houghton. Beyond 'Revenge Porn': The Continuum of Image-Based Sexual Abuse. *Feminist Legal Studies*, 25:25–46, 2017.

[59] Annelise Mennicke, Jessi Bowling, Jennifer Gromer, and Caitlin Ryan. Factors Associated With and Barriers to Disclosure of a Sexual Assault to Formal On-Campus Resources Among College Students. *Violence Against Women*, 27(2):255–273, 2021.

[60] Janice M. Morse. Perfectly Healthy, But Dead: The Myth of Inter-Rater Reliability. *Qualitative Health Research*, 7(4):445–447, November 1997.

[61] Margi Murphy. 'Nudify' apps that use AI to undress women in photos are soaring in popularity, December 2023. Accessed 10 February 2024.

[62] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. Privacy Expectations and Preferences in an IoT World. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 399–412, Santa Clara, CA, July 2017. USENIX Association.

[63] Stuart Napshin, Jomon Paul, and Justin Cochran. Individual Responsibility Around Deepfakes: It's No Laughing Matter. *Cyberpsychology, Behavior, and Social Networking*, 2024.

[64] Yuval Nirkin, Yosi Keller, and Tal Hassner. FSGAN: Subject Agnostic Face Swapping and Reenactment. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[65] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, Stanford, CA, 2010.

[66] Lindsay M. Orchowski, Amy S. Untied, and Christine A. Gidycz. Social Reactions to Disclosure of Sexual Victimization and Adjustment Among Survivors of Sexual Assault. *Journal of Interpersonal Violence*, 28(10):2005–2023, 2013.

[67] Eyal Peer, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. Data Quality of Platforms and Panels for Online Behavioral Research. *Behavior Research Methods*, 54(4):1643–1662, 08 2022. Published on August 1, 2022.

[68] Stephen R. Porter, Michael E. Whitcomb, and William H. Weitzer. Multiple Surveys of Students and Survey Fatigue. *New Directions for Institutional Research*, 121:63–73, 2004.

[69] Jeremy Prichard, Richard Wortley, Paul A Watters, Caroline Spiranovic, Charlotte Hunn, and Tony Krone. Effects of automated messages on internet users attempting to access "barely legal" pornography. *Sexual Abuse*, 34(1):106–124, 2022.

[70] Prolific. Why Prolific?, 2020. Accessed 29 January 2024.

[71] Lisa R Pruitt. Her own good name: Two centuries of talk about chastity. *Md. L. Rev.*, 63:401, 2004.

[72] Albert Pumarola, Antonio Agudo, Adrià Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer. Ganimation: Anatomically aware facial animation from a single image. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 818–833, 2018.

[73] Reddit. Never Post Intimate or Sexually Explicit Media of Someone Without Their Consent. Accessed 29 January 2024.

[74] Elissa M. Redmiles, Mia M. Bennett, and Tadayoshi Kohno. Power in computer security and privacy: A critical lens. *IEEE Security & Privacy Magazine*, March/April 2023.

[75] Elissa M Redmiles, Sean Kross, and Michelle L Mazurek. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1326–1343. IEEE, 2019.

[76] Nataniel Ruiz, Sarah Adel Bargal, Cihang Xie, and Stan Sclaroff. Practical Disruption of Image Translation Deepfake Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(12):14478–14486, Jun. 2023.

[77] Yanet Ruvalcaba and Asia A Eaton. Nonconsensual Pornography among US Adults: A Sexual Scripts Framework on Victimization, Perpetration, and Health Correlates for Women and Men. *Psychology of Violence*, 10(1):68, 2020.

[78] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, Devi Parikh, Sonal Gupta, and Yaniv Taigman. Make-A-Video: Text-to-Video Generation without Text-Video Data, 2022.

[79] Gerald R Smith. Of malice and men: The law of defamation. *Val. UL Rev.*, 27:39, 1992.

[80] StopNCII. Frequently Asked Questions. Website, Accessed 2024.

[81] Daniel Story and Ryan Jenkins. Deepfake Pornography and the Ethics of Non-Veridical Representations. *Philosophy & Technology*, 36(3):56, August 26 2023.

[82] Yolande Strengers, Jathan Sadowski, Zhuying Li, Anna Shimshak, and Florian 'Floyd'Mueller. What can HCI learn from sexual consent? A feminist process of embodied consent for interactions with emerging technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[83] Grace Tatter. Consent at Every Age, December 2018. Accessed 9 February 2024.

[84] Brian Timmerman, Pulak Mehta, Progga Deb, Kevin Gallagher, Brendan Dolan-Gavitt, Siddharth Garg, and Rachel Greenstadt. Studying the Online Deepfake Community. *Journal of Online Trust and Safety*, 2(1), Sep. 2023.

[85] Rebecca Umbach, Nicola Henry, Gemma Beard, and Colleen Berryessa. Attitudes Towards and Knowledge of Non-Consensual Synthetic Intimate Imagery in 10 Countries, 2024.

[86] Nadia de Vries. "Porsche Girl": When a Dead Body Becomes a Meme. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, (Summer 2022), aug 26 2022. https://mit-serc.pubpub.org/pub/porsche-girl.

[87] Miranda Wei, Pardis Emami-Naeini, Franziska Roesner, and Tadayoshi Kohno. Skilled or Gullible? Gender Stereotypes Related to Computer Security and Privacy. In *IEEE Symposium on Security and Privacy*, 2023.

[88] Mika Westerlund. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11):9–16, 2019.

[89] Robin Whittemore, Susan K. Chase, and Carol Lynn Mandle. Validity in Qualitative Research. *Qualitative Health Research*, 11(4):522–537, July 2001.

[90] David Gray Widder, Dawn Nafus, Laura Dabbish, and James Herbsleb. Limits and Possibilities for "Ethical AI" in Open Source: A Study of Deepfakes. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, page 2035–2046, New York, NY, USA, 2022. Association for Computing Machinery.

[91] Kaylee Williams. Exploring Legal Approaches to Regulating Nonconsensual Deepfake Pornography, 2023.

[92] Rhiannon Williams. Text-to-image AI models can be tricked into generating disturbing images. *MIT Technology Review*, 2023.

[93] Andrea L. Wirtz, Tonia C. Poteat, Mannat Malik, and Nancy Glass. Gender-based violence against transgender people in the united states: A call for research and programming. *Trauma, Violence, & Abuse*, 21(2):227–241, 2020.

[94] Pamela J Wisniewski and Xinru Page. Privacy theories and frameworks. In *Modern Socio-Technical Perspectives on Privacy*, pages 15–41. Springer International Publishing Cham, 2022.

[95] Jay Zhangjie Wu, Yixiao Ge, Xintao Wang, Stan Weixian Lei, Yuchao Gu, Yufei Shi, Wynne Hsu, Ying Shan, Xiaohu Qie, and Mike Zheng Shou. Tune-A-Video: One-Shot Tuning of Image Diffusion Models for Text-to-Video Generation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7623–7633, October 2023.

[96] Zhiliang Xu, Zhibin Hong, Changxing Ding, Zhen Zhu, Junyu Han, Jingtuo Liu, and Errui Ding. Mobile-faceswap: A lightweight framework for video face swapping. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 2973–2981, 2022.

[97] Hanqing Zhao, Wenbo Zhou, Dongdong Chen, Tianyi Wei, Weiming Zhang, and Nenghai Yu. Multi-Attentional Deepfake Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2185–2194, June 2021.

[98] Douglas Zytko, Jane Im, and Jonathan Zong. Consent: A Research and Design Lens for Human-Computer Interaction. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*, pages 205–208, 2022.

[99] Carl Öhman. Introducing the Pervert's Dilemma: A Contribution to the Critique of Deepfake Pornography. *Ethics and Information Technology*, 22(2):133–140, June 2020.

A Full Survey Instrument

[Informed Consent] This is a survey about generative AI systems that can be used to create realistic-looking, but fake, images, videos, and audio of people. As part of this survey, we discuss different scenarios about this type of media. Be aware that some of these scenarios discuss people using such technology to create synthetic intimate images of you.

The [anonymized] reviewed our study and determined that it was exempt from federal human subjects regulation. We do not expect that this survey will put you at more risk than you might encounter in everyday conversations.

To participate, you must be at least 18 years old and able to complete the survey in English. We expect this survey will take about 8 minutes to complete. If you have any questions about this survey, you may email us at [anonymized].

I am 18 years or older. ◯ Yes ◯ No

I have read and I understand the information above. ◯ Yes ◯ No

I want to participate in this survey and continue with the task. ◯ Yes ◯ No

[Background, task description, & re-consent] A generative AI system is a type of artificial intelligence that learns patterns and relationships from a dataset of human-created content. When given a prompt, it uses this learned knowledge to generate new content in a similar style or format. This enables it to autonomously create realistic and coherent outputs, ranging from images to text.

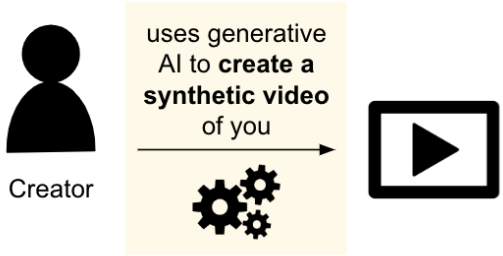
For example, a generative AI system might learn from a set of photographs taken by people. When prompted to generate an image of the sky, it would use patterns from this data set of photographs to generate the requested content.

These systems can be used to create realistic-looking, but fake, images and videos of people. For example, generative AI systems can create synthetic intimate images - images of people naked or engaged in sexual acts.

In this survey, we will show you 3 scenarios about synthetic, AI-generated videos. After presenting each scenario, we will ask you some follow-up questions. Then we will ask a series of questions about non-synthetic intimate imagery and other topics, including topics related to sex and sexuality. We will end the survey by asking general demographic questions.

I understand what this survey involves and would like to continue. ◯ Yes ◯ No

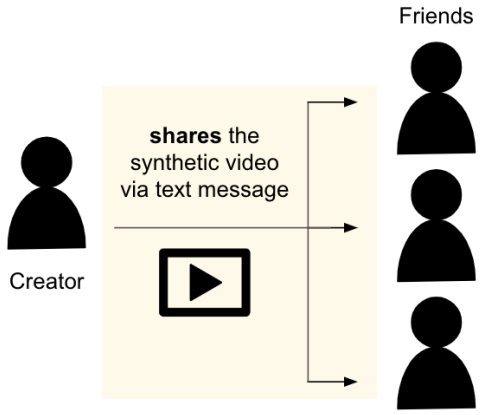
[Vignettes (this section is repeated three times with randomized vignettes)] Imagine that an intimate partner uses generative AI to create a synthetic video of you performing a sexual act for the purpose of sexual pleasure. Assume that you are unaware of the video's creation and existence.



In my opinion, the creation of this video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

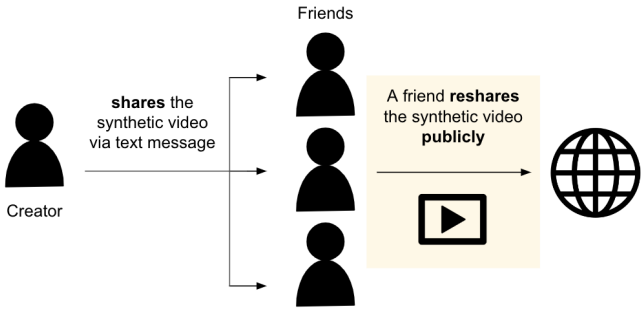
Please write at least 1-2 sentences about why you feel this scenario in particular is (un)acceptable. _____

Now, imagine that the creator of this video shares it with other people in a private channel, like a group chat with friends.



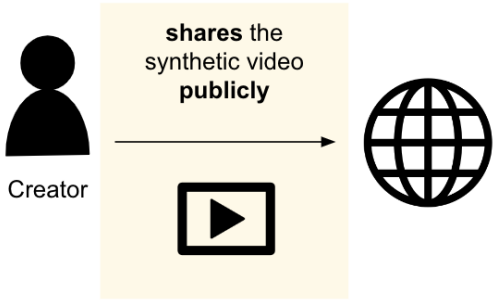
In my opinion, this sharing of the video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

Now, imagine that one of the people in the group chat, who received the video from the creator through the private channel, reshares the video publicly, like posting it on Reddit.



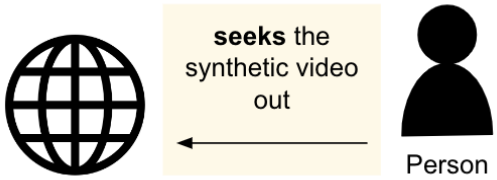
In my opinion, this resharing of the video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

We are now going to consider an alternative scenario. Imagine the creator of this video shares it in a public format, like posting it on Reddit.



In my opinion, this sharing of the video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

Now, imagine someone who has not created this video or had it shared with them seeks the video out, like by searching online for a video matching its description.



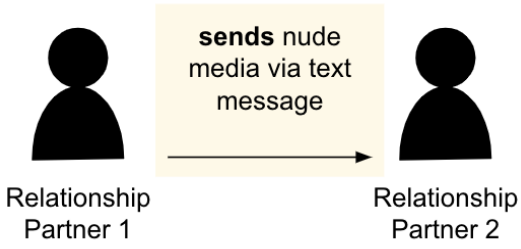
In my opinion, the seeking out of this video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

People take many different approaches to handling online content and experiences. If the video was shared in a public format, which of the following, if any, would you do: ☐ Report this content to the platform for removal ☐ Reach out to a helpline or support service organization for assistance getting the content removed ☐ Reach out to a lawyer ☐ Reach out to the police ☐ Reach out to friends or family ☐ Other _____ ☐ None of the above ☐ Prefer not to answer

This marks the end of questions about this scenario.

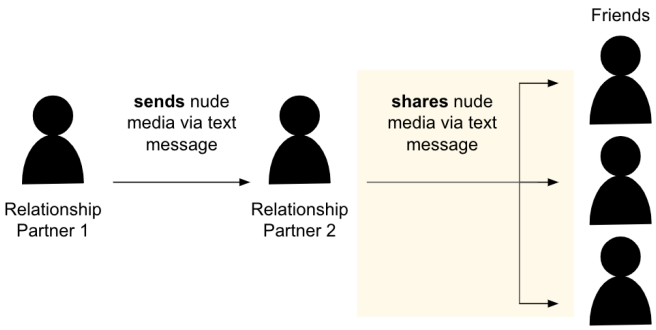
[Authentic intimate imagery questions] The following questions are about non-synthetic, intimate imagery.

Imagine two people in an intimate relationship (e.g., dating, married) send nude media (images or videos) of themselves to each other.



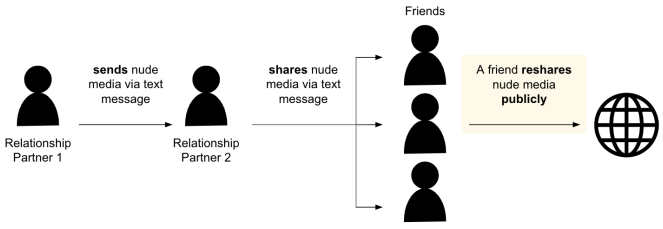
In my opinion this is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

Now, imagine that the intended recipient of this media shares it with other people in a private channel, like a group chat with friends, without informing the original sender.



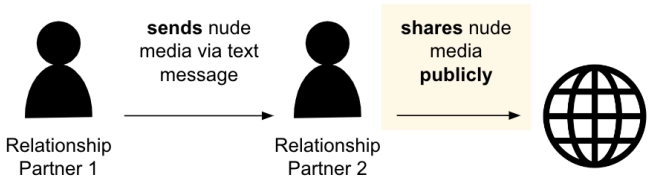
In my opinion, this sharing of the media is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

Now, imagine that one of the people in the group chat, who received the media from the intended recipient through the private channel, reshares the media publicly, like posting it on Reddit.



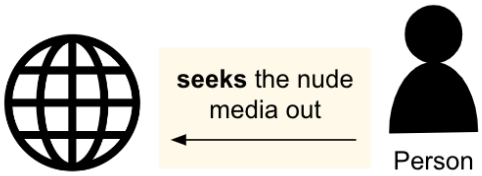
In my opinion, this resharing of the video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

We are now going to consider an alternative scenario. Imagine someone in the relationship shares nude media that they received from the other person in a public format, like posting it on Reddit.



In my opinion, this sharing of the media is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

Now, imagine someone outside of the relationship seeks the video out, like by searching online for media matching its description.



In my opinion, the seeking out of this video is: ◯ Totally unacceptable ◯ Somewhat unacceptable ◯ Neutral ◯ Somewhat acceptable ◯ Totally acceptable ◯ Prefer not to answer

People take many different approaches to handling online content and experiences. If you were the subject of this media and it was shared in a public format, which of the following, if any would you do: ☐ Report this content to the platform for removal ☐ Reach out to a helpline or support service organization for assistance getting the content removed ☐ Reach out to a lawyer ☐ Reach out to the police ☐ Reach out to friends or family ☐ Other _____ ☐ None of the above ☐ Prefer not to answer

Indicate your agreement with the following: "If a nude photo or video is shared with anyone, it will eventually be seen by people who were not the intended recipients." ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

[Deepfake community attitude agreement (these statements were sourced from a study of online deepfake communities [84] and were not used in analysis)] Realistic-looking fake porn can now be created using generative AI to swap the faces

of pornographic actors with other people's faces, so that it looks like they're in a porn video. This is sometimes referred to as 'deepfake pornography' as it looks very realistic and can be hard to recognize as fake or digitally created.
Imagine reading the following statement online: "Using generative AI to create pornography another person without their consent is acceptable." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "A person should be able to create deepfake pornography of anyone without their consent." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "Creating deepfake pornography of someone whose authentic nude media is publicly available does not cause additional harm." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "Violence is an act committed against a person that results in bodily harm, physical abuse is violence for example, sexual assault is violence, a video is not violence. Even a fake video of fake violence is not violence." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "Even clay modeling and cave paintings were used for porn. There is always resistance to novelty. But even if there might have been a temporary ban on those Arts, they still went on. Same here. We just gotta show some diversity." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "Technically, this is a cropping of one face onto another body. Has been done since decades. But for pictures. Now the pictures are moving." Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

Imagine reading the following statement online: "I'd be thrilled that someone found me attractive enough to be worth making a deepfake of lol!" Indicate your agreement with this statement: ◯ Strongly disagree ◯ disagree ◯ Neither agree nor disagree ◯ Agree ◯ Strongly agree ◯ Prefer not to answer

[Sexual Consent Scale - Revised Subscale 2: Positive attitudes towards establishing sexual consent]
Response options: ◯ Strongly disagree ◯ Disagree ◯ Somewhat disagree ◯ Neither agree nor disagree ◯ Somewhat agree ◯ Agree ◯ Strongly agree

1. I feel that sexual consent should always be obtained before the start of any sexual activity
2. I believe that asking for sexual consent is in my best interest because it reduces any misinterpretations that might arise
3. I think it is equally important to obtain sexual consent in all relationships regardless of whether or not they have had sex before
4. I feel that verbally asking for sexual consent should occur before proceeding with any sexual activity
5. When initiating sexual activity, I believe that one should always assume they do not have sexual consent
6. I believe that it is just as necessary to obtain consent for genital fondling as it is for sexual intercourse
7. Most people that I care about feel that asking for sexual consent is something I should do
8. I think that consent should be asked before any kind of sexual behavior, including kissing or petting
9. I feel it is the responsibility of both partners to make sure sexual consent is established before sexual activity begins
10. Before making sexual advances, I think that one should assume "no" until there is a clear indication to proceed
11. Not asking for sexual consent some of the time is ok

Scoring instructions: Strongly disagree = 1, Disagree = 2, Somewhat disagree = 3, Neither agree nor disagree = 4, Somewhat agree = 5, Agree = 6, Strongly agree = 7. Reverse item 11 (1 = 7, 2 = 6, 3 = 5, 4 = 4, 5 = 3, 6 = 2, 7 = 1) and average all items.

[Sexual Consent Scale - Revised Subscale 4: Sexual consent norms]
Response options: ◯ Strongly disagree ◯ Disagree ◯ Somewhat disagree ◯ Neither agree nor disagree ◯ Somewhat agree ◯ Agree ◯ Strongly agree

1. I think that obtaining sexual consent is more necessary in a new relationship than in a committed relationship

2. I think that obtaining sexual consent is more necessary in a casual sexual encounter than in a committed relationship

3. I believe that the need for asking for sexual consent decreases as the length of an intimate relationship increases

4. I believe it is enough to ask for consent at the beginning of a sexual encounter

5. I believe that sexual intercourse it the only sexual activity that requires explicit verbal consent

6. I believe that partners are less likely to ask for sexual consent the longer they are in a relationship

7. If consent for sexual intercourse is established, petting and fondling can be assumed

Scoring instructions: Strongly disagree = 1, Disagree = 2, Somewhat disagree = 3, Neither agree nor disagree = 4, Somewhat agree = 5, Agree = 6, Strongly agree = 7. Average all items.

[Demographics] What is your age? _____

What is your gender? _____

In politics today, do you consider yourself a Republican, a Democrat, an Independent, or something else? ◯ Republican ◯ Democrat ◯ Independent ◯ Don't know ◯ Refuse to answer ◯ Other

(If answered 'Independent,' 'Don't know,' 'Refuse to answer,' or 'Other') As of today, do you lean more to the Republican Party or the Democratic Party? ◯ Republican ◯ Democrat ◯ Refuse to answer

[Attention check] What would you like to see elected leaders in Washington get done during the next few years? Please give as much detail as you can. _____

[Debrief] If you or someone you know experiences or has experienced non-consensual intimate image abuse, support is available. Visit https://StopNCII.org/ and https://cybercivilrights.org/ccri-safety-center/ for comprehensive resources and information. Additionally, you can contact the National Domestic Violence Hotline [1(800)799-SAFE (7233)] for confidential assistance and guidance. Select "Next Page" to complete the survey and be redirected to Prolific.

B Participant demographics

Participants' gender, age, and political orientation is presented in Table B.

C Media action regression results

Results for the regression are presented in Table 6.

D Qualitative Codebook

The codebooks from qualitatively analyzing explanations for why the creation of the synthetic video in each vignettes is either acceptable or unacceptable. Codes were not mutually exclusive.

Rationales for acceptability

No Harm: Will not cause harm

Relationship: Trust in an intimate partner

Indifference: No impact; 'I don't care'

Compliment: Indicates attraction

Fantasy: Indulges fantasy

Pro-Tech: Technology and AI are interesting

Humor: 'It's funny'

Dislike: Although acceptable, elicits negative feelings; The video is 'weird,' 'creepy,' 'disgusting,' 'uncomfortable,' etc.

Rationales for unacceptability

Consent: Absence of consent or permission

Awareness: Lack of awareness about video's creation and existence

Dislike: Elicits negative feelings; The video is 'weird,' 'creepy,' 'disgusting,' 'uncomfortable,' etc.

Harm: Creates or could create harm

Ethics: Violation of ethics, morality, or law; The video is 'wrong'

Privacy: Violation of privacy

Fake: Fake nature, inauthentic

Stranger: Created by a stranger

Relationship: Violation of trust in an intimate partner

E Additional Figures

Odds ratios and confidence intervals for the creation, private_sharing, public_sharing, resharing, and seeking_out models are visualized in Figure 4. Figure 5 depicts heatmaps comparing acceptability involving AIG-NCII and GII created in an intimate partnership. Acceptability of creation and private_sharing across all vignettes and compared between creators is visualized in Figure 6. Figure 7 and Figure 8 visualize the distribution of codes for justifications of acceptable and unacceptable creation, respectively, across vignettes and actions.

Gender		Age		Political Orientation	
		18-24	17.8%		
Woman	49.5%	25-34	33.0%	Democrat	48.6%
Man	47.6%	35-44	24.4%	Republican	16.2%
Non-binary	1.9%	45-54	13.3%	Leans Democrat	18.4%
Agender	0.6%	55-64	7.9%	Leans Republican	8.9%
Prefer not to say	0.3%	65+	2.9%	Refuse to answer	7.9%
		Prefer not to say	0.6%		

Table 5: Breakdown of participant demographics by gender, age, and political orientation.

		OR; Confidence Interval
Intercepts	Totally unacceptable | Somewhat unacceptable	2.42; [1.89, 3.1]***
	Somewhat unacceptable | Neutral	7.41; [5.73, 9.59]***
	Neutral | Somewhat acceptable	28.59; [21.65, 37.76]***
	Somewhat acceptable | Totally acceptable	89.53;[65.78, 121.85]***
Content Action	private_sharing	0.47; [0.37, 0.58]***
	public_sharing	0.26; [0.21, 0.33]***
	resharing	0.42; [0.33, 0.52]***
	seeking_out	5.43; [4.45, 6.62]***

Table 6: Results from a single regression exploring the relationship between acceptability (first row, intercepts) and action being preformed with the synthetic media (second row, content action). Reference level of content action is creation. Significance of OR: $p < 0.001$ = *** .

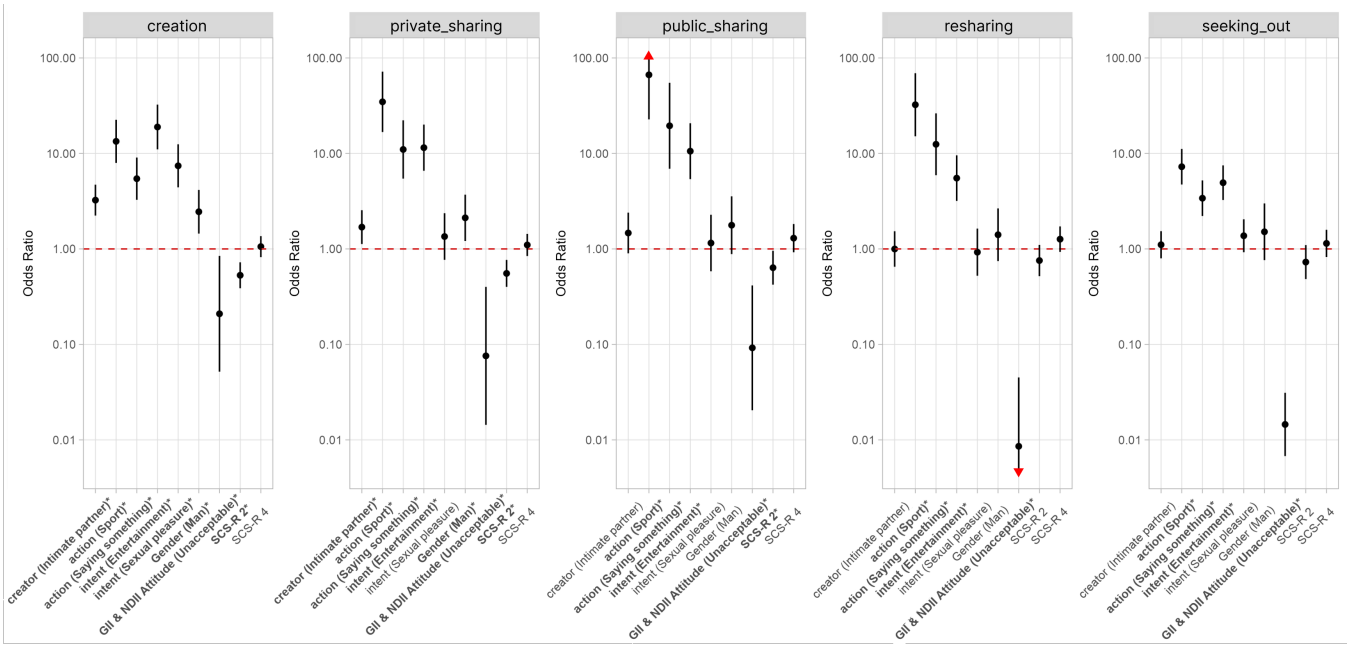


Figure 4: Odds ratios and confidence intervals for the controlled and uncontrolled IVs in the models for creation, private_sharing, public_sharing, resharing, and seeking_out. Statistically significant IVs ($p < 0.05$) are in bold with an asterisk. Y-axis is on a logarithmic scale and trims are indicated by red triangles.

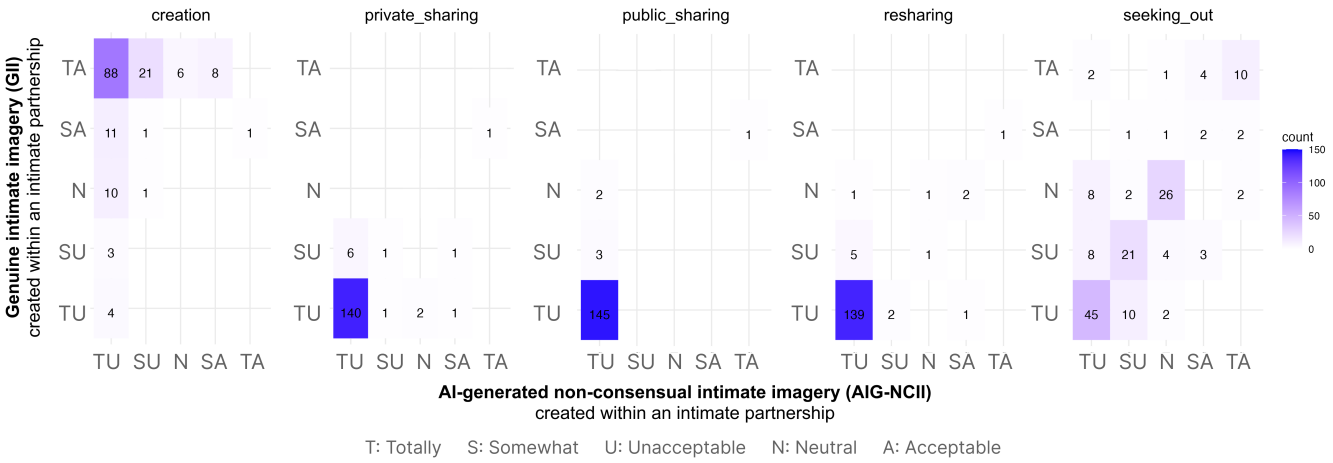


Figure 5: Heatmap of acceptability for creation and seeking_out when the action is performing a sexual act. Darker indicates more respondents.

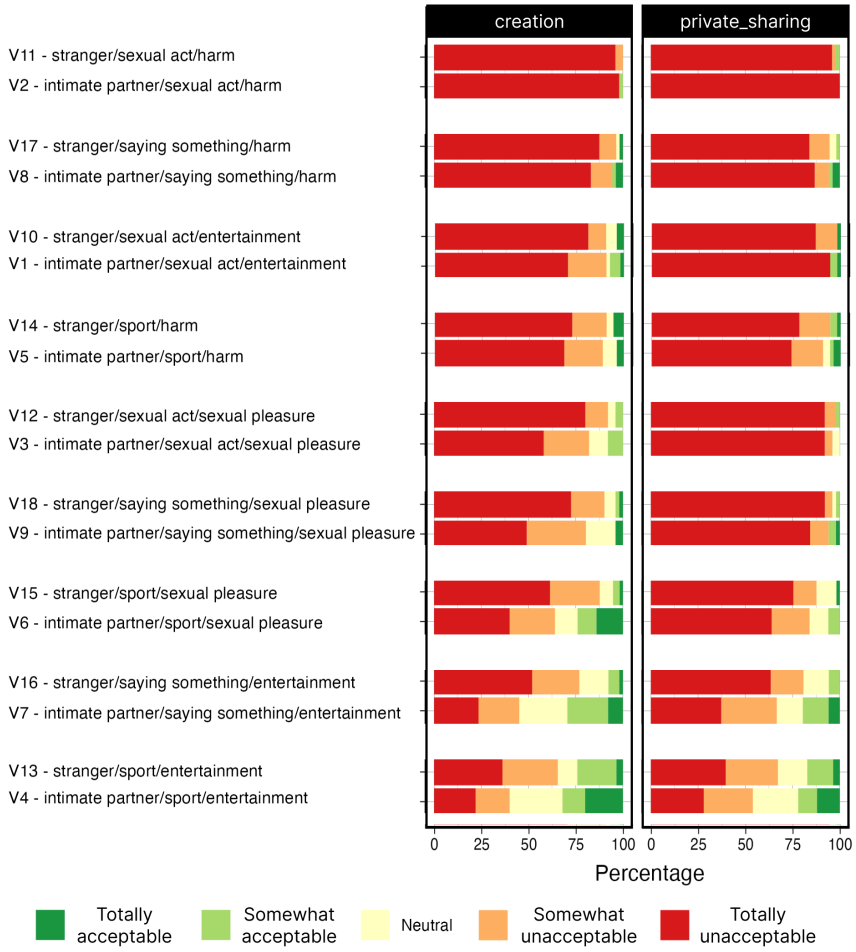


Figure 6: Participants' perception of acceptability for creation and private_sharing across all vignettes. Vignettes are defined by the creator/action/intent and grouped by vignettes with the same action and intent.

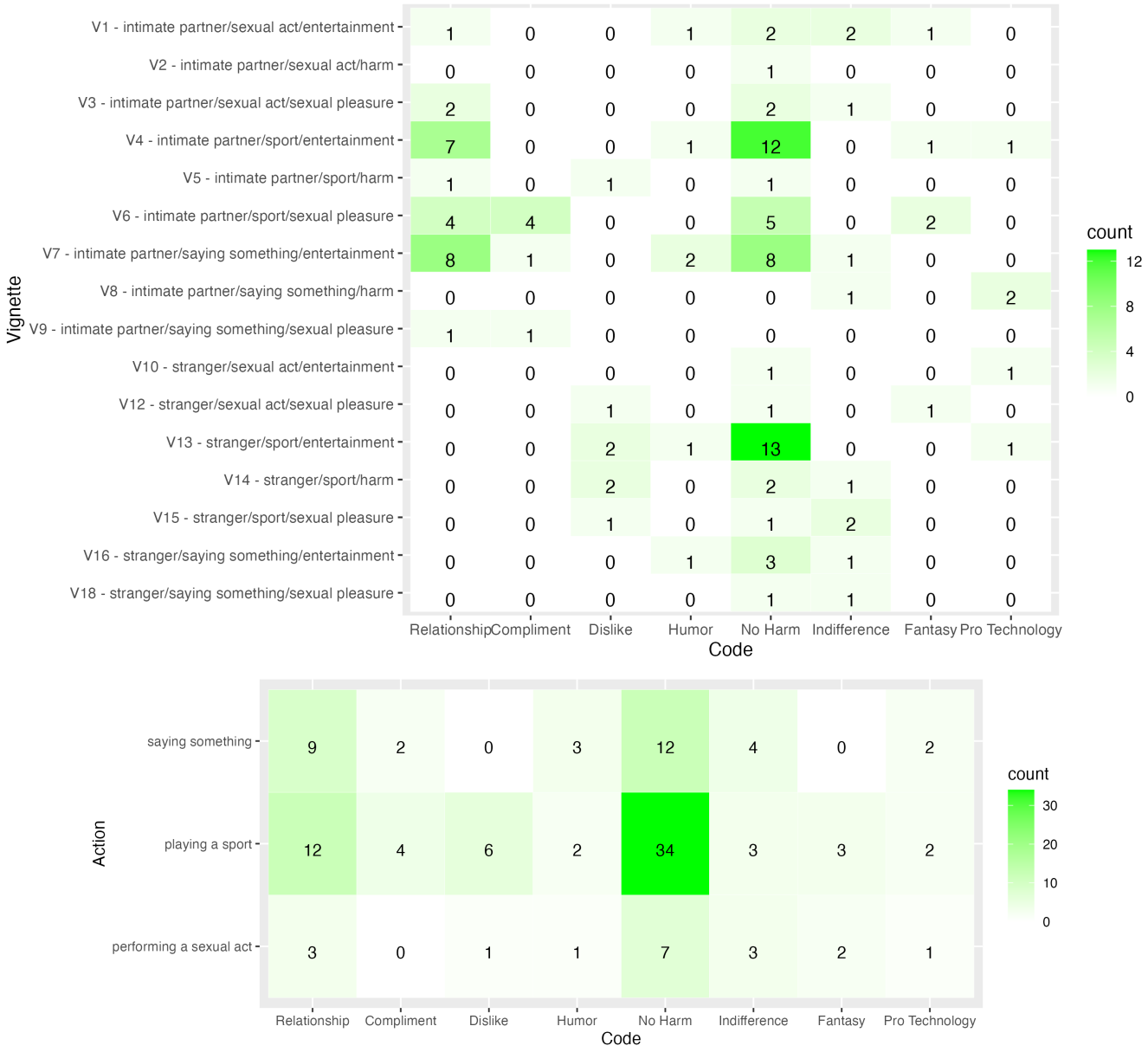


Figure 7: Heatmaps of the frequency of codes for justifications of acceptable creation. The top heatmap shows occurrences for each vignette and the bottom heatmap shows occurrences across actions. Darker green indicates a higher count of occurrences in the qualitative data.

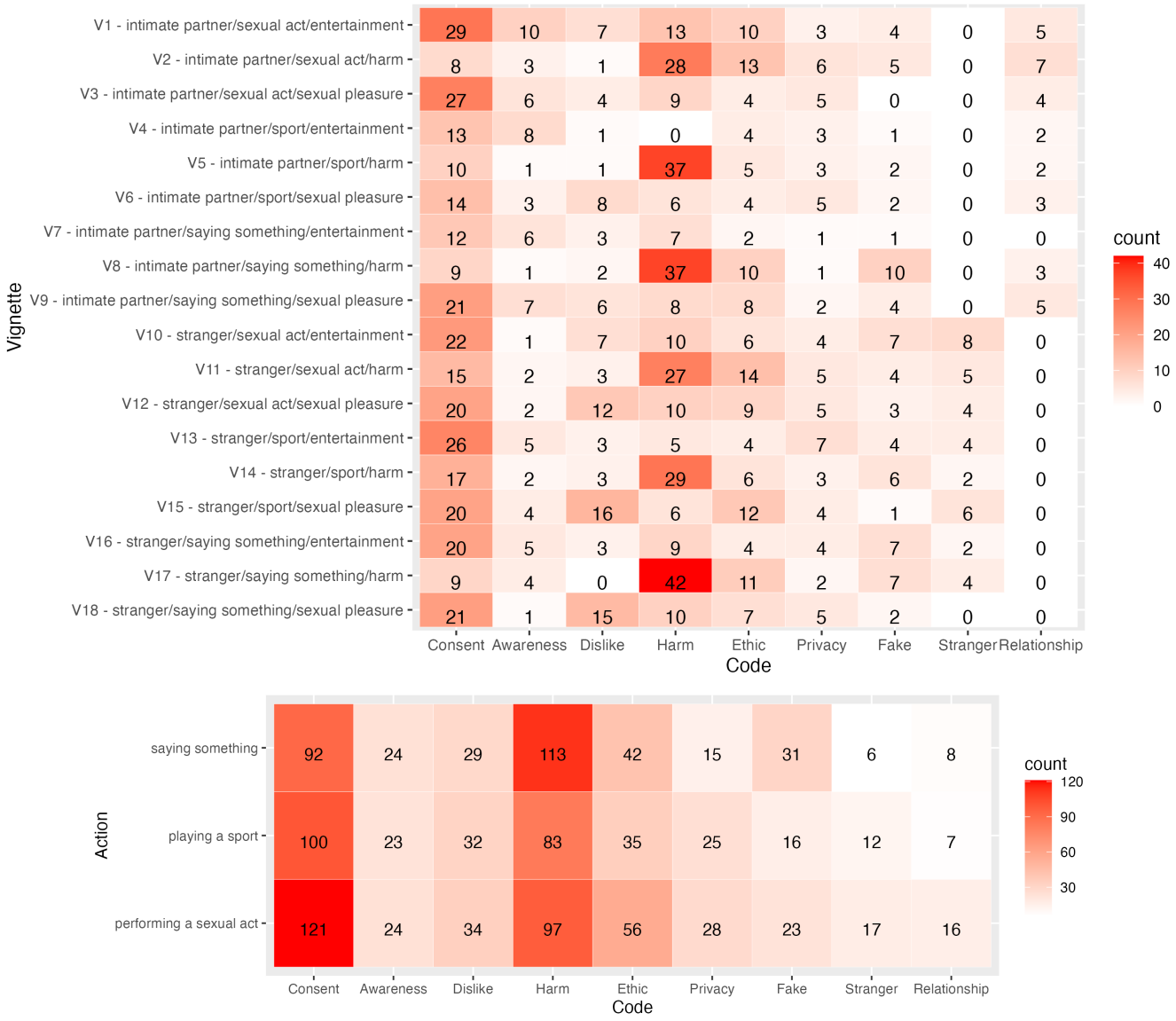


Figure 8: Heatmaps of the frequency of codes for justifications of unacceptable creation. The top heatmap shows occurrences for each vignette and the bottom heatmap shows occurrences across actions. Darker red indicates a higher count of occurrences in the qualitative data.