Bottom-up psychosocial interventions for interdependent privacy: Effectiveness based on individual and content differences

Renita Washburn
University of Central Florida
renitawashburn@knights.ucf.edu

Tangila Islam Tanni
University of Central Florida
tanni@knights.ucf.edu

Yan Solihin
University of Central Florida
Yan.Solihin@ucf.edu

Apu Kapadia
Indiana University
kapadia@indiana.edu

Mary Jean Amon
University of Central Florida
MJAmon@ucf.edu

ABSTRACT

Although a great deal of research has examined interventions to help users protect their own information online, less work has examined methods for reducing interdependent privacy (IDP) violations on social media (i.e., sharing of other people's information). This study tested the effectiveness of concept-based (i.e., general information), fact-based (i.e., statistics), and narrative-based (i.e., stories) educational videos in altering IDP-relevant attitudes and multimedia sharing behaviors. Our study revealed concept and fact videos reduced sharing of social media content that portrayed people negatively. The narrative intervention backfired and increased sharing among participants who did not believe IDP violations to be especially serious; however, the narrative intervention decreased sharing for participants who rated IDP violations as more serious. Notably, our study found participants preferred narrative-based interventions with real world examples, despite other strategies more effectively reducing sharing. Implications for narrative transportation theory and advancing bottom-up (i.e., user-centered) psychosocial interventions are discussed.

CCS CONCEPTS

• **Security and privacy**; • **Human and societal aspects of security and privacy**; • **Social aspects of security and privacy**;

KEYWORDS

behavior change, interdependent privacy, psychosocial intervention, social media/online communities

ACM Reference Format:
Renita Washburn, Tangila Islam Tanni, Yan Solihin, Apu Kapadia, and Mary Jean Amon. 2023. Bottom-up psychosocial interventions for interdependent privacy: Effectiveness based on individual and content differences. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3544548.3581117

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581117

1 INTRODUCTION

Interdependent privacy (IDP) violations occur when people share potentially sensitive information about others without their consent [1]. Multimedia IDP violations are one particularly common type that involves social media users sharing other people's photos [2] or images online [1], which may include private information about others such as drug usage, location, medical history, sexual history, or other embarrassing or personal information [3, 4]. Photos can reach unintended audiences [5] and migrate across platforms as they are shared and re-shared, such that the person depicted no longer has control over their online image. Photos can also undergo 'context collapse' [6] as they are modified or taken out-of-context with accompanying text captions. Although some content may be shared maliciously, oftentimes potential IDP violations are the product of misalignments between people's privacy preferences, misunderstandings, posting while emotional, or underdeveloped privacy standards [7]. IDP violations can have a variety of personal and professional consequences, ranging from psychological distress, harassment, job loss, stalking, and damaged interpersonal relationships [8, 9].

Although a great deal of research has examined nudges and educational interventions to help users protect their own information online [10-13], considerably less work has been dedicated to encouraging users to reconsider the ways in which they share other people's photos and information in social media. Moreover, most IDP-relevant interventions are of a technical nature [1] in that they must be implemented within a particular user interface and are 'top-down' in typically requiring buy-in from social media platforms, which can, in turn, constrain scalability. Little research has focused on 'psychosocial' interventions for IDP preservation, or interventions that rely on psychological and social processes to promote particular courses of action. Psychosocial interventions have the advantage of targeting change from the 'bottom-up' (i.e., user-centered change), influencing change at the source, and empowering communities to leverage their collective social norms and values to enact change [14]. Accordingly, there is a significant need for experimentally rigorous research to change the status quo of online sharing, especially to limit sharing of other people's sensitive images and information in social media.

Limited research on psychosocial interventions for IDP preservation has shown privacy prompts and perspective-taking nudges to be ineffective or even backfire to increase the sharing of other people's information [15]. However, there is a path forward, as interventions in the public health and computer science sectors

point to the utility of public messaging as scalable solutions to change attitudes and behaviors [16]. In particular, video interventions have been demonstrated as effective for promoting positive social behavior [17] and internet usage [18]. But the success of video interventions is ultimately driven by their content and the context in which the interventions are applied [19]. Fact-based (i.e., statistics) and narrative-based (i.e., stories) content is especially common, with relative effectiveness varying considerably based on context and topic [20-23]. Research is needed to investigate the types of intervention content that effectively reduce photo and other multimedia based IDP violations.

Although some psychosocial interventions for public welfare have large-scale impact [24, 25], other psychosocial interventions are ineffective [26] and even backfire [15, 27, 28]. Such failures may be due to the interventions' inadequacies in properly anticipating various effects based on individual recipient characteristics [29-31]. Thus, a more complete understanding of intervention effectiveness requires consideration of individual and social media content differences that modulate recipient responses. For example, social media users are more likely to share positively-valenced information about other people, suggesting people consider these instances of sharing to be relatively acceptable even if the information is being shared without permission [15]. However, to our knowledge, the significance of individual and content differences in modulating IDP intervention effectiveness has not been tested previously.

We investigated the extent to which concept-based (i.e., explanation of IDP), fact-based (i.e., statistics about IDP prevalence and consequences), and narrative-based (i.e., stories about people effected by IDP violations) video interventions were effective in altering IDP-relevant photo-sharing attitudes and behaviors compared to a control condition with no intervention. We utilized a mix of quantitative and qualitative methods to examine how the following changed based on intervention type: 1) attitudes toward IDP as assessed via survey, 2) sharing likelihood of photo-based memes depicting potential IDP violations, and 3) qualitative responses to the interventions. In addition, we examined whether intervention effectiveness in changing meme sharing decisions varied based on individuals' beliefs in the seriousness of IDP violations, as well as the valence of each social media meme. Valence was rated by an independent group of participants, who coded each photo-based meme for how positively or negatively the photo target was portrayed.

In doing so, the proposed project examined the overarching research question of how psychosocial interventions can promote IDP preservation in social media. To answer the overarching research question, four sub-questions were addressed:

- RQ1. To what extent do users' IDP-relevant photo-sharing attitudes change in response to the interventions?
- RQ2. To what extent do users' IDP-relevant photo-sharing decisions change in response to the interventions?
- RQ3. To what extent is an intervention's effectiveness in altering photo-sharing decisions modulated by social media content differences in how meme targets are portrayed?
- RQ4. To what extent is an intervention's effectiveness in altering photo-sharing decisions modulated by individual differences in IDP seriousness ratings?

Our research complements previous literature by examining video-based educational interventions for IDP preservation, representing a crucial step in advancing empirically-validated educational programs to alter photo-sharing social media sharing norms. We developed new and varied intervention content that can be readily disseminated to the general public and adopted a mixed analytical approach to understanding intervention preferences, IDP attitudes, and behaviors following the experimental manipulation. This research is also among the first IDP intervention study to account for individual and contextual differences that may alter recipient responses. Thus, we examine both normative and variable behavioral responses to the three intervention types to provide a more holistic account of intervention effectiveness, including conditions under which the interventions decrease sharing of other people's information versus backfire to increase sharing.

2 BACKGROUND

2.1 Interdependent privacy

IDP refers to the ways in which people's privacy not only depends on their own actions but also the actions of others [1]. IDP violations can include a wide range of behaviors, ranging from the sharing of genomic data, which can have implications for even distant relatives, or sharing location data that happens to reveal the location of others nearby [1]. An especially common form of IDP violation occurs with the sharing of multimedia data that often has images or information about individuals other than the person who captured, shared, or re-shared the information online [1]. Beyond including identifying information about other people (e.g., image of face or audio of voice), multimedia content can also include potentially personal information about the featured individuals. Definitions of what constitutes 'private' content varies. For example, some work includes information about a person's gender, political or religious orientation, race or ethnicity, interests, and relationships to be a potential IDP violation when shared without permission [1, 32-34]. In addition, potential IDP violations can include information such as drug activity, medical history, sexual history, embarrassing moments, or shaming people for non-normative behaviors [35].

Given that individuals vary in their personal privacy preferences [15], it can be difficult for users to anticipate the potential sensitivity or subsequent harm that may follow from sharing multimedia content depicting others. For instance, some parents post photos of their children on social media sites, possibly leaving damaging digital footprints far before the age of consent, while other parents consider the risks carefully before sharing photos online. As these children grow older and view the images their parents post, some may experience nostalgia and delight, whereas others may feel misunderstood and embarrassed of even seemingly benign images of themselves. Similarly, whereas some people pursue lifestyles as public influencers, other people actively minimize their digital footprint [36, 37]. For this reason, many instances of sharing another person's photos and other information online may be seen as a *potential* IDP violation, where the impact of sharing is not immediately ascertainable.

Even though people generally report they prefer to grant permission before being posted about by others [38], most social media

users report that they have learned by chance about photos of themselves online [39]. It has been estimated that more than one-fourth of social media users have requested for another user to remove a shared picture, highlighting that many such posts are unwanted [40]. Although IDP violations may sometimes be associated with 'internet trolls,' or people who purposefully seek to provoke or harm others online [41], social media users are often most concerned about the posting behaviors of close friends, family members, and employers rather than strangers [42-44].

A portion of online content is co-opted to become popular internet 'memes,' or units of culture spread throughout a society and often shared in social media [45, 46]. As memes are shared and reshared, the meme 'target,' or person depicted in the photo, may feel helpless to stop the spread of their own image [47], leading to an increase in privacy concerns [48, 49]. This type of viral sharing can lead to a disconnect between the originally-intended audience and actual audience [5, 50]. Iterations of sharing are also accompanied by changes to the original image or caption, altering the original context and portrayal of the meme target (i.e., 'context collapse [6, 51, 52]). In this way, photo-based memes represent particularly salient examples of potential IDP violations, in that they include personally identifying information (e.g., the meme target's face), are spread widely, out of the control of the meme target, and it is typically unclear if the meme target consented to their photo being used in this manner. Moreover, the viral nature of photo-based memes demonstrates the scale at which multimedia IDP violations occur.

2.2 Interventions for interdependent privacy preservation

Many of the researched methods for IDP preservation are in reference to users' existing mitigation strategies. Users' strategies for preventing IDP violations include avoiding appearances in undesirable photos, filtering friend requests, maintaining anonymity online, minimizing social media interactions, posting to select audiences, personally defining posting rules, strategic tagging or untagging, and 'unfriending' users who tend to compromise others' privacy [42, 53-58]. When regulating content that has already been posted, individual users tend to delete associated comments and untag or disconnect themselves from the privacy-violating content [56]. Thus, users engage in a wide variety of individual behaviors aimed at protecting their information from undesired sharing by others.

Given the interpersonal nature of IDP violations, many user-adopted strategies for maintaining privacy also require collaboration with others. Collaborative IDP-preservation efforts include users proactively seeking permission from others before posting, educating other users, discussing settings, and negotiating sharing rules with people outside of social media platforms [42, 59, 60]. Following an IDP violation, users often request that other people delete the content [56]. Collaborative strategies can have the benefit of improving interpersonal relationships [42], fostering social cohesion, and supporting privacy management as a collective effort [59]. Despite these benefits, research suggests that users prefer controlling their own posting behaviors, versus trying to influence others' decisions [56]. Moreover, users significant use of side-channels and other makeshift solutions to enforce personal privacy preferences highlights a lack of formal mechanisms for preventing and managing potential IDP violations [58].

Even though research examining non-technical IDP interventions largely examines existing individual user strategies, there are some exceptions. Amon and colleagues [15] compared two nudges - or aspects of a choice architecture intended to alter people's behavior [5] - to a control condition to examine whether subsequent sharing of memes with potential IDPs lessened. A series of studies revealed that, whereas a nudge to adopt the perspective of the meme target was ineffective, prompts to consider the target's privacy consistently backfired to increase sharing. A follow-up study demonstrated that participants exposed to the privacy prompt intervention went on to note that the memes were 'not private' 15 times more than in the control condition. Thus, apparently straightforward methods for IDP preservation can backfire [61], emphasizing the importance of empirical testing. Indeed, prior educational and nudge research highlight the perils of inadequate testing, including the potential for wasted resources (e.g., $1.3 billion D.A.R.E program [26]) or counter-intuitive effects that invalidate the intervention altogether (e.g., pregnant women drink more after viewing reminders to not drink [28]), privacy education interventions can negatively interact [27] or backfire [61].

Although the present work focuses on non-technical solutions, it is worth noting that a variety of technical mechanisms for promoting IDP preservation have been explored with great promise. For example, users tend to prefer obfuscation methods that mask people's identities with avatars or full removal, as compared to face blurring, as the former methods maintain visual appeal and continuity of the original photo [62]. Xu and colleagues [63] propose a face-recognition algorithm that also discriminates based on graph neighbors and their relative closeness to the person depicted in multimedia content, with the goal of informing people about potential IDP violations. Additional solutions have been proposed that include voting schemes and privacy-preference algorithms [64, 65], access control strategies that allow for posting after all parties consent or vote [66, 67], shared control over the allowed audience [68], and a conflict-resolution model [69]. Although a full review is out of the scope of the present work (see [1] for review), technical solutions are currently used for large-scale content moderation (e.g., algorithmic approaches in social media), with additional effective solutions on the horizon.

In summary, research examining strategies for IDP preservation typically center on users' pre-existing personal strategies, whereas research testing new IDP interventions largely focuses on technical solutions. In contrast, research on psychosocial strategies, or those that capitalize on psychological and social processes to promote particular courses of action, are highly limited. To date, psychosocial interventions for IDP preservation have produced counterintuitive effects by backfiring to increase sharing of others' potentially personal information [15]. However, computer privacy and other interdisciplinary research highlight the large-scale benefits of psychosocial interventions.

2.3 Bottom-up psychosocial interventions

Technical strategies for mitigating IDP violations often require 'top-down' (i.e., corporate or government [70]) commitment, implementation, and management. That is, corporate buy-in is typically needed to reach users with new interface features (e.g., blocking, tagging, or unfollowing), technical applications (e.g., tools that blur faces), and algorithmic approaches (e.g., for detecting and mitigating existing violations). Thus, strategies that inherently rely on top-down implementation have limited scalability. For instance, even if a new obfuscation tool proves effective in reducing IDP violations, there is no guarantee that one or more social media platforms or other applications will adopt the tool in its most privacy-preserving form. Although top-down solutions are essential to supporting IDP preservation, this limitation highlights the need for complementary approaches.

Interventions that target change from the 'bottom-up' (i.e., user-centered change) can be disseminated to users and by users, influencing change at the source and empowering communities to leverage their collective social norms and values to enact change [14]. Bottom-up strategies are often psychosocial in that they capitalize on psychological and social processes to promote particular courses of action. Bottom-up psychosocial interventions have the benefit of being preventative as a 'primary intervention' strategy, versus reacting to existing violations as a 'secondary intervention' strategy [71, 72]. Bottom-up psychosocial interventions are also scalable in that they can be efficiently disseminated to large audiences and have been highly efficacious in addressing public health concerns (e.g., anti-smoking campaigns [24, 73] and social issues (e.g., 'me too' movement [74]). These interventions promote change among a critical mass of people as information spreads, altering broader societal norms [75, 76]. Notably, social media users recruited to co-design solutions to multiparty privacy conflicts have highlighted users' interests in preventative strategies aimed at educating users about community standards [77].

Common types of psychosocial interventions include educational interventions and nudges. Educational interventions have been explored in the realm of privacy and security, but this research usually centers on encouraging users to protect their own privacy. Educational interventions often seek to inform users about the data being collected about them in social media and how it may be used to make inferences about their characteristics and behaviors [78, 79]. Educational interventions differ in the amount and type of information offered to users, such as crowd-sourced user setting recommendations [27], or information about how social media posts might be perceived by others [80]. Online privacy education recommendations also include incorporating relatable stories, enhancing users' decision-making abilities, and conveying a range of privacy consequences [81]. Despite advancements in this area, the effects of educational interventions vary and sometimes in unexpected ways. For example, crowd-sourced setting recommendations and self-reflection both help users adopt more stringent privacy settings, but, when combined, these two strategies negatively interact [27].

Nudges are another common psychosocial avenue for supporting users' personal security and privacy in social media. Contrasting educational approaches aimed to inform and enhance awareness, nudges focus on changing the decision-making environment. A review and meta-analysis of papers utilizing nudge interventions to alter personal information sharing [5, 10, 77, 82] revealed no statistical differences in effectiveness between presentation, information, default setting, or incentive nudges, suggesting a general benefit of nudging to promote personal privacy, though nudges to increase disclosure were more effective than those intended to reduce disclosure [10]. In the realm of IDP, nudge research indicates that social media users are more likely to use tagging settings when framed positively or as default options [11]. Similarly, precautionary mechanisms that force users to collaborate by default are generally preferred to dissuasive mechanisms aimed at deterring uploaders from sharing without consent [82]. However, in contrast to findings by Anaraky et al. [11], Masaki et al. [12] demonstrate that negative framing of risky choices is more effective than positive framing, with conflicting findings suggesting the context sensitivity of nudge effectiveness.

Given that most privacy-oriented psychosocial interventions focus on limiting the degree to which people share their personal information, there remains a significant need for experimentally rigorous research for improving users' sharing decisions as they pertain to interpersonal information. A significant gap in the privacy literature was identified by Pinter and colleagues [83] in their review of 132 privacy articles: few studies have gone beyond identifying self-reported privacy attitudes to establish novel intervention strategies to change the status quo. Social media users already make some efforts to protect their own [60] and other's privacy [84] but demonstrate generally "lax attitudes" toward interpersonal privacy violations [43], raising questions about how to encourage more responsible sharing. Citing limited regulations and policies to protect users from IDP violations, Kamleitner and Mitchell [85] propose a framework to promote IDP preservation through 'the 3Rs:" Realize (some data is transferred to a third party), Recognize (data has privacy implications for others) and Respect (others' rights). Additionally, Kamleitner and Mitchell [85] outline strategies for combatting IDP violations, which include 'educating for respect.'

Although work on psychosocial interventions to promote IDP preservation is limited, significant research on bottom-up psychosocial interventions highlights the potential of narratives and facts to alter public behavior. In particular, emotional and personal narratives are powerful sources of influence in health promotion and disease prevention and may be especially relevant to preventing interpersonal privacy violations within highly interconnected social networks, where individuals may be both influenced by narratives and, in turn, participate in influencing others within their network. Humans are natural story tellers [86], using narratives to communicate information about characters experiencing different situations often with specific contexts, goals, and intentions [87]. A wide range of literature in advertising, entertainment education, and health communication demonstrates the power of interventions employing personal and emotional stories in altering public attitudes and behaviors due to their ability to engage and encourage meaning making [88, 89]. Such narrative-based interventions are especially effective when high in emotion, exemplifying healthy behaviors, advocating prevention (vs. cessation), and illustrating cause-and-effect [90, 91].

However, the utility of narrative persuasion compared to non-narrative persuasion in altering public attitudes and behaviors is

complicated [92]. Studies comparing fact-based versus narrative interventions indicate that their relative effectiveness can vary based on topic [20, 22, 23, 93], with statistical evidence sometimes surpassing narratives in effectiveness [94, 95]. In the realm of computer science, a number of studies have examined stories to improve users' personal security behaviors online, with cyber security research indicating that stories are effective in improving security setting use and reducing personal disclosures [86]. However, stories may be less effective than fact-based interventions in some settings like those aimed at improving phishing detection [87]. Thus, the effectiveness of narrative interventions, as compared to fact interventions, varies considerably based on context.

In addition to effectiveness varying based on intervention type, effectiveness may differ based on users' individual differences. For example, interventions to reduce electricity bills were more effective in liberal versus conservative households [29] and efforts to increase tax reporting by emphasizing social norms were effective except for with those carrying the most debt [30]. In the realm of online privacy, Peer et al. [31] investigated how responses to privacy nudges could be improved through personalization, finding that nudges that differed based on users' decision-making style enhanced efficacy. The effectiveness of different types of bottom-up psychosocial interventions, as well as how effectiveness is modulated by individual differences and content differences, remains an open question. Researchers have also identified user preferences on types of nudges vary based on sharing habits gender, age [82], perceived risks from sharing [12], and valence of content [11].

2.4 The present study

The present study extends empirical testing of bottom-up psychosocial interventions to the IDP literature, leveraging educational videos with the aim of supporting change in user's IDP-relevant attitudes and behaviors and, in turn, reducing privacy violations. Intervention strategies build on prior literature that suggests the varied utility of fact- and narrative-based intervention in other applications. Additionally, each intervention built on findings from nudge-based interventions regarding effectiveness of dissuasive mechanisms (i.e., cautioning and discourage sharing content without permission) to reduce IDP violations [11, 77, 82] by promoting collaboration between the sharer and media subject [82]. Specifically, we compared four groups of participants: 1) Control group who did not receive an intervention, 2) Concepts intervention group who learned about the concept of IDP violations or—as we termed it in the videos—'privacy pirating,' 3) Facts intervention group who were provided statistics about the prevalence and consequences of IDP violations, in addition to being introduced to the concept of privacy pirating, and 4) Narratives intervention group who were shown two stories about people negatively affected by IDP violations, as well as the concept of privacy pirating. All three experimental groups were provided information about the concept of 'privacy pirating,' as this provided necessary context for subsequent facts or narrative videos. A mix of quantitative and qualitative methods were utilized to assess attitudes and decisions to share potential IDP violations following the interventions, including how sharing decisions following the interventions were modulated based on individual differences and features of the social media content. In addition, we qualitatively analyzed participants' feedback regarding the interventions to determine the extent to which the psychosocial interventions for IDP preservation were perceived as effective and appealing.

Considering narrative transportation theory and prior research highlighting the effectiveness of narrative-based interventions, including in reducing personal disclosures online [86], we hypothesized that emotional stories depicting the consequences of IDP violations would be especially well-received by users and effective in reducing subsequent sharing of other people's information (H1). We further hypothesized that concept- and fact-based interventions would reduce sharing, compared to the control condition (H2), given that users may not otherwise have much awareness of IDP violations or 'privacy pirating.'

Given that intervention effectiveness is likely to vary based on individual and content differences [29-31], it follows that features of social media content may modulate intervention effectiveness as well. That is, psychosocial interventions might only reduce sharing of content that appears especially harmful, whereas sharing of 'less serious' IDP violations may continue in line with broader social media norms. On the other hand, it is possible that many people already avoid sharing what appears to be especially harmful content, such that those influenced by the interventions reduce sharing of 'less serious' IDP violations. Thus, we hypothesized that the extent to which participants reduced sharing following the interventions would vary based on the content valence, or the extent to which social media posts portrayed people in a positive or negative light (H3). However, the hypothesis was not directional.

Lastly, we predicted that participant ratings regarding the seriousness of IDP violations would modulate sharing decisions following the interventions. This hypothesis follows from previous work showing that people's responses to a single individual privacy preference question were a primary factor in predicting their sharing of other people's information online [15]. In particular, we expected that participants who rated IDP violations as less serious would be less responsive to the interventions, particularly in the narrative-based intervention, which we predicted would have the largest effect in altering sharing decisions (H4).

3 METHOD

3.1 Participants

An Institutional Review Board in the Southeast of the United States approved this study. Participants were recruited via Amazon's Mechanical Turk's online participant panel and compensated $5 upon completion. To be eligible, participants had to be living in the United States, fluent in English, ages 18 to 60 years old, have normal or corrected-to-normal vision (i.e., for the video interventions and photo-based images), have normal or corrected-to-normal hearing (i.e., for the video interventions), and pass attention checks during the experiment. The study used three reading-based attention checks and one open-ended question to evaluate the quality of responses. For reading-based attention checks, participants were required to provide matching age submission in two instances of the survey. Additionally, two multiple-choice questions required specific responses to the attention check embedded within a study

Table 1: Summary of participant demographics by experiment group

	Control	Concept	Facts	Narrative	Total	
Total N (%)	95	101	98	101	395	
Age (M (SD))	35.8 (10.5)	37.6 (10.9)	35.6 (10.5)	36.1 (11.2)	36.3 (10.8)	
Gender					n	%
Female	40	56	48	48	192	48.6%
Male	53	45	50	53	201	50.8%
Non-binary/third gender	1				1	0.3%
Prefer not to say	1				1	0.3%
Race						
American Indian or Alaska Native	3	3	3	1	10	2.6%
Asian	5	4	4	5	18	4.5%
Black or African American, Non-Hispanic	7	4	5	9	25	6.3%
Hispanic, Latinx, or Spanish American	0	3	5	3	11	2.8%
White, Non-Hispanic	80	87	80	83	330	83.5%
Other			1		1	0.3%
Highest Degree Earned						
High school diploma or equivalency (GED)	6	10	11	8	35	9%
Associate degree (Junior College)	9	10	6	6	31	8%
Bachelor's degree	62	63	59	58	242	61%
Master's degree	18	15	19	27	79	20%
Doctorate/Professional (MD, JD, DDS, etc.)			1	2	3	0.75%
None of the above (less than high school)		2	1		3	0.75%
Other		1	1		2	0.5%

task (see Appendix A.3). Reading-based attention checks were supplemented with an open-ended question as recommended by prior research [96]. Responses to the open-ended question, "Please indicate the main reasons that you share or re-share photos of other people online" were reviewed for each respondent to ensure the response was applicable the question, not a duplication of the question or other responses (i.e., behaviors of a bot). Participants failing either of the reading-based or open-ended attention checks were removed from the study. The final sample included 395 participants, with participant demographics detailed by experiment group in Table 1.

3.2 Experimental Manipulation

Participants were randomly assigned to one of four groups. Those in the control group did not receive any information related to IDP, whereas participants in the three experimental groups viewed videos on the topic of IDP (see Appendix A for video transcripts). The video interventions were developed in Vyond [97], combining audio, animations, and text to explain issues related to IDP. Participants in the experimental groups had the option of re-watching videos before moving on to the rest of the study. Additionally, participants in the manipulation groups were required to respond to two knowledge-check questions as an attention check and to reinforce understanding of interpersonal privacy violations. The narrative group participants responded to two additional questions, given they viewed two key narrative scenarios. The design of the experiment required video content, and therefore video length, to vary between each group. The concept video was 2 minutes and 16 seconds in length, facts was 3 minutes and 14 seconds, and the narrative video totaled 4 minutes and 16 seconds.

3.2.1 Concept Intervention. The concept intervention video served as a foundation to all video interventions, as it explained the notion of IDP violations through sharing photos and other multimedia online. The concept intervention was designed to achieve consciousness raising, which is used to target changes in behaviors by providing information about the causes, effects (e.g., consequences), and alternative behaviors for a situation [98, 99]. The video included an explanation of IDP, actions which cause violations, which was rebranded as 'privacy pirating' to avoid technical jargon, general effects and consequences for victims, and recommendations on sharing decision making. See Appendix A.1.1 for the full script. For all interventions, videos were chosen as the delivery mechanism, because the use of pictures and entertainment-education are shown to be especially engaging [100], and we suspected participants may not read a series of long text passages.

3.2.2 Facts Intervention. The facts intervention video built on the concept components and included fabricated facts on the prevalence of IDP violations and consequences for victims. Prior research into fact-based interventions have mixed results on relative benefits and effectiveness in changing opinions or behaviors [101, 102] and is complicated by individual preferences and starting opinions on the subject [102]. Research into IDP-relevant benchmarks and consequences, specifically in regard to photos and other multimedia, was limited. Consequently, facts were fabricated because there was an insufficient number of existing benchmarks in this area, as well as to strengthen the experimental manipulation. See Appendix A.1.2

Image 1. Concept condition introduction

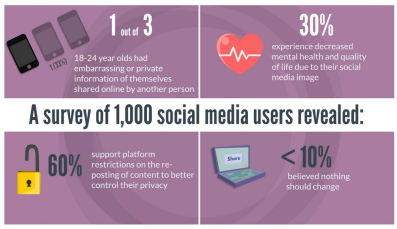

Image 2. Facts condition sample manipulated facts

Image 3. Narrative condition story example

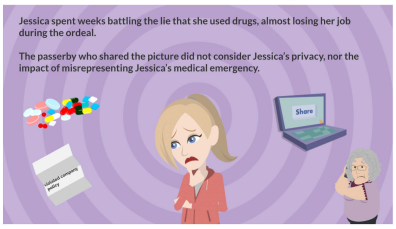

Image 4. Narrative condition story example

Figure 1: Screenshots from intervention videos

for the full script. Since the facts were fabricated in an effort to have a stronger manipulation, an IRB-approved debriefing statement was provided to participants to explain this point following study completion.

3.2.3 Narrative Intervention. The narrative intervention video built on the concept components as well, adding two stories that highlighted consequences of privacy violations for the victims. The narrative videos were designed using the Narrative Immersion Model and focused on experience narratives that were temporarily ordered, explained the relevant outcomes, and reinforced optimal decision-making and sharing behaviors [103]. Both narratives were designed to be somewhat tame, allowing for relatability to participants as 'it could happen to anyone.' See Appendix A.1.3 for the full script. Additionally, screenshots from each condition are shown in Figure 1.

3.3 Questionnaires

Participants completed a series of online questionnaires pertaining to their social media usage, intervention feedback, and attitudes about IDP interactions These questionnaires are described in detail below.

3.3.1 Social Media Usage Questionnaire. The Social Media Usage questionnaire included seven survey items developed to assess the participants' social media photo sharing activity. Participants rated on a scale from 1 (*never*) to 8 (*multiple times a day*) how frequently they shared or re-shared photos on social media based on source (i.e., taken by themselves, their friends, family or discovered on the internet), content (photos were of themselves or others), and intended audience for the shared photos (i.e., friends/connections, general viewers/public, or both). Participants also provided information regarding the social media platforms where they have accounts and which they use for sharing photos.

3.3.2 Intervention Response Questionnaire. The Intervention Response questionnaire included 13 survey items to measure participant ratings of intervention quality, knowledge about interpersonal privacy prevalence, and perceptions regarding how different parties are responsible for managing privacy (i.e., parents/guardians, peers or friends, schools, social media platforms or governments). Questionnaire items related to the quality of intervention were based on the Client Satisfaction Questionnaire adapted to Internet-based interventions (CSQ-I), which is often used to evaluate web-based health interventions [104]. Additionally, participants in the experimental conditions were given an opportunity to provide open-ended written feedback on ways the educational videos could be improved. See Appendix A.2 for full questionnaire.

3.4 Meme decision-making task

Participants also completed a meme decision-making task in which they viewed 68 photo-based memes and rated the likelihood that they would want to share those memes on their own social media profiles (1 = *extremely unlikely*; 5 = *extremely likely*). The photo-based memes were collected from social media and each included at least one person with a clear photo of their face and a brief text caption. A clear photo of the face meant that the meme included a stranger's identifiable information, and the text caption provided a common point of reference to ensure participants had a similar understanding of the meme and might potentially find it interesting enough to share. The context varied between the memes (e.g., personal information such as passport photos, drug use, medical information, or sexual history of the photo subject), but each meme contained potentially sensitive information about strangers who were possibly unaware information was being spread on the internet by strangers. These context categories were generated by evaluating a large data set of photos via a qualitative approach [105] similar to Amon et al. [15]. To narrow the focus to potential IDP violations occurring when sharing of information about strangers, we excluded photos containing celebrities as the privacy of public

figures may be perceived differently. Additionally, we excluded photos that involved polarizing topics such as sexism, racism, or bigoted themes.

3.5 Procedure

After consenting to participate, participants completed eligibility questions to confirm they met inclusion and exclusion criteria for the study. Participants who proceeded also received a content warning, noting that they may view offensive material should they choose to proceed with the study. This was necessary due to the nature of some of the stimuli in the meme decision-making task. Participants were randomly assigned to a condition. Those in the control condition completed the questionnaires, except for the subscale assessing opinions about the video interventions, which did not apply to this group. In contrast, those in the experimental conditions viewed their respective IDP video intervention before proceeding with the rest of the experiment. For experimental groups, the order of the meme decision-making task and the Intervention Response questionnaire were counterbalanced. This was due to concerns that 1) responses to the video interventions could be modulated by viewing memes depicting potential IDP violations, and 2) responses to the meme decision-making task could be modulated after the Intervention Response Questionnaire, which could increase reflection about IDP violations. Notably, order was further controlled through its inclusion as a covariate in a number of models presented below. Lastly, participants completed the Social Media Usage and demographic questionnaires. The average time to complete the study was 36 minutes. As noted earlier, participants in the facts condition received an IRB-approved debriefing statement following the study to explain some of the points in the video were fabricated for the purposes of the study.

3.6 Valence ratings

A separate and independent study was conducted to obtain the perceived valence (i.e., how positively or negatively the meme target was portrayed) for the 68 photo-based memes in the decision-making task (§3.4). Participants were recruited via Amazon's Mechanical Turk recruitment system. To be eligible for the study, participants had to be living in the United States, proficient in English, between the ages of 18 and 60, and regular social media users with an active social media account (i.e., logged in at least once per week). The 104 participants were presented with each meme in random order and rated each meme on the following dimension, "To what extent does this post portray the person in the photo negatively or positively?". Participants provided responses on a Likert scale (1= *very negatively*, 5 = *very positively*). Those ratings were not shared with the current study's participants. However, the average valence ratings for each image are used during the study analysis to understand the relationship of perceived positive or negative portrayals on sharing decisions.

3.7 Qualitative analysis of intervention feedback

Participant's receiving a video-based intervention were asked an option open-ended question, "How could this material be improved to make it a more effective learning experience?". Thematic qualitative analysis was performed using NVIVO software. Two coders independently reviewed each comment and constructed initial themes using open coding. The coders performed a comparative analysis of each response individually to develop and assign themes to note similarities and differences. Once each coder developed their themes and subthemes, they compared their frameworks in detail to converge on a final set of themes and subthemes, ultimately arriving at seven key themes. The analysis included two rounds of each coder separately re-coding each response to align to the seven themes and then discussing the codebook development. In the third round, each response was reviewed collaboratively for consensus on seven core themes divided into eight subthemes. As the thematic coding process involved multiple rounds of reviews, recoding, and revising themes and subthemes to discover the final emergent codebook, we did not measure the inter-rater reliability score [106].

4 RESULTS

4.1 Preliminaries

Participants completed a meme decision-making task where they viewed a series of photo-based memes from social media and rated the likelihood that they would share each online. The average sharing likelihood rating across memes was 2.69 (SD = 1.00) out of 5, with higher ratings indicating a greater likelihood of sharing. The second study used to rate memes on valence, or how positively or negatively they portrayed the photo target, demonstrated an average valence rating of 3.06 (SD = .38) out of 5, with higher ratings indicating a more positive portrayal.

On average, participants rated IDP violations as relatively serious (M = 3.83; SD =0.97) (see Figure 2 for distributions) and felt that social media platforms (74%) were primarily responsible for managing inappropriate private content that leads to IDP violations, followed by the original person who posted (59%), other social media users (37%) and governments (13%; note that participants could select multiple options for the latter). Participants were asked to indicate one or multiple parties they perceived as responsible for educating social media users about the dangers of IDP violations. Most participants indicated social media platforms (81%) and parents/guardians (60%) have a role in educating users, with fewer identifying peers or friends (46%), schools (43%) and governments (e.g., social programs) (32%) (see Figure 3 for distributions). Participants were also asked how often they share or re-share photos on social media and most indicated that they shared online photos multiple times a week (22%), followed by multiple times a month (16%), once a week (12%), less than once in a month (12%), multiple times per day (12%), once in a day (12%), once in a month (10%) and never (5%). Additionally, participants most frequently shared photos with friends or connections (58%), followed by friends and general public (28%), general public (10%) and no sharing or re-sharing of photos (4%). The primary social media platforms used by participants are shown in Table 2.

4.2 User IDP-relevant attitudes and perceptions of interventions

To address RQ1, we first investigated individual attitudes and perceptions of IDP based on the four conditions using a one-way

Table 2: Summary of social media platforms used by participants

Name of platform	Participants who have an account on platform	Participants who share photos on platform
Discord	15%	5%
Facebook	91%	71%
Flicker	3%	2%
Instagram	89%	67%
Myspace	5%	2%
Pinterest	33%	9%
Reddit	39%	6%
Snapchat	31%	12%
TikTok	33%	12%
Twitch	16%	2%
Twitter	69%	32%
WhatsApp	50%	34%
YouTube	78%	22%
Other (e.g., Gab, MeWe, Parler, Quora)	1%	1%
NA: I do not share photos on social media.		3%
Total number of platforms	M(SD) = 5.57(2.28)	M(SD) = 2.79(1.92)

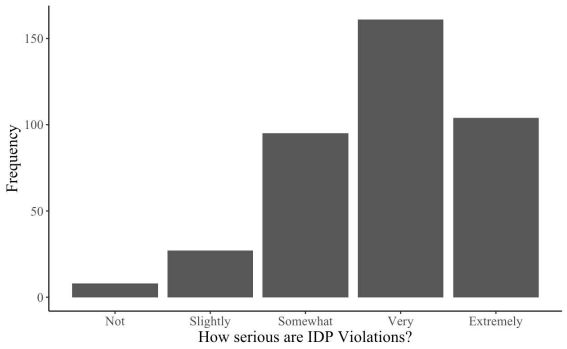

Figure 2: Perceived seriousness of IDP violations

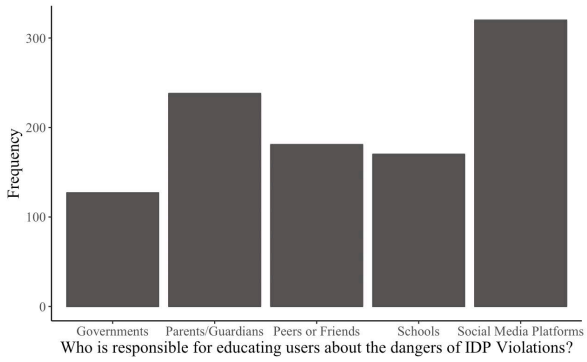

Figure 3: Responsibility for education of IDP violation dangers

multivariate analysis of variance (MANOVA) with condition as the independent variable and dependent variables including ratings of IDP commonality, IDP seriousness, and the need for IDP precautions as indicated on the Intervention Response Questionnaire. We performed an a priori power analysis with an ANCOVA F-Test (four groups with $\alpha = 0.5$) and estimated 400 participants were needed to reach a power of 90% to detect an effect size f equal to 0.25. Considering the statistically significant one-way MANOVA for the main effect of condition ($F(3,391) = 2.56$, $p = .006$, $\eta^2 = .02$) and plotting of residuals identified no concerns with distribution, univariate one-way ANOVAs were used to examine each dependent variable separately to identify which contributed to the significant overall effect. Ratings of IDP commonality varied significantly based on video intervention type, $F(3,390) = 6.60$, $p < .001$, $\eta^2 = .05$. Pairwise comparisons revealed that the narrative intervention ($M = 3.76$; $SD = 0.83$) resulted in significantly higher ratings of IDP commonality than the facts ($M = 3.33$; $SD = 1.07$ or concept ($M = 3.47$; $SD = 0.94$) interventions, $p < .05$. All other effects were statistically non-significant, $p > .05$.

Next, we investigated individual attitudes and perceptions specific to the three video interventions, excluding the control condition from analysis due to control participants not receiving a video intervention. The dependent variables included self-reported amount of new knowledge learned, video effectiveness, need for increased IDP awareness, and IDP topic importance, and the independent variable was video intervention type with order as a covariate (i.e., order of Intervention Response Questionnaire or meme decision-making task presentation). The MANOVA was statistically non-significant for intervention type, indicating that video interventions were perceived as relatively equivalent in effectiveness and did not result in different ratings of IDP seriousness, $p > .05$. However, there was a small but statistically significant effect of order for the MANOVA ($F(1,296) = 2.44$, $p = .04$, $\eta^2 = .03$). A one-way ANOVA follow-up test demonstrated that participants who completed the meme decision-making task before the Intervention Response Questionnaire ($M = 4.31$; $SD = .68$), versus after

(M = 4.06; SD = .80), indicated a greater need for IDP awareness, $F(1,296) = 8.08$, $p = .005$, $\eta^2 = .03$. Thus, viewing meme examples of potential IDP violations led participants to view IDP awareness as more important.

Given the three video intervention types were rated relatively similar in effectiveness by participants, additional analyses were used to test if there was a net positive effect of the interventions. A one-sample t-test examined whether change in knowledge about IDP differed significantly from a null hypothesis of no change. Participants reported a significant increase in IDP knowledge (M = 0.90; SD = 1.20) based on the video interventions, $t(299) = 12.95$, p < .001, 95% CI [.76, 1.04]. An additional one-sample t-test compared participants' perceptions of video effectiveness to a null hypothesis of neutral on the 5-point rating scale (i.e., a value of 2.5). Taken together, participants were relatively optimistic about the potential effectiveness of the interventions (M = 4.14; SD = 0.75), with an effectiveness rating significantly higher than neutral, $t(299) = 38.11$, p < .001, 95% CI [4.06, 4.22]. A one-sample t-test was also used to examine whether those exposed to the intervention agreed that more IDP awareness was needed, with participants scoring significantly higher in agreement than predicted by the null hypothesis of neutral (i.e., a value of 2.5 on the 5-point rating scale), $t(299)$ = 38.79, p < .001, 95% CI [4.10, 4.27]. Lastly, comparing ratings of topic importance (M = 4.39; SD = .64) to a null hypothesis of neutral (2.5) on the 5-point scale, confirmed that participants found the topic of importance, $t(299) = 51.05$, p < .001, 95% CI [4.32, 4.47]. These effects remained statistically significant with the conservative Bonferroni correction for multiple tests (adjusted α = .01).

4.3 Intervention effectiveness in changing sharing decisions

To assess RQ2-4, we analyzed influence of intervention type, content differences in terms of meme valence, and individual differences in terms of IDP-attitudes on sharing decisions. We investigated the degree to which participant's short-term decisions to share photos of other people in social media differed based on condition using an incremental set of mixed-effects regression models. For our first model, we hypothesized that the emotional characteristics of narratives depicting consequences of IDP violations would lead to reduced sharing (H1), and concept- and fact-based interventions would reduce sharing because of increases awareness of the IDP violations (H2). To test these hypotheses, we regressed sharing ratings from the meme decision-making task condition on condition, with participant as a random intercept.[1] Using this approach, the main effect of condition was non-significant, p > .05, therefore not conclusively supporting H1 and H2. See Table 3 Model 1 for full results.

The aim of the interventions was not necessarily to decrease all types of sharing equally. Instead, it was considered potentially more feasible and, in some cases, desirable to decrease sharing of photos which have the most potential to cause harm. Along these lines, the meme decision-making task included photo-based memes that varied in their portrayal of the target. Whereas some memes portrayed more extreme examples of potential IDP violations (e.g., pictures of passports or crude remarks about the target's sexual history), other memes portrayed targets positively (e.g., showing a fun family moment). A range of IDP violation severity was included for two reasons: First, an intervention that decreases sharing of the most harmful information about others would be considered successful. Second, it is possible that some people may be against sharing any type of information about other people without permission, including photos that portray other people positively [107]. For example, France has enacted legal ramifications for posting other people's photos without permission [108], and this policy does not distinguish between the sharing of positive or negative information about others. Thus, we examined intervention effectiveness in light of the valence ratings described earlier.

We examined the extent to which the interaction between condition and valence ratings predicted participants' decisions to share information about others in social media using a mixed-effects multiple linear regression with participant as a random intercept (see Table 3 Model 2 for full results). For this interaction, we hypothesized (H3) that the extent to which participants reduced sharing following the interventions would vary based on the content valence, without specific assumptions on directionality. There was a significant main effect of meme valence, β = .26, p < .001, such that more positively valenced memes were more likely to be shared. Controlling for valence resulted in a marginally significant main effect of condition, where the concept (β = -.25, p = .10) and facts (β = -.30, p = .06) conditions resulted in a decrease of sharing during the meme decision-making task.[2] This main effect was superseded by the significant interaction between condition and valence ratings. As demonstrated in Figure 4, the likelihood of sharing was lower when photo targets were portrayed more negatively. This effect was enhanced in the facts condition compared to the control, such that more negatively-valenced memes were shared less following the facts intervention, β = .06, p = .001. The opposite was true in the narrative condition: Compared to control, participants in the narrative condition were slightly more likely to share negative memes and less likely to share positive memes, β = -.04, p = .04. The narrative condition was also more likely to share negative memes compared to facts (β = -.10, p < .001) and concept (β = -.06, p < .01) participants. Overall, participants were less likely to share memes that portrayed people negatively, and this effect was amplified in the concept and facts conditions. No other comparisons were statistically significant, p > .05.

We hypothesized that the interventions would not be equally effective across all participants; specifically that participant ratings regarding the seriousness of IDP violations would modulate sharing decisions following the interventions (H4). Although a number of variables could be used to examine individual differences in intervention responsiveness, it was noted in earlier findings that participant ratings regarding the degree to which IDP violations constituted a serious problem were stable regardless of condition. Thus, we considered this the most straightforward single measurement of individuals' social media attitudes and willingness to respond to IDP interventions, allowing for model parsimony similar

[1] Note that order was originally included in the models. However, order was removed for the sake of parsimony due to its non-significance as a covariate across the mixed-effects models. All effects described in the section were similarly significant or non-significant regardless of the inclusion of the order covariate.

[2] Our confidence in rejecting the null hypothesis lies along a continuum, consistent with the notion that p-values and effect sizes should be interpreted as a continuous variable [109]. Thus, we acknowledge marginal effects in our findings.

Table 3: Estimates and standard errors for models predicting participants' meme sharing likelihood based on condition (M1), condition by meme valence (M2), and condition by participants' ratings of IDP seriousness (M3)

	Dependent Variable: Sharing Likelihood		
	Model 1(M1)	Model 2 (M2)	Model 3 (M3)
Predictors	β (SE)	β (SE)	β (SE)
Concept intervention	-0.20 (0.14)	-0.25 (0.15)	0.05 (0.60)
Facts intervention	-0.12 (0.14)	-0.30 (0.15)	0.02 (0.55)
Narrative intervention	-0.01 (0.14)	0.11 (0.15)	1.72** (0.57)
Meme valence		0.26*** (0.01)	
Concept interventions * Meme valence		0.02 (0.02)	
Facts intervention * Meme valence		0.06*** (0.02)	
Narrative intervention * Meme valence		-0.04* (0.02)	
IDP seriousness belief			0.02 (0.10)
Concept intervention * IDP seriousness belief			-0.07 (0.15)
Fact intervention * IDP seriousness belief			-0.04 (0.14)
Narrative intervention * IDP seriousness belief			-0.44**(0.15)
Intercept	2.77*** (0.10)	1.97*** (0.11)	2.71*** (0.39)
Observations	26,860	26,860	26,860
Log Likelihood	- 36,475	-35,569	-36,466
Akaike Inf. Crit.	72,961	71,158	72,952
Bayesian Inf. Crit.	73,011	71,240	73,034
Note:		*p<0.1; **p<0.05; ***p<0.01	

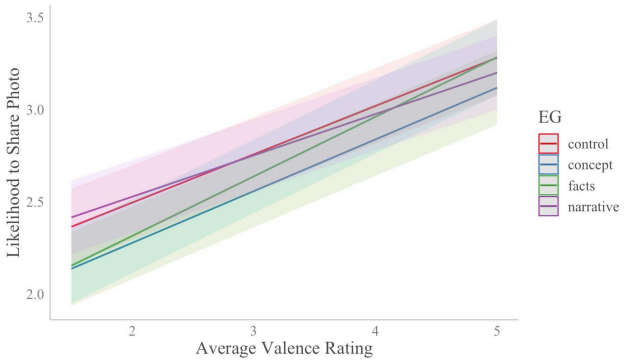

Figure 4: Interaction effect of average valence rating on likelihood to share a photo.

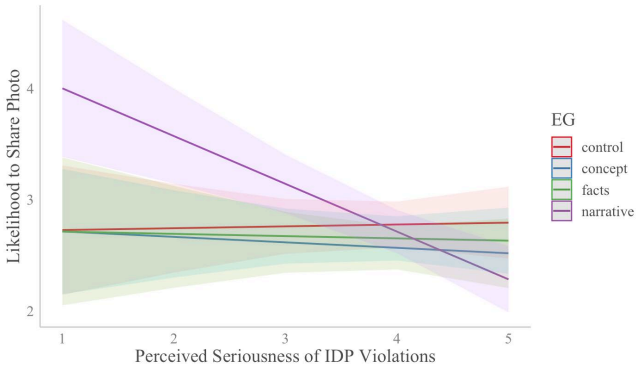

Figure 5: Interaction effect of perceived seriousness of IDVPs on likelihood to share a photo

to that presented in Model 2. A mixed-effects multiple regression model was used to examine the interaction between condition and ratings of IDP violation seriousness in predicting meme sharing decisions. Accounting for individual differences in beliefs about IDP violation seriousness, the findings revealed a significant main effect of condition. Specifically, there was an overall increase of meme sharing likelihood in the narrative condition compared to control, $\beta = 1.72$, $p = .003$, concept ($\beta = 1.66$, $p = .007$) and facts ($\beta = 1.69$, $p = .003$) conditions. Moreover, the statistically significant two-way interaction depicted in Figure 5 between condition and self-reported IDP seriousness indicated that the narrative condition backfired to increase the desire to share among those who do not believe IDP violations are a serious problem, $\beta = -.44$, $p = .003$. In contrast, those who do believe IDP violations are serious responded to the narrative intervention by decreasing their sharing likelihood. All other comparisons were statistically non-significant, $p > .05$.

Lastly, we used an ANOVA to test the relative fit of the mixed-effects models described above, along with Akaike information criterion (AIC) comparisons to provide converging evidence for the model comparisons. Using this approach, we confirmed that the model accounting for condition and valence (AIC = 71158) better accounted for the data than the model with condition alone (AIC = 72961 X^2 (4) = 1811.1, $p < .001$). This was also true when comparing the model with participants' IDP seriousness ratings in addition to condition (AIC = 72952): IDP seriousness ratings contributed positively to model fit, X^2 (4) = 17.40, $p = .002$. We conclude that individual and content differences are important considerations

when examining IDP intervention effectiveness; thus the findings support H3 and H4.

4.4 Qualitative analysis of user open-ended feedback

4.4.1 Themes: How could this material be improved to make it a more effective learning experience? Participants were asked to provide open-ended feedback on their intervention experience and suggested improvements for learning effectiveness, which we examined quantitatively through thematic coding. Overarching concepts included strategies to improve intervention effectiveness (e.g., more relatable stories and examples of desirable user behavior), training experience (e.g., audio and visual quality), and satisfaction with the material (see table in Appendix A.4 for summary of themes). In terms of satisfaction with the video interventions, all three interventions were well received by participants, with 83 comments either stating no improvements were needed or general positive statements about the videos (e.g., "No improvements needed. The video gets the point across well."), compared to six comments related to negative opinions about the intervention quality or effectiveness (e.g., "different concern from many people's opinion was too much"). In addition, some participants (n = 5) doubted the video intervention would be effective in changing some user's sharing behaviors (e.g., "I worry that nothing will change the habits of the new generation of Snap users.").

The remaining comments were organized into themes of video quality and overall training experience ($n = 41$), request for more relatable content and stories ($n = 65$), suggestions to expand training scope to address personal strategies for IDP violation prevention ($n = 19$), and strategies to achieve increased effectiveness through dissemination of the material ($n = 26$). For example, in the dissemination of the material subtheme, participants suggested having an awareness campaign to educate social media users regarding interpersonal privacy violations (e.g., "There need to be awareness campaigns on social media and television and in the schools"), as well as provided suggestions for the target audience (e.g., "Awareness for those who frequently use social media without much knowledge about privacy pirating"). Coders were unable to attribute a theme for 52 responses due to ambiguity in participant's comments or non-answers (e.g., "prefer not to say", "Not sure"). The findings from participant's feedback will be used to improve future versions of each video intervention and associated evaluation materials.

4.4.2 Comparison of condition differences. The most common theme identified was increasing effectiveness through more relatable content and narratives, specifically "real word examples" of IDP violations users might face on social media platforms. This theme appeared most in the concept condition ($n = 29$), but also appeared in the facts ($n = 22$) and narratives ($n = 14$) conditions. The experiment intentionally deprived the concept and facts conditions of strong narrative examples. These participants requested personal testimonies related to IDP violations, such as a concept group participant noting the importance of "having some testimonials by people who were actually affected" and a facts group participant indicating "... adding articles about real incidents and people's interview" would support video intervention effectiveness. Notably, the narrative participants shared similar feedback, albeit less frequently.

The narrative conditions videos begin with the statement that "This experience is based on a true story with names changed" and focuses on the action of the photo sharer and consequences for the victim, which include impacts to social and economic well-being (Appendix A.1.3). However, participants in the narrative condition still recommended "real world examples." Specifically, narrative participants mentioned the desire to see the real victim (e.g., "show videos of live people that have gone through situations that have ruined their lives due to privacy pirating" or " Personal testimony from someone affected"). We conclude that first-person perspectives may be more effective in immersing participants in the narrative, with the potential to alter intervention effectiveness, though this requires further testing.

5 DISCUSSION

As noted by Bak-Coleman et al. [109], "There is no reason to believe that human social dynamics will be sustainable or conducive to wellbeing if left unmanaged." The rise of social media as a major channel of communication has drastically increased the threat of IDP violations. Despite the pressing need for interventions to support privacy attitudes and behaviors, empirical research geared toward identifying evidence-based bottom-up IDP privacy preservation strategies is limited [83]. Moreover, previous attempts at testing psychosocial interventions for IDP preservation have backfired to increase users' sharing [15]. The objective of the present study was to identify psychosocial interventions that promote users' IDP awareness (RQ1) and decrease sharing of other people's private information through photos and other multimedia (RQ2). We also examined how intervention effectiveness was modulated by content (RQ3) and individual (RQ4) differences to develop a more holistic understanding of intervention efficacy. In doing so, we identified intervention features that decrease sharing of potential IDP violations, as well as factors that lead to intervention backfiring and increase sharing of potentially sensitive information.

5.1 Intervention effects on social media sharing likelihood

An important question in designing bottom-up psychosocial strategies concerns what intervention types reduce sharing of potential IDP violations. Prior literature in the computer science and public health domains suggests two primary intervention types: Those that include facts and those that include narratives [20, 22, 23, 86, 87]. A third 'concept' intervention category was also developed, because it was unclear how users would respond to the mere idea of IDP (or 'privacy pirating'). Unexpectedly, the latter concept intervention had the greatest effect in reducing sharing of potential IDP violations, but only for social media content that depicted people relatively negatively. Paired with findings that participants rated IDP violations as relatively high in seriousness, these results suggest that social media users are receptive to general information about IDP and responsive in changing some of their sharing behaviors, at least in the short-term. The finding that users only reduced sharing of negative portrayals highlights that, even when influenced by the intervention, sharing decisions were likely still informed by users'

pre-existing social media beliefs about what is acceptable to share. It is promising that the concept intervention reduced sharing of negative information about others, but users have been developing and strengthening their sharing preferences for extended periods of time through ongoing social media usage. The concept intervention represents a path forward for IDP preservation, but it is to be expected that a one-time intervention will not necessarily transform a person's social media sharing preferences.

Because IDP is inherently a social phenomenon, we hypothesized that the narrative condition would lead to the greatest reduction in IDP-relevant sharing due to its ability to convey complex social information (e.g., other people's perspectives, emotions, and consequences) [110, 111]. We crafted stories that depicted individual's different privacy preferences and how they were violated by other people's sharing decisions, linking the social media sharing to consequences for the victim that sounded like 'it could happen to anyone.' In contrast to our hypothesis, the narrative condition backfired to increase sharing among users who rated IDP violations as less serious. It is possible that, for users already in the mindset that IDP violations are not very serious, the narratives that we provided may have seemed mild and inadvertently reinforced their preexisting IDP attitudes. However, more research is needed to examine exactly which aspects of the narrative intervention caused it to backfire.

The facts intervention was slightly less effective than the concept video in decreasing sharing of negatively-valenced content. This reduction in effectiveness occurred despite the facts about IDP prevalence and consequences being embedded into the concept video to provide context to the presented facts. It may seem surprising that the facts diluted the effectiveness of the concept information. However, as with the narratives, we suspect that the facts we developed for the intervention were not perceived as especially compelling. Taken together, that the facts and narratives lowered effectiveness of the concept information highlights the importance of ensuring all portions of an intervention are impactful and have added value, so as not to detract from the main messaging.

Whereas previous literature highlights facts (i.e., statistics) and narratives (i.e., emotional stories) as especially impactful interventions [112, 113], our findings suggest that there are cases in which mere awareness of concepts is an important first step toward social change. Previous research has examined facts as a means of reducing phishing attacks [87] or stories to reduce personal disclosures [86], but the basic premises of phishing and personal privacy are likely more well-known than IDP. Given the lack awareness of the term IDP violation, we coined a user-friendly term like 'privacy pirating' to address this point. More research is needed to examine specific features of facts and narratives that enhance the concept intervention, as well as long-term impact of the interventions.

5.2 Interdependent privacy attitudes

In addition to examining sharing decisions following intervention, we investigated the extent to which the interventions altered IDP-relevant attitudes. Taken together, the interventions were associated with a self-reported increase in IDP knowledge, interventions were perceived as effective, the IDP topic was rated as important, and participants agreed that the intervention was needed. Overall, survey responses support that IDP interventions are viewed as informative and will be well-received by the general public.

Notably, the interventions did not alter users' ratings of IDP seriousness compared to control, nor did interventions alter participant ratings regarding the need for more IDP precautions. One explanation is that these particular privacy preferences are more trait-like in that they represent characteristics of thinking, feeling, and behaving that generalize across situations, whereas 'states' are relevant to a specific time and context [114]. For example, a person's overall rating of IDP seriousness might be relatively stable over time, but their sharing of specific material may vary based on privacy-relevant contextual factors. This notion is consistent with research demonstrating the interplay between individual and situational traits in determining privacy decisions [115]. For this reason, a reduction of IDP-relevant sharing based on intervention type may not be accompanied by a change in overall IDP attitudes.

The narrative condition was associated with a significant increase in participant's beliefs that IDP violations are common and occur frequently. As previously mentioned, we focused on developing narratives with the intention of communicating 'it could happen to anyone.' The focus was on ordinary rather than extraordinary scenarios and consequences. For example, the story about 'Matteo' noted he was working as a cashier when a stranger took a picture of him that subsequently went viral. Thus, while the narrative may have backfired by increasing sharing among those who thought IDP violations were not especially serious, the narratives appear effective in communicating IDP violations as something that happen with relative regularity.

5.3 User's intervention preferences

The qualitative results underline users' strong preference for real-world stories as a part of IDP interventions. Even in the narrative condition, participants noted that they wanted to hear more stories or even stories in the first person (versus third person). The findings are in line with narrative transportation theory, or the idea that narratives are a powerful source of social influence that serve to engage and encourage meaning making [88, 89, 92, 116]. In working to understand a fairly new concept like IDP or 'privacy pirating,' it is understandable that users would want to hear concrete examples. On the one hand, our findings demonstrated that narrative interventions could backfire to increase sharing of potential IDP violations for certain types of users. On the other hand, user responses to the interventions highlight that stories are likely still an important element of bottom-up psychosocial interventions that require more research to optimize effectiveness.

Additionally, participant's requested prescriptive strategies to prevent IDP violations as a sharer and victim. Prescriptive narratives can be effective in making concepts more relatable, even when contrasting with previously held beliefs [117]. Hearing real people share their stories with prescriptive techniques to prevent IDP violations can help users emotionally and intellectually connect with a concept that is inherently social in nature. Overall, the qualitative user feedback provides additional direction for future intervention strategies.

6 LIMITATIONS AND FUTURE DIRECTIONS

6.1 Limitations

This experiment examined changes in user's IDP-relevant attitudes and decisions to share photos of varying valence and content following video-based educational interventions. Given the focus on limiting interdependent privacy violations, versus sharing of personal information, the focus was necessarily limited to IDP violations that involved the sharing or re-sharing of content others, not of oneself. Moreover, the study also limited its scope by using photo-based memes to evaluate users' decisions to share potentially sensitive information about other people. To understand the effect of photo valence on each area, a second set of participants were used to rate meme valence to reduce survey burden and the potential for order effects. The researchers acknowledge the limitations of using valence ratings from a different group of participants with potentially varying perceptions of the memes. However, the use of a sufficient sample size (104 participants and 68 memes) and similar general eligibility requirements (e.g., same age requirements) allows for the ratings to remain relevant for the current study. In measuring changes in knowledge for IDP violations, the study used self-reported assessments of pre-and post-knowledge. Self-assessments have their limitations as participants may be unable to accurately assess their knowledge. However, prior research has found retrospective pre- and post-test self-assessments (i.e., performing the self-assessment after the intervention) can address specific self-assessment limitations such as response-shift bias (i.e., participant's understanding of a concept changes between pre and post-tests) [118], therefore allowing for representative assessments of knowledge shifts. Lastly, conditions had different durations for intervention exposure inherent to each condition's design. The narrative condition had the fewest participant comments about video length, despite having the longest intervention length. It is unclear whether duration of an intervention impacts effectiveness (i.e., participant fatigue). The researchers acknowledge the limitation of not directly evaluating duration as a confounding variable.

6.2 Future Directions

This research has a number of future directions related to the intervention framework (e.g., concept vs narrative-based), content (i.e., images used in sharing decision task), and evaluating additional IDP-relevant attitudes. We aimed to compare effectiveness of three major types of psychosocial interventions. The current study identified concept-based interventions were most effective in reducing the sharing of negatively-valanced photos of other people. This finding suggests a major gap in social media users' knowledge about IDP and suggests that even learning about the basic concept of IDP or 'privacy pirating' has benefits in raising participant's consciousness of the issue and reducing prevalence of the least desirable behaviors. We recommend that IDP awareness interventions be incorporated into more traditional cyber security and privacy trainings. Additional research is also needed to identify the strongest drivers for the behavior change from the concept-based interventions (e.g., awareness of consequences of IDP violations for all parties or providing examples of model sharing behavior).

Similarly, participants expressed a strong preference for relatable narratives to explain the causes, consequences, and prevention strategies related to IDP violations. Thus, despite the fact that narrative interventions backfired to increase sharing among some users, our research suggests that narrative-based interventions may still be a relevant tool to change photo-sharing behaviors [119, 120]. Future intervention studies should test narratives with extremely negative and sensitive scenarios (in contrast to our 'it could happen to anyone' scenarios), from both the first and third person. Prior research suggest that narratives that effectively allow participants to shift perspectives (e.g., seeing a situation from the perspective of another person) can be effective in changing behaviors [99].

Future studies can include follow-up assessments regarding each participant's overall IDP attitudes and sharing likelihood at set increments to determine if influence fades over time, or, alternatively, if repeated exposure to intervention content can sustain changes in attitudes and behaviors. Another research direction is to understand to what extent individual IDP-relevant attitudes modulate intervention effectiveness. Specifically, researchers should consider the relationship between commonly observed sharing behaviors and social capital factors in intervention effectiveness. For example, researchers may use recent real-world photo examples to make the content relatable. If participants were exposed to this content in their real lives, they may perceive this sharing behavior as socially acceptable and replicate the observed sharing behaviors despite the intervention. This could be achieved by asking participants if they have seen similar content before or measuring the importance of social capital for each participant. Additionally, research into the influence of individual attitudes should focus on how to design interventions to target different types of users, with the ability to tailor content and behavior changing strategies based on these differences.

7 CONCLUSIONS

The present research provides empirical evidence on the relative effectiveness of bottom-up psychosocial interventions that include concepts (i.e., general information), facts (i.e., statistics), or narratives (i.e., emotional stories) in changing IDP-relevant attitudes and photo-sharing behaviors compared to a control condition. These experimental groups used educational videos that incorporated emerging findings regarding strategies for influencing attitudes and behaviors across a variety of public health, political, and social issues. Overall, users rated the IDP interventions positively and felt that the interventions were needed. In particular, the interventions explaining the general 'concept' of IDP and 'facts' about IDP decreased sharing of especially negatively-valenced memes depicting other people, compared to the control and narrative conditions. In contrast, the condition sharing emotional 'narratives' backfired to increase sharing likelihood of potential IDP violations among users who did not feel IDP violations were especially serious, whereas narrative-condition users who rated IDP violations as serious reduced memes sharing. Notably, qualitative analysis revealed that participants overwhelmingly desired to hear more IDP narratives, despite this being the least effective condition in terms of altering user's sharing behaviors. The concept- and fact-based IDP interventions hold promise in reducing IDP violations in social media, but careful consideration of narratives and their potential to backfire is required even in the face of positive participant feedback.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the National Science Foundation under grant CNS-2053152. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

[1] Mathias Humbert, Benjamin Trubert, and Kévin Huguenin. 2019. A survey on interdependent privacy. ACM Computing Surveys (CSUR), Vol. 52, No. 6, 1-40.
[2] Kate Knibbs. 2014. 1.8 billion images are uploaded every day. Daily Dot. Accessed August 26, 2022.
[3] Huina Mao, Xin Shuai, and Apu Kapadia. 2011. Loose tweets: An analysis of privacy leaks on twitter. In Proceedings of the 10th annual ACM workshop on Privacy in the Electronic Society, 1-12.
[4] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. 2020. Towards a taxonomy of content sensitivity and sharing preferences for photos. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1-14.
[5] Alessandro Acquisti *et al.* 2017. Nudges for privacy and security: Understanding and assisting users' choices online. ACM Computing Surveys (CSUR), Vol. 50, No. 3, 1-41.
[6] Helen Nissenbaum. 2009. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.
[7] Qingya Wang, Wei Chen, and Yu Liang. 2011. The Effects of Social Media on College Students. Johnson & Wales University. MBA Student Scholarship.
[8] Justin W. Patchin and Sameer Hinduja. 2020. Sextortion among adolescents: Results from a national survey of US youth. Sexual Abuse, Vol. 32, No. 1, 30-54.
[9] Yok-Fong Paat and Christine Markham. 2021. Digital crime, trauma, and abuse: Internet safety and cyber risks for adolescents and emerging adults in the 21st century. Social Work in Mental Health, Vol. 19, No. 1, 18-40.
[10] Athina Ioannou, Iis Tussyadiah, Graham Miller, Shujun Li, and Mario Weick. 2021. Privacy nudges for disclosure of personal information: A systematic literature review and meta-analysis. PloS one, Vol. 16, No. 8. https://doi.org/10.1371/journal.pone.0256822.
[11] Reza Ghaiumy Anaraky, Tahereh Nabizadeh, Bart P Knijnenburg, and Marten Risius. 2018. Reducing default and framing effects in privacy decision-making. SIGHCI 2018 Proc. Assoc. for Info. Sys.(AIS), Atlanta, GA, USA, Vol. 7.
[12] Hiroaki Masaki, Kengo Shibata, Shui Hoshino, Takahiro Ishihama, Nagayuki Saito, and Koji Yatani. 2020. Exploring nudge designs to help adolescent SNS users avoid privacy and safety threats. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, 1-11.
[13] Tehila Minkus, Kelvin Liu, and Keith W Ross. 2015. Children seen but not heard: When parents compromise children's online privacy. In Proceedings of 24th International Conference on World Wide Web, 2015, 776-786.
[14] William Easterly. 2008. Institutions: Top Down or Bottom Up? American Economic Review, Vol. 98, No. 2, 95-99. https://doi.org/10.1257/aer.98.2.95.
[15] Mary Jean Amon, Rakibul Hasan, Kurt Hugenberg, Bennett I. Bertenthal, and Apu Kapadia. 2020. Influencing Photo Sharing Decisions on Social Media: A Case of Paradoxical Findings. In 2020 IEEE Symp. on Secu. and Privacy (SP). IEEE, San Francisco, California, USA, 79–95.
[16] Jon Roozenbeek, Sander van der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. 2022. Psychological inoculation improves resilience against misinformation on social media. Science Advances, Vol. 8, No. 34, eabo6254. https://doi.org/doi:10.1126/sciadv.abo6254.
[17] Ciara Keenan, Allen Thurston, and Karolina Urbanska. 2017. Video-based interventions for promoting positive social behaviour in children with autism spectrum disorders: a systematic review and meta-analysis. The Campbell Collaboration.
[18] Ofir Turel, Michele Mouttapa, and Elaine Donato. 2015. Preventing problematic Internet use through video-based interventions: A theoretical model and empirical test. Behaviour & Information Technology, Vol. 34, No. 4, 349-362.
[19] Murtaza Dahodwala, Rose Geransar, Julie Babion, Jill de Grood, and Peter Sargious. 2018. The impact of the use of video-based educational interventions on patient outcomes in hospital settings: A scoping review. Patient Education and Counseling, Vol. 101, No. 12, 2116-2124.
[20] E. James Baesler and Judee K. Burgoon. 1994. The temporal effects of story and statistical evidence on belief change. Communication Research, Vol. 21, No. 5, 582-602.
[21] Sally Dunlop, Melanie Wakefield, and Yoshihisa Kashima. 2010. Pathways to Persuasion: Cognitive and Experiential Responses to Health-Promoting Mass Media Messages. Communication Research - COMMUN RES, Vol. 37, 02/01, 133-164. https://doi.org/10.1177/0093650209351912.
[22] Kathryn Greene and Laura S. Brinn. 2003. Messages influencing college women's tanning bed use: Statistical versus narrative evidence format and a self-assessment to increase perceived susceptibility. Journal of Health Communication, Vol. 8, No. 5, 443-461.
[23] Jenifer E. Kopfman, Sandi W. Smith, James K. Ah Yun, and Annemarie Hodges. 1998. Affective and cognitive reactions to narrative versus statistical evidence organ donation messages. https://doi.org/10.1080/00909889809365508.
[24] SarahJ. Durkin, Lois Biener, and Melanie A. Wakefield. 2009. Effects of Different Types of Antismoking Ads on Reducing Disparities in Smoking Cessation Among Socioeconomic Subgroups. American Journal of Public Health, Vol. 99, No. 12, 2217-2223. https://doi.org/10.2105/ajph.2009.161638.
[25] S. Emery *et al.* 2005. Televised state-sponsored antitobacco advertising and youth smoking beliefs and behavior in the United States, 1999-2000. Arch Pediatr Adolesc Med, Vol. 159, No. 7, Jul, 639-45. https://doi.org/10.1001/archpedi.159.7.639.
[26] Steven L. West and Keri K. O'Neal. 2004. Project DARE outcome effectiveness revisited. American Journal of Public Health, Vol. 94, No. 6, 1027-1029.
[27] Isadora Krsek, Kimi Wenzel, Sauvik Das, Jason I. Hong, and Laura Dabbish. 2022. To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1-17.
[28] Meenakshi S. Subbaraman and Sarah C. M. Roberts. 2019. Costs associated with policies regarding alcohol use during pregnancy: Results from 1972-2015 Vital Statistics. PloS one, Vol. 14, No. 5, e0215670.
[29] Dora L. Costa and Matthew E. Kahn. 2013. Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. Journal of the European Economic Association, Vol. 11, No. 3, 680-702.
[30] David Halpern. 2015. Inside the Nudge Unit: How Small Changes Can Make a Big Difference. Random House.
[31] Eyal Peer, Serge Egelman, Marian Harbach, Nathan Malkin, Arunesh Mathur, and Alisa Frik. 2020. Nudge me right: Personalizing online security nudges to people's decision-making styles. Computers in Human Behavior, Vol. 109, 106347.
[32] Siyao Fu, Haibo He, and Zeng-Guang Hou. 2014. Learning race from face: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, No. 12, 2483-2509.
[33] Yan Shoshitaishvili, Christopher Kruegel, and Giovanni Vigna. 2015. Portrait of a privacy invasion. Proc. Priv. Enhancing Technol., Vol. 2015, No. 1, 41-60.
[34] Kurt Thomas, Chris Grier, and David M. Nicol. 2010. Unfriendly: Multi-party privacy risks in social networks. In Proceedings of the International Symposium on Privacy Enhancing Technologies Symposium, 2010: Springer, 236-252.
[35] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. 2011. " I regretted the minute I pressed share": A qualitative study of regrets on Facebook. In Proceedings of seventh symposium on Usable Privacy and Security, 1-16.
[36] Mary Madden. 2012. Privacy management on social media sites. Pew Research Center. [Online]. Available: https://www.pewresearch.org/internet/2012/02/24/privacy-management-on-social-media-sites/. Accessed September 9, 2022.
[37] Pamela Wisniewski, A.K.M. Najmul Islam, Bart P. Knijnenburg, and Sameer Patil. 2015. Give Social Network Users the Privacy They Want. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada. [Online]. Available: https://doi.org/10.1145/2675133.2675256.
[38] David H. Nguyen *et al.* 2009. Encountering SenseCam: Personal recording technologies in everyday life. In Proceedings of the 11th international conference on Ubiquitous Computing, 165-174.
[39] Benjamin Henne, Christian Szongott, and Matthew Smith. 2013. SnapMe if you can: Privacy threats of other peoples' geo-tagged media and what we can do about it. In Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks, 95-106.
[40] Chutikulrungsee Tharntip Tawnie and Burmeister Oliver Kisalay. 2017. Interdependent Privacy. The ORBIT Journal, Vol. 1, No. 2, 1-14.
[41] Ahmed Al Marouf, Rasif Ajwad, and Adnan Ferdous Ashrafi. Looking behind the mask: A framework for detecting character assassination via troll comments on social media using psycholinguistic tools. In Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE): IEEE, 1-5.
[42] Andrew Besmer and Heather Richter Lipford. Moving beyond untagging: Photo privacy in a tagged world. 2010. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1563-1572.
[43] Bernhard Debatin, Jennette P. Lovejoy, Ann-Kathrin Horn, and Brittany N. Hughes. 2009. Facebook and online privacy: Attitudes, behaviors, and unintended consequences. Journal of Computer-Mediated Communication, Vol. 15, No. 1, 83-108.
[44] Kate Raynes-Goldie. 2010. Aliases, Creeping, and Wall cleaning: Understanding Privacy in the Age of Facebook. First Monday.
[45] Richard Dawkins and Nicola Davis. 2017. The Selfish Gene. Macat Library.
[46] Susan Blackmore and Susan J. Blackmore. 2000. The Meme Machine. Oxford Paperbacks.

[47] Michael Massimi, Khai Truong, David Dearman, and Gillian Hayes. 2009. Understanding recording technologies in everyday life. IEEE Pervasive Computing, Vol. 9, No. 3, 64-71.

[48] Jin Chen, Jerry Wenjie Ping, Yunjie Xu, and Bernard C. Y. Tan. 2015. Information privacy concern about peer disclosure in online social networks. IEEE Transactions on Engineering Management, Vol. 62, No. 3, 311-324.

[49] Alessandro Acquisti and Ralph Gross. 2006. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In Proceedings of International workshop on privacy enhancing technologies, 28 (June 2006), 36-58.

[50] Amanda Lenhart, Kristen Purcell, Aaron Smith, and Kathryn Zickuhr. 2010. Social Media & Mobile Internet Use Among Teens and Young Adults. Millennials. Pew Internet & American Life Project.

[51] Danah Michele Boyd. 2008. Taken out of context: American teen sociality in networked publics. University of California, Berkeley.

[52] Danah Boyd. 2014. It's Complicated: The Social Lives of Networked Teens. Yale University Press.

[53] Justin Lee Becker. 2009. Measuring privacy risk in online social networks. University of California, Davis.

[54] Pritam Gundecha, Geoffrey Barbier, and Huan Liu. 2011. Exploiting vulnerability to secure user privacy on a social networking site. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge Discovery and Data Mining, 511-519.

[55] Pritam Gundecha, Geoffrey Barbier, Jiliang Tang, and Huan Liu. 2014. User vulnerability and its reduction on a social networking site. ACM Transactions on Knowledge Discovery from Data (TKDD), Vol. 9, No. 2, 1-25.

[56] Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen. 2011. We're in it together: interpersonal management of disclosure in social network services. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 3217-3226.

[57] Alice E. Marwick and Danah Boyd. 2014. Networked privacy: How teenagers negotiate context in social media. New Media & Society, Vol. 16, No. 7, 1051-1067.

[58] Pamela Wisniewski, Heather Lipford, and David Wilson. 2012. Fighting for my space: Coping mechanisms for SNS boundary regulation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 609-618.

[59] Hichang Cho and Anna Filippova. 2016. Networked privacy management in Facebook: A mixed-methods and multinational study. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, 503-514.

[60] Yasmeen Rashidi *et al.* 2018. " You don't want to be the next meme": College Students' Workarounds to Manage Privacy in the Era of Pervasive Photography. In Proceedings of the fourteenth symposium on Usable Privacy and Security (SOUPS 2018), 143-157.

[61] Tobias Kroll and Stefan Stieglitz. 2021. Digital nudging and privacy: improving decisions about self-disclosure in social networks. Behaviour & Information Technology, Vol. 40, No. 1, 1-19.

[62] Yifang Li, Nishant Vishwamitra, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos. In Proceedings of the ACM on Human-Computer Interaction, Vol. 1, No. CSCW, 1-24.

[63] Kaihe Xu, Yuanxiong Guo, Linke Guo, Yuguang Fang, and Xiaolin Li. 2015. My privacy my decision: Control of photo sharing on online social networks. IEEE Transactions on Dependable and Secure Computing, Vol. 14, No. 2, 199-210.

[64] Yu Pu and Jens Grossklags. 2014. An Economic Model and Simulation Results of App Adoption Decisions on Networks with Interdependent Privacy Consequences. In Proceedings of the international conference on Decision and Game Theory for Security: Springer International Publishing, 246-265.

[65] Anna Cinzia Squicciarini, Mohamed Shehab, and Federica Paci. 2009. Collective privacy management in social networks. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain. [Online]. Available: https://doi.org/10.1145/1526709.1526780.

[66] B. Carminati and E. Ferrari. 2011. Collaborative access control in on-line social networks. In Proceedings of the 7th international conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 15-18 (October 2011), 231-240. https://doi.org/10.4108/icst.collaboratecom.2011.247109.

[67] Arunee Ratikan and Mikifumi Shikida. 2014. Privacy Protection Based Privacy Conflict Detection and Solution in Online Social Networks. In Proceedings of the international conference on Human Aspects of Information Security, Privacy, and Trust, 433-445.

[68] Anna C. Squicciarini, Heng Xu, and Xiaolong Zhang. 2011. CoPE: Enabling collaborative privacy management in online social networks. Journal of the American Society for Information Science and Technology, Vol. 62, No. 3, 521-534. https://doi.org/10.1002/asi.21473.

[69] Jose M. Such and Natalia Criado. 2014. Adaptive Conflict Resolution Mechanism for Multi-party Privacy Management in Social Media. In Proceedings of the 13th Workshop on Privacy in the Electronic Society, Scottsdale, Arizona, USA. [Online]. Available: https://doi.org/10.1145/2665943.2665964.

[70] Jay P. Kesan and Andres A. Gallo. 2006. Why are the United States and the European Union failing to regulate the internet efficiently? Going beyond the bottom-up and top-down alternatives. European Journal of Law and Economics, Vol. 21, No. 3, 2006/05/01, 237-266. https://doi.org/10.1007/s10657-006-7422-y.

[71] Janina M. Björk, Pernilla Bolander, and Anna K. Forsman. 2021. Bottom-Up Interventions Effective in Promoting Work Engagement: A Systematic Review and Meta-Analysis. Frontiers in Psychology, 3754.

[72] Silvia Riva and Ezekiel Chinyio. 2018. Stress Factors and Stress Management Interventions: The Heuristic of "Bottom Up" an Update From a Systematic Review. Occupational Health Science, Vol. 2, No. 2, 2018/06/01, 127-155. https://doi.org/10.1007/s41542-018-0015-7.

[73] Sherry Emery *et al.* 2005. Televised State-Sponsored Antitobacco Advertising and Youth Smoking Beliefs and Behavior in the United States, 1999-2000. Archives of Pediatrics & Adolescent Medicine, Vol. 159, No. 7, 639-645. https://doi.org/10.1001/archpedi.159.7.639.

[74] Shana L. Maier. 2022. Rape Victim Advocates' Perceptions of the # MeToo Movement: Opportunities, Challenges, and Sustainability. Journal of Interpersonal Violence. https://doi.org/10.1177/08862605221081929.

[75] Karine Nyborg *et al.* 2016. Social norms as solutions. Science, Vol. 354, No. 6308, 42-43. https://doi.org/doi:10.1126/science.aaf8317.

[76] R.M. Kanter. 1977. Men and Women of the Corporation. Basic Books. New York.

[77] Kavous Salehzadeh Niksirat, Evanne Anthoine-Milhomme, Samuel Randin, Kévin Huguenin, and Mauro Cherubini. 2021. "I thought you were okay": Participatory Design with Young Adults to Fight Multiparty Privacy Conflicts in Online Social Networks. In Proceedings of the Designing Interactive Systems Conference 2021, 104-124.

[78] Jennifer Golbeck and Matthew Louis Mauriello. 2016. User Perception of Facebook App Data Access: A Comparison of Methods and Privacy Concerns. Future Internet, Vol. 8, No. 2, 9. [Online]. Available: https://www.mdpi.com/1999-5903/8/2/9.

[79] Anatoliy Gruzd, Joanne McNeish, Lilach Dahoah Halevi, and Martin Phillips. 2021. Seeing Self in Data: The Effect of a Privacy Literacy Intervention on Facebook Users' Behaviour. Available at SSRN 3946376.

[80] J Jennifer M. Walton, Jonathan White, and Shelley Ross. 2015. What's on YOUR Facebook profile? Evaluation of an educational intervention to promote appropriate use of privacy settings by medical students on social networking sites. Medical Education Online, Vol. 20, No. 1, 2015/01/01, 28708. https://doi.org/10.3402/meo.v20.28708.

[81] Priya Kumar *et al.* 2018. Co-designing online privacy-related games and stories with children. In Proceedings of the 17th ACM Conference on Interaction Design and Children, Trondheim, Norway. [Online]. Available: https://doi.org/10.1145/3202185.3202735.

[82] Mauro Cherubini, Kavous Salehzadeh Niksirat, Marc-Olivier Boldi, Henri Keopraseuth, Jose M Such, and Kévin Huguenin. 2021. When forcing collaboration is the most sensible choice: Desirability of precautionary and dissuasive mechanisms to manage multiparty privacy conflicts. In Proceedings of the ACM on Human-Computer Interaction, Vol. 5, No. CSCW1, 1-36.

[83] Anthony T. Pinter, Pamela J. Wisniewski, Heng Xu, Mary Beth Rosson, and Jack M. Caroll. 2017. Adolescent Online Safety: Moving Beyond Formative Evaluations to Designing Solutions for the Future. In Proceedings of the 2017 Conference on Interaction Design and Children, Stanford, California, USA. [Online]. Available: https://doi.org/10.1145/3078072.3079722.

[84] Haiyan Jia and Heng Xu. 2016. Autonomous and Interdependent: Collaborative Privacy Management on Social Networking Sites. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, California, USA. [Online]. Available: https://doi.org/10.1145/2858036.2858415.

[85] Bernadette Kamleitner and Vince Mitchell. 2019. Your Data Is My Data: A Framework for Addressing Interdependent Privacy Infringements. Journal of Public Policy & Marketing, Vol. 38, No. 4, 433-450. https://doi.org/10.1177/0743915619858924.

[86] Amanda Nosko *et al.* 2012. Examining priming and gender as a means to reduce risk in a social networking context: Can stories change disclosure and privacy setting use when personal profiles are constructed? Computers in Human Behavior, Vol. 28, No. 6, 2012/11/01/, 2067-2074. https://doi.org/10.1016/j.chb.2012.06.010.

[87] Rick Wash and Molly M. Cooper. 2018. Who provides phishing training? Facts, stories, and people like me. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada. [Online]. Available: https://doi.org/10.1145/3173574.3174066.

[88] Jerome Bruner. 1996. The Culture of Education. Harvard University Press.

[89] Leda Cosmides. 1989. The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. Cognition, Vol. 31, No. 3, 187-276.

[90] Anneke De Graaf, José Sanders, and Hans Hoeken. 2016. Characteristics of narrative interventions and health effects: A review of the content, form, and context of narratives in health-related narrative persuasion research. Review of Communication Research, Vol. 4, 88-131.

[91] Fuyuan Shen, Vivian C. Sheer, and Ruobing Li. 2015. Impact of narratives on persuasion in health communication: A meta-analysis. Journal of Advertising,

Vol. 44, No. 2, 105-113.
[92] Helena Bilandzic and Rick Busselle. 2013. Narrative persuasion. The Sage Handbook of Persuasion: Developments in Theory and Practice, Vol. 2, 200-219.
[93] Sally M. Dunlop, Melanie Wakefield, and Yoshihisa Kashima. 2010. Pathways to persuasion: Cognitive and experiential responses to health-promoting mass media messages. Communication Research, Vol. 37, No. 1, 133-164.
[94] Mike Allen and Raymond W. Preiss. 1997. Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. Communication Research Reports, Vol. 14, No. 2, 125-131.
[95] Amber Marie Reinhart. 2006. Comparing the persuasive effects of narrative versus statistical messages: A meta-analytic review. State University of New York at Buffalo.
[96] Jenny Tang, Eleanor Birrell, and Ada Lerner. 2022. Replication: How well do my results generalize now? The external validity of online privacy and security surveys. In Proceedings of the Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022), 367-385.
[97] Vyond. 2022. Video Animation Software for Businesses. Retrieved September 1, 2022 from https://www.vyond.com.
[98] Parviz Birjandi and Ali Derakhshan. 2014. Pragmatic comprehension of apology, request and refusal: An investigation on the effect of consciousness-raising video-driven prompts. Applied Research on English Language, Vol. 3, No. 1, 67-86.
[99] LK Bartholomew Eldredge, Christine M. Markham, Robert AC Ruiter, Maria E. Fernández, Gerjo Kok, and Guy S. Parcel. 2016. Planning Health Promotion Programs: An Intervention Mapping Approach. John Wiley & Sons.
[100] Maya Adam, Shannon A McMahon, Charles Prober, and Till Bärnighausen. 2019. Human-centered design of video-based health education: an iterative, collaborative, community-based approach. Journal of medical Internet research, Vol. 21, No. 1, e12128.
[101] Oscar Barrera, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya. 2020. Facts, alternative facts, and fact checking in times of post-truth politics. Journal of public economics, Vol. 182, 104123.
[102] Jan G Voelkel, Mashail Malik, Chrystal Redekopp, and Robb Willer. 2022. Changing Americans' attitudes about immigration: Using moral framing to bolster factual arguments. The ANNALS of the American Academy of Political and Social Science, Vol. 700, No. 1, 73-85.
[103] Victoria A. Shaffer, Elizabeth S. Focella, Andrew Hathaway, Laura D. Scherer, and Brian J. Zikmund-Fisher. 2018. On the usefulness of narratives: an interdisciplinary review and theoretical model. Annals of Behavioral Medicine, Vol. 52, No. 5, 429-442.
[104] Leif Boß*et al.* 2016. Reliability and validity of assessing user satisfaction with web-based health interventions. Journal of Medical Internet Research, Vol. 18, No. 8, e5952.
[105] Barney G. Glaser and Anselm L. Strauss. 2017. The Discovery of Grounded Theory: Strategies for Qualitative Research. Routledge.
[106] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. In Proceedings of the ACM on Human-Computer Interaction, Vol. 3, No. CSCW, 1-23.
[107] M. J. Amon, N. Kartvelishvilli, B. Bertenthal, K. Hugenberg, and A. Kapadia. 2022. Sharenting and children's privacy: Parenting style, practices, and perspectives on sharing young children's photos on social media. Computer Supported Cooperative Work.
[108] Jess Staufenberg. 2016. French parents 'could face prison' for posting photos of their children on Facebook. The Independent. Retrieved on September 1, 2022 from https://www.independent.co.uk/news/world/europe/french-parents-told-their-children-might-sue-them-for-pictures-put-on-facebook-a6906671.html
[109] Joseph B. Bak-Coleman *et al.* 2021. Stewardship of global collective behavior. In Proceedings of the National Academy of Sciences, Vol. 118, No. 27, e2025764118.
[110] Meghan B. Moran, Sheila T. Murphy, Lauren B. Frank, and Lourdes Baezconde-Garbanati. 2015. The ability of narrative communication to address health-related social norms. International Review of Social Research.
[111] Adebanke L. Adebayo, Rochelle Davidson Mhonde, Nathaniel DeNicola, and Edward Maibach. 2020. The effectiveness of narrative versus didactic information formats on pregnant women's knowledge, risk perception, self-efficacy, and information seeking related to climate change health risks. International Journal of Environmental Research and Public Health, Vol. 17, No. 19, 6969.
[112] Matthew W. Kreuter *et al.* 2010. Comparing narrative and informational videos to increase mammography in low-income African American women. Patient Education and Counseling, Vol. 81, S6-S14.
[113] Josephine Pui-Hing Wong *et al.* 2019. Exploring the use of fact-based and story-based learning materials for HIV/STI prevention and sexual health promotion with South Asian women in Toronto, Canada. Health Education Research, Vol. 34, No. 1, 27-37.
[114] Manfred Schmitt and Gabriela S. Blum. 2020. State/Trait Interactions, in Encyclopedia of Personality and Individual Differences, Virgil Zeigler-Hill and Todd K. Shackelford Eds. Cham: Springer International Publishing, 5206-5209.
[115] Jennifer Fries Taylor, Jodie Ferguson, and Pamela Scholder Ellen. 2015. From trait to state: understanding privacy concerns. Journal of Consumer Marketing, Vol. 32, No. 2, 99-112. https://doi.org/10.1108/JCM-07-2014-1078.
[116] Melanie C. Green and Timothy C. Brock. 2000. The role of transportation in the persuasiveness of public narratives. Journal of Personality and Social Psychology, Vol. 79, No. 5, 701.
[117] Julie S. Downs. 2014. Prescriptive scientific narratives for communicating usable science. In Proceedings of the National Academy of Sciences, Vol. 111, No. 4, 13627-13633.
[118] Ayebo E Sadoh, Clement Osime, Damian U Nwaneri, Bamidele C Ogboghodo, Charles O Eregie, and Osawaru Oviawe. 2021. Improving knowledge about breast cancer and breast self examination in female Nigerian adolescents using peer education: a pre-post interventional study. BMC Women's Health, Vol. 21, No. 1, 1-9.
[119] Purnima Menon *et al.* 2020. Lessons from using cluster-randomized evaluations to build evidence on large-scale nutrition behavior change interventions. World Development, Vol. 127, 104816.
[120] Howard S. Muscott, Eric L. Mann, and Marcel R. LeBrun. 2008. Positive behavioral interventions and supports in New Hampshire: Effects of large-scale implementation of schoolwide positive behavior support on student discipline and academic achievement. Journal of Positive Behavior Interventions, Vol. 10, No. 3, 190-205.

APPENDICES

A INTERVENTION EDUCATION VIDEO TRANSCRIPTS

A.1 CONCEPT CONDITION

Figure 6: Screenshot from Concept experimental manipulation video.

You might have heard of the term pirating, where people take copyrighted content without permission. But have you heard of privacy pirating? Privacy pirating happens when people take and then share other people's photos or information on social media without permission and without considering the other person's privacy preferences.

It is common for people to post photos and information about their friends, family, and other people they know. However, doing so without permission can turn a friend or family member into a victim of privacy pirating, if they feel the information or photos shared were embarrassing, private, or simply makes them feel uncomfortable.

Privacy pirating can also occur when people re-share stranger's photos and information they find online, contributing to the spread of another person's private information. A popular way to share content about other people is through memes, which often include photos of people with an entertaining or funny caption. Even though memes may be funny, touching, or relatable, it is difficult to know if the person in a meme provided permission for their image to be used by the public.

Victims of privacy pirating may be left feeling helpless as they wonder who will see their information, what others will say about them, and how the information may affect their relationships. In

addition to experiencing significant distress, some victims may become targets of harassment, and suffer other negative personal and professional consequences.

It's important to remember that every person has their own level of comfort when it comes to what they consider too private or embarrassing to share on social media. Moreover, even if someone shares information about themselves, they may not have intended for it to be re-shared by other people. When you post online without considering other people's privacy preferences, you could be creating a victim of privacy pirating.

The goal of this video is to raise awareness about 'privacy pirating.' We hope you have learned about its impact and will share this information with others.

A.2 FACTS CONDITION

Note that 'facts' were embedded in the 'concept' condition video to provide context about IDP or 'privacy pirating.' Thus, content specific to the facts condition are italicized below.

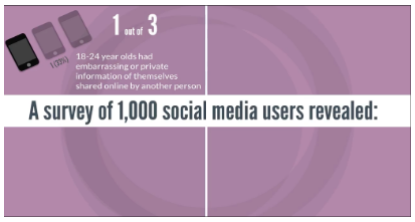

Figure 7: Screenshot from Facts experimental manipulation video.

You might have heard of the term pirating, where people take copyrighted content without permission. But have you heard of privacy pirating? Privacy pirating happens when people take and then share other people's photos or information on social media without permission and without considering the other person's privacy preferences.

It is common for people to post photos and information about their friends, family, and other people they know. However, doing so without permission can turn a friend or family member into a victim of privacy pirating, if they feel the information or photos shared were embarrassing, private, or simply makes them feel uncomfortable.

A survey of 1000 social media users revealed:

One third of users 18-24 had embarrassing or private information of themselves shared online by another person;

30% experience decreased mental health and quality of life due to their social media image

60% support platform restrictions on the re-posting of content to better control their privacy

Less than 10% believed nothing should change about other's sharing practices

Privacy pirating can also occur when people re-share stranger's photos and information they find online, contributing to the spread of another person's private information. A popular way to share content about other people is through memes, which often include photos of people with an entertaining or funny caption. Even though memes may be funny, touching, or relatable, it is difficult to know if the person in a meme provided permission for their image to be used by the public.

Victims of privacy pirating may be left feeling helpless as they wonder who will see their information, what others will say about them, and how the information may affect their relationships. In addition to experiencing significant distress, some victims may become targets of harassment, and suffer other negative personal and professional consequences.

Last year, there were over 10,000 police reports in the US related to online harassment or privacy violations from sharing of photos and videos. Majority of those victims reported new or worsening symptoms for anxiety and depression disorders;

60% of those victims said photos were initially shared by friends or a social acquaintance without consent.

It's important to remember that every person has their own level of comfort when it comes to what they consider too private or embarrassing to share on social media. Moreover, even if someone shares information about themselves, they may not have intended for it to be re-shared by other people. When you post online without considering other people's privacy preferences, you could be creating a victim of privacy pirating.

The goal of this video is to raise awareness about 'privacy pirating.' We hope you have learned about its impact and will share this information with others.

A.3 NARRATIVE CONDITION

Note that 'narratives' followed the 'concept' condition video to provide context about IDP or 'privacy pirating.'

A.3.1 Mateo: Admired Cashier. This experience is based on a true story with names changed.

Figure 8: Screenshot from Narrative experimental manipulation video.

Mateo woke up to hundreds of new friend requests. His first thought was he was targeted by an online scam. He received a text from a co-worker that said a picture of him had gone viral and he was being called the "World's Sexiest Cashier". Mateo was initially flattered and quickly accepted his new friend requests.

However, followers quickly identified his girlfriend from older posts and began online attacks. Admirers were showing up at his job to take pictures with him. It became so disruptive he could no longer work the register. And his girlfriend could not handle the unjustified hate.

Mateo had to quit his job to escape the constant attention. And his girlfriend broke-up with to get away from the online harassment. The original customer may have believed they were sharing

their admiration of Mateo's good looks. However, it led to rapid international recognition and unwanted attention for him and others in his life.

A.3.2 Jessica: Don't Do Drugs. This experience is based on true stories with names changed.

Figure 9: Screenshot from Narrative experimental manipulation video.

One of the scariest moments in Jessica's life went viral. She was experiencing terrible headaches and dizziness because of the flu and drove herself to the emergency room. But before she could make it inside, she fainted.

A passerby took a photo of Jessica and shared it online with the caption "DON'T DO DRUGS" assuming that she was passed out from an overdose. Her photo rapidly spread on social media sites. She endured people laughing at her and received concerned calls from friends, family and co-workers.

Jessica spent weeks battling the lie that she used drugs, almost losing her job during the ordeal. The passerby who shared the picture did not consider Jessica's privacy, nor the impact of misrepresenting Jessica's medical emergency.

B INTERVENTION RESPONSE QUESTIONNAIRE

Client Satisfaction Questionnaire adapted to Internet-based interventions [104] (CSQ-I) and Social Media Privacy Perception Questions. Note: * indicates questions adopted from CSQ-I

1. Rate your knowledge of privacy pirating on social media before the course.*

•Not at all, Slightly, Moderately, or Extremely knowledgeable

2. Rate your knowledge of privacy pirating on social media now after the course.*

•Not at all, Slightly, Moderately, or Extremely knowledgeable

3. How common do you think it is for people to post private or embarrassing information (or photos) about others on social media without permission?

•Not, Slightly, Somewhat, Very or Extremely common

4. How serious of a problem is it when people post private or embarrassing information (or photos) about others on social media without permission?

•Not, Slightly, Somewhat, Very or Extremely serious

5. Do you think that social media users should take more or less precautions to reduce the number of inappropriate posts on social media (e.g., posts that include embarrassing or private information about others)?

•No, Fewer, Same, More or Many more precautions

6. Who ought to be responsible for educating social media users about the dangers of posting embarrassing or private photos and information about other people without permission? Select all that apply.

•Parents/guardians, Peers or friends, Schools, Social media platforms or Governments (e.g., social programs)

7. Who do you believe should be responsible for managing inappropriate private content on social media that is posted without permission? Select all that apply.

•The original person who posted, Other social media users, Social media platforms or Governments (e.g., social programs)

8. I think this video covered an important topic.*

•Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

9. I wish more people knew about this information.*

•Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

10. I think the videos and materials were high quality.*

•Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

11. This information would be effective in changing people's behaviors.*

•Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree

12. How could this material be improved to make it a more effective learning experience? (Open-ended)*

C READING-BASED ATTENTION CHECKS

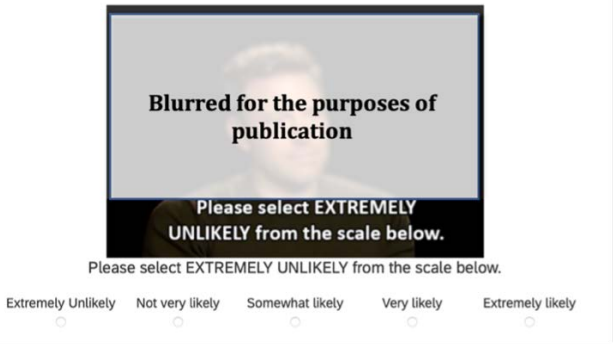

Figure 10: Example A of reading-based attention check

Figure 11: Example B of reading-based attention check

D QUALITATIVE ANALYSIS CODEBOOK

Table 4: Frequency of themes identified for optional open-ended question: How could this material be improved to make it a more effective learning experience?

Theme	Description	Example	Responses by Condition Group			
			Narrative	Facts	Concept	Total
Strategies to Improve Intervention Effectiveness						
1. Recommend Disseminating Content	Encouraged to share intervention videos with larger audience, including specific strategies or target audience	*"spread this concept to each and every one whom use the social media platforms"*	8	9	9	26
1.a Included recommendation on sharing strategy		*"Maybe post it on social media sites themselves"*	3	3	2	8
1.b Included recommendation on target audience		*"by teaching it to the people who is not aware of it"*	1	3	3	7
2. More Relatable Content and Stories	Requested content that includes "real word examples", more current material or greater focus on impact to sharer and victims	*"show an example that is relatable"*	14	22	29	65
3. Include Prevention Strategies	Requested content that addresses personal actions to prevent IDVPs from multiple user roles	*"provide information on how to mitigate the issue"*	5	8	6	19
Training Experience						
4 Video Quality & Training Experience	Comment on specific aspects of the intervention videos that impact the experience (not directly content)		12	13	16	41
4.a Add Interactive component and more knowledge checks		*"more interactive elements"*	3	6	3	12
4.b Nonspecific recommendation for more examples		*"more examples"*	5	0	4	9
4.c Video audio experience		*improve the audio quality"*	3	1	3	7
4.d Video length or pace		*""make the video little slow"*	1	5	5	11
4.e Video visual style		*""The material should have more animation rather than the still picture"*	2	2	0	4
Satisfaction with Material						
5 Critical Feedback	Participant expressed dissatisfaction with intervention or doubts regarding effectiveness	*"different concern from many people's opinion was too much"*	2	1	3	6
5.a No improvements, but doubt it will change behaviors		*"The material was fine, I just don't think people will care enough to stop"*	1	1	3	5
6. No Improvements	Satisfied with content in current state	*"i think it was enough to understand. I don't know, it seems pretty effective to me"*	12	17	9	38
7. Statement of Positive Experience	Expressed satisfaction with training, but did not explicitly state no improvements required	*"Nothing to improve. It's great"*	19	13	13	45

[a] Note: Core theme counts are inclusive of subtheme responses.