Disclosure by Design: How Dark Patterns Reduce Users' Social Privacy

Dominique Kelly[1,*] and Jacquelyn Burkell[1]

[1] *Western University, London, Ontario, Canada*

Abstract

Privacy dark patterns are interface design tactics deployed by online services to influence users to make choices that reduce their online privacy. While some dark patterns undermine users' *institutional privacy* by facilitating the collection and use of their personal data by institutions, others weaken users' *social privacy* by increasing other people's access to that data. In this study, we focused specifically on how the design of social networking sites (SNSs) popular with teens (Snapchat, TikTok, Instagram, Twitter, and Discord) steer users to make choices that reduce their social privacy. To this end, we recorded attempts to register an account, configure account settings, and log in and out for each of the five SNSs in our sample. We then coded the recordings for the presence of design strategies that could influence users to take actions that increase other people's access to their personal data. As a result of our content analysis, we identified two major types of dark patterns that reduce users' social privacy (*Obstruction* and *Obfuscation*) and seven subtypes. We discuss why social media companies are incentivized to promote social sharing among users through design, as well as the challenges associated with regulating privacy dark patterns.

Keywords

dark patterns, social privacy, interface design, social media

1. Introduction

Manipulative interface design tactics, or dark patterns [1], are common online [2] and effective in influencing user behaviour [3, 4]. Privacy dark patterns refer to a subset of dark patterns specifically crafted to steer people to make choices that reduce their online privacy [5, 6]. Research examining privacy dark patterns in the context of social media shows, for instance, that some social networking sites (SNSs) suggest the user's account to off-site connections by default [7], frame opting into facial recognition technology positively [8], or require unnecessary additional clicks to opt out of targeted advertising [9].

Some privacy dark patterns reduce users' *institutional privacy* [10] by facilitating the collection and use of their personal data by institutions. Other privacy dark patterns weaken users' *social privacy* [10] by nudging them to make choices that increase other people's access to their personal data. A setting that controls targeted advertising affects users'

Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices Workshop at CHI conference on Human Factors in Computing Systems, May 12, 2024, Honolulu, HI (Hybrid Workshop)

* Corresponding author.

dkelly48@uwo.ca (D. Kelly); jburkell@uwo.ca (J. Burkell)

0000-0003-1222-6237 (D. Kelly); 0000-0003-2645-8127 (J. Burkell)

institutional privacy because this setting pertains to platform collections and use of personal data, whereas a setting that controls whether posts are public or private affects users' social privacy because that setting determines the visibility of user posts.

The risks and harms associated with weakened social privacy include identify theft, stalking, embarrassment, and blackmail [11], as well as user regret [12] and reputational damage [13]. Research shows that perceptions of social norms on SNSs can affect users' own levels of self-disclosure [14, 15, 16, 17, 18, 19]. If dark patterns nudge users to accept settings that allow much of their data to be exposed to a wide audience, other users will presumably perceive this level of sharing to be acceptable and, in turn, share more of their own data with more people. Social media companies promote this social sharing among users because they profit from the data users disclose [17, 20]. The privacy policies and/or terms of services for many major social media companies give them the right to collect, and in certain cases use, the user-generated data that is posted on their platforms [21, 22, 23, 24, 25].

This study is part of a broader research agenda to identify privacy dark patterns on popular SNSs and determine how teens perceive and respond to these strategies. Our decision to focus on teens was driven by the vulnerability of this demographic to the effects of dark patterns on their privacy decision-making, as well as the dearth of research, to date, that has examined young people's reactions to manipulative design tactics [26]. In this study, we document dark patterns that specifically undermine users' social privacy on a sample of SNSs popular among teens: Snapchat, TikTok, Instagram, Twitter, and Discord [27]. We recorded attempts to register an account, configure account settings, and log in and out for each of the five SNSs, and coded the recordings for the presence of design strategies that could influence users to make choices that increase other people's access to their personal data, either by expanding the scope of data shared or by broadening the size of the audience to whom the information is revealed.

We identified two major types of dark patterns that undermine users' social privacy (*Obstruction* and *Obfuscation*) and seven subtypes. Our findings show that SNSs often deploy privacy dark patterns in combinations that complement and reinforce one another. We discuss why social media companies are incentivized to promote users' social sharing as well as the challenges faced by regulators aiming to combat these manipulative practices.

2. Background

2.1. Social Privacy Invasions and the Role of SNS Design

The term "dark patterns" was coined in 2010 [1], but researchers, journalists, and users alike have long recognized that SNSs can deliberately undermine people's social privacy through design. Many of Facebook's privacy incidents have been tied to their use of overly permissive defaults; users have routinely been forced to opt out of, rather than into, settings and features that enable the widespread sharing of their personal data. In 2006, for instance, Facebook automatically enrolled all members in the News Feed, a feature that served users a stream of updates about the actions taken by their Friends, such as newly uploaded pictures and changes in relationships [28, 29]. Following much dissent, Facebook introduced privacy settings to control what would be shown on people's News Feeds [28].

A similar controversy ensued a year later with the launch of Beacon, an advertising platform that "shared users' actions with external partner Web sites via the News Feed" [28] (para. 13). Because users were enrolled in the feature by default, many were surprised to see information about their off-site purchases broadcast to their Facebook Friends [28, 29]. Again, user outrage forced Facebook to backtrack, this time by converting Beacon to an opt-in model and tweaking its privacy notice; eventually, Beacon was simply discontinued [29].

While the News Feed and Beacon incidents were relatively short-lived, Facebook's privacy settings have been the subject of ongoing concern and debate. Writing in 2010, boyd and Hargittai note that every time Facebook "introduced new options for sharing content, the default was to share broadly" [28] (para. 11). For instance, when Facebook prompted users to reconsider their privacy settings in December 2009, all of the defaults for sharing content were set to "Everyone" instead of "Old settings," and users were forced to respond before they could access the rest of the site [28].

Other sites have also raised privacy concerns. In 2010, for example, Google launched Buzz, an SNS intended to compete directly with Facebook. Without offering "prior notice or the opportunity to consent," Buzz automatically set up Gmail users with "followers" drawn from their email contact lists and made this information publicly accessible to anyone viewing a user's profile [29] (p. 1386). Although Google implemented more granular privacy controls in response to user complaints, the company was forced to settle a class action lawsuit and Federal Trade Commission (FTC) complaint, and announced in fall 2011 that it was officially discontinuing Buzz [29]. These incidents reveal that public concern for privacy-undermining design tactics predates both the popularization of the term "dark patterns" and the recent surge in academic and regulatory attention devoted to these strategies.

2.2. Risks and Harms Associated with Social Privacy Invasions

Increasing other people's access to one's personal data on social media exposes users to a range of potential risks and harms, including identify theft, stalking, embarrassment, and blackmail [11]. Users may experience regret over posts that contain sensitive content, strong sentiments, and lies or secrets [12]. In some cases, revealing personal data to others can result in long-term reputational damage [13]. This risk is vividly illustrated by accounts of small missteps on social media provoking online shaming campaigns that aim to humiliate and discipline their chosen targets [30]. An ill-conceived comment or photo, posted impulsively or many years ago, could have an enduring effect on one's reputation and prospects; as Solove notes, people grow and change, but the disclosure of information to others risks making "a person 'a prisoner of his recorded past'" [13] (p. 145).

The fact that a user's online audience may be much larger than expected contributes to, and sometimes directly causes, these issues. Social media users "consistently underestimate the audience size for their posts" [31] (p. 29), and their "imagined audience" [32] (p. 115) may be quite different from the actual group of people they are sharing with. As several of Facebook's privacy incidents have proven, SNSs may deliberately promote widespread data sharing by setting defaults for content to be shared publicly rather than only with one's direct connections (e.g., Facebook Friends). Furthermore, people "tend to accept friend requests from weak ties" due to factors like social pressure [33] (p. 3). Consequently, even

if the visibility of a user's information is limited to direct connections, users may nonetheless be sharing with acquaintances or even strangers alongside family members and close friends [33].

These risks and harms are particularly salient for teens, whose active usage of social media [34], "susceptibility to peer pressure," and "limited capacity for self-regulation" [35] (p. 800) may lead them to publicly share information that could have a lasting negative impact on their lives.

3. Methods

In this study, we aimed to identify dark patterns that undermine users' social privacy on a sample of popular SNSs. Following the methodological approach adopted in prior studies [2, 7, 36], we recorded user-site interactions and analyzed the recordings for the presence of dark patterns.

3.1. Data Collection

The first author recorded her attempts to register an account, configure account settings, and log in and out on five SNSs popular with teen users (Snapchat, TikTok, Instagram, Twitter, and Discord) [27] from March to May 2022. A temporary Protonmail email account was used to sign up for all accounts. Snapchat and TikTok were downloaded to a mobile device (a Samsung Galaxy tablet), while Instagram, Twitter, and Discord were accessed through a desktop browser (Google Chrome). Recordings were captured by a built-in Android screen recorder on the mobile device and Windows Game Bar on the desktop browser.

For each procedure, the first author adhered to the following protocol to ensure consistency in her actions across all five sites:

- Registering an account: Review terms and conditions; fill in mandatory fields; select privacy-friendly options when possible; log out.
- Configuring account settings: Log in; navigate to account settings; review all settings chronologically; adjust settings to be more privacy-friendly when possible; log out.
- Logging in and out: Log into the account three additional times on separate dates; select privacy-friendly options when possible; log out.

The first author attempted to make the most privacy-friendly choices possible with respect to both social and institutional privacy, meaning that she selected options that limited other people's access to the user's personal data, as well as the site's collection and use of that data. The user-SNS interactions for each procedure were carried out and recorded on separate days.

3.2. Data Analysis

In line with the basic approach for content analysis outlined by Krippendorff [37], the data were coded for the presence of design strategies that could influence users to make privacy-

invasive choices (i.e., "privacy dark patterns"). For the purposes of this study, we considered choices to be privacy-invasive if they increased other people's access to the user's personal data – by enlarging the audience and/or increasing the range of data exposed – thereby reducing the user's social privacy. Thus, an account with highly privacy-invasive settings would expose a wide range of the user's personal data (e.g., full name, birthdate, phone number, online status, etc.) to a large audience (all site members or even the general public).

When coding, we specifically focused on design strategies that increased the user's workload, confused or mislead the user, or depicted certain choices positively or negatively through language and/or visuals (adapted from the broad modes of influence described in prior work [38]). We paid close attention to visual and verbal UI design elements (e.g., buttons, text, pop-ups, pre-selected options, and images) as well as stylistic choices like size, colour, contrast, and placement that enhanced the visibility of or attention to privacy-invasive options.

We imported the recording files into a software program for qualitative data analysis (NVivo) and manually assigned codes, representing privacy dark patterns, to temporal segments of the recordings. This original set of codes was based on notes taken by the first author immediately after each recording. Through an iterative process, we re-reviewed the recordings, added new codes to capture strategies missed in our initial inspection, and updated our codebook to include these additions. We ultimately identified seven privacy dark pattern subtypes and thematically organized these patterns into two major types (*Obstruction* and *Obfuscation*). The codebook contains detailed instructions for coders as well as clear descriptions of each privacy dark pattern subtype.

In parallel, we coded the data for the presence of dark patterns that could undermine users' institutional privacy. This separate work is the focus of a study currently under review.

4. Results

As a result of our content analysis of the recordings, we identified two major types of dark patterns and seven subtypes that compromise users' social privacy in SNSs. Table 1 summarizes the major types and subtypes.

Table 1
Typology of dark patterns that undermine users' social privacy in SNSs, adapted from our broader typology in a working paper [39]

Privacy dark pattern types and subtypes	Description
1 Obstruction	The site increases the effort that users must exert to make a privacy-friendly choice (i.e., by requiring more actions).
1.1 Defaults	Privacy-invasive options are selected by default prior to user interaction, requiring the user to locate and change them.
1.2 Confirmations	Attempts to make privacy-friendly choices are accompanied by pop-ups that require the user to confirm their decision by clicking an additional button.

Privacy dark pattern types and subtypes	Description
1.3 Interruptions	Pop-ups asking the user to make a privacy-invasive choice appear and must be manually dismissed. The requests are irrelevant to the user's current activity.
1.4 Missing Bulk Options	Three or more closely-related privacy-invasive defaults are presented together without a corresponding bulk option (e.g., a "reject all" button).
2 Obfuscation	The site obscures, hides, or omits relevant information and options, with the intent of confusing or misleading the user into making a privacy-invasive choice.
2.1 Attention Manipulation	The buttons for privacy-invasive choices are given greater salience than privacy-friendly choices through their size, colour, placement, and/or contrast.
2.2 False Private Account	The user is given the opportunity to set their account to "private," but enacting this setting does not alter all privacy-invasive defaults.
2.3 Concealed Settings	After account registration, the site does not suggest that the user check their account settings to ensure that the current defaults align with the user's preferences.

4.1. Privacy Dark Patterns in Account Registration, Configuring Account Settings, and Logging in and Out

The sections below summarize how sites in our sample employed the privacy dark patterns identified in our typology across the procedures of registering an account, configuring account settings, and logging in and out of the account.

4.1.1. Registering an Account

During account registration, two sites allowed users to adjust settings that controlled other people's access to their personal data. Twitter, for example, gave users the option to protect their Tweets, meaning that their posts would only be visible to their followers. These Tweets, however, were public by default (representing our *Defaults* dark pattern), and the user needed to tick a relatively small box to enact the privacy-protective setting. Other examples of defaults employed by Twitter included pre-ticked boxes that allowed other site members to find the user by their phone number and email address. Moreover, the site design served to distract the user from changing these defaults (an example of our *Attention Manipulation* dark pattern): a high-contrast "sign up" button enticed the user to complete the registration process, while a link leading to the pre-ticked boxes was hidden in a block of fine print. The user could easily miss this link, and therefore not review the default selections. As a result, the user could remain unaware that their account will be discoverable by their email and phone contacts.

Another site, Snapchat, employed *Attention Manipulation* to push the user to sync their contacts (i.e., upload information about their contacts from their device's address book) to "Find [their] Friends." Syncing contacts allows other site members to find the user's account based on their contact information, meaning that an account that otherwise contains no identifying information could be linked to the user's real identity. The button to sync contacts naturally attracted the user's attention through its large size, bright colour, and central placement; meanwhile, the button to skip this step was small, faint, and hidden in the top right corner of the interface. Consequently, some users – especially those eager to finish signing up – might not realize that an option to skip syncing contacts, and therefore restrict the ability of other users to find their account, is available.

4.1.2. Configuring Account Settings

All five sites in our sample used the *Concealed Settings* dark pattern, meaning that users were not prompted to check their account settings after registering an account. Considering that all five sites pre-selected some options that reduced the user's social privacy (*Defaults*), failing to remind users to check their account settings could result in more of the user's data being exposed – to more people – than anticipated or desired. The permissive defaults set by the sites controlled features including user-to-user interactions (who could mention or tag the user, send them direct messages, and share their posted content), the visibility of information about the user (such as their birthday, online status, and on-site activities), and the discoverability of the user's account (by allowing other site members to find the user through their phone number or email address, or suggesting the user's account to others). For example, Instagram allowed Tags from "Everyone" by default, when more privacy-friendly options (i.e., "People You Follow" and "No One") were also available. These permissive defaults weakened the user's social privacy, as they could allow other people to find and learn about the user based on content that they did not personally post.

Two sites gave users the option to set their account to "Private," but enacting this setting confusingly only altered a few aspects of the user's privacy – a dark pattern we term *False Private Account*. For example, Instagram's "Private Account" mode, which had to be manually selected, restricted the audience for the user's photos and videos to only people they approved. However, other privacy-related defaults, including those that allowed other people to see the user's activity status and share the user's Stories as messages, remained enacted. These defaults revealed the time when the user was last active to other site members, and permitted those members to share the user's Stories with a wider audience than the user may have expected when they initially uploaded that content. Similarly, even after its "Private account" mode, which was untoggled by default, was clicked, TikTok kept in place four defaults that allowed the user's account to be widely suggested to others – including phone contacts, Facebook friends, people with mutual connections, and people who sent links to the user or opened links sent to them. These settings promoted a wider audience for the user's posts and, in some cases (i.e., phone contacts and Facebook friends) enabled other site members to link the user's real identity to their account. To further complicate the process of opting out, each closely-related default needed to be manually de-selected, representing a dark pattern we call *Missing Bulk Options*.

Two sites, employing a strategy we refer to as *Confirmations*, required the user to confirm their choice when attempting to enact a privacy-friendly setting. For example, if the user attempted to protect their Tweets (i.e., restrict the audience for their posts and account information) in their account settings on Twitter, a pop-up appeared asking them to confirm their choice by clicking an additional button.

4.1.3. Logging In and Out

Two sites interrupted the user at login with prompts to add personal data or enact settings that would undermine their social privacy, a strategy we call *Interruptions*. Discord presented a pop-up to users at login that encouraged them to add their school email address, potentially allowing them to connect with other people known in-person through an otherwise anonymous account. TikTok similarly asked users to sync their Facebook friend list through a pop-up, simultaneously encouraging the user to expand the audience for their posts and making it possible for other site members to link the user's real identity to their account. Notably, TikTok also made the option to sync contacts ("OK") slightly more salient than the option to decline ("Don't allow"), representing a subtle instance of *Attention Manipulation*.

5. Discussion

Our findings reveal that SNSs utilize a variety of design strategies to encourage users to share a large amount of personal data with a wide audience. These strategies can be categorized by two primary modes of influence: requiring users to exert unnecessary effort to protect their privacy, and confusing and/or misleading users into maintaining or enacting privacy-invasive settings. The privacy dark patterns we observed were often deployed in combinations that reinforced one another; for instance, all five sites selected at least some defaults that weakened the user's social privacy and failed to prompt users to check their account settings – where they could review and change those defaults – after registering an account.

In the sections below, we discuss why social media companies are incentivized to promote social sharing as well as the challenges associated with regulating these manipulative practices.

5.1. Social Media Companies' Interest in Promoting Social Sharing

There is some evidence that perceptions of social norms on SNSs can impact users' own levels of self-disclosure [14, 15, 16, 17, 18, 19]. For example, Lewis et al. found that students were more likely to have a private Facebook profile if this setting was shared by their friends and roommates [17], and Burke et al. determined that "newcomers who see their friends contributing [to Facebook] go on to share more content themselves" [14] (p. 945). Presumably, when dark patterns push users to accept settings that allow large volumes of their personal data to be accessed by a wide audience, other users will in turn perceive this broadscale social sharing to be appropriate and share more of their own data.

Social media companies are incentivized to promote this social sharing because the data users disclose is of value to them [17, 20]. People's traits and attributes – such as their age, gender, and preferences, as well as online comments and social media photos – "are increasingly regarded as business assets that can be used to target services or offers, provide relevant advertising, or be traded with other parties" [20] (p. 494). A survey of the privacy policies and/or terms of service for the five SNSs in our study conducted in August 2023 revealed that these social media companies have the right to collect and, in some cases, use user-generated data that is posted on their platforms [21, 22, 23, 24, 25].

5.2. Challenges to Regulating Privacy Dark Patterns

In recent years, the need to protect children and teens from the effects of dark patterns has received some regulatory attention. For example, in 2022, the FTC engaged in enforcement actions against Epic Games Inc. for its alleged violation of the COPPA (Children's Online Privacy Protection Act) Rule "by collecting personal information from children under 13 who played Fortnite, a child-directed online service, without notifying their parents or obtaining their parents' verifiable consent" [40] (para. 7). Moreover, Epic allegedly "violated the FTC Act's prohibition against unfair practices by enabling real-time voice and text chat communications for children and teens by default" [40] (para. 7).

In general, however, existing regulatory frameworks across Canada, the United States, and the European Union do not specifically target dark patterns. One notable exception, the California Privacy Rights Act, "defines a dark pattern as 'a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision-making, or choice, as further defined by regulation'" [41] (para. 14). In practice, regulating dark patterns remains a complicated task. Many regulatory frameworks rely on user reporting mechanisms to address privacy issues – a critical weakness given that dark patterns are, by design, hard to detect [2, 42]. Moreover, the point at which a dark pattern clearly subverts or impairs user autonomy is open to debate, and even when specific patterns appear relatively benign, their effects may be amplified when deployed in combinations that complement and reinforce one another.

6. Conclusion

This work is part of a broader research project to document privacy dark patterns on SNSs and ascertain how teens perceive and respond to these strategies. In this study, we focused on how the design of SNSs can influence users to make choices that reduce their social privacy. Specifically, we content-analyzed recordings of user-site interactions on five SNSs popular among teens for evidence of design strategies that could steer users to take actions that increase other people's access to their personal data. We identified seven privacy dark patterns that weakened users' social privacy and found that these strategies were often deployed in mutually-reinforcing combinations. As institutions that profit from the collection of users' data, social media companies are incentivized to deploy privacy dark patterns that increase social sharing. These is a need for regulatory frameworks to target privacy dark patterns, although, in practice, combatting these tactics presents a clear challenge.

Acknowledgements

This project has been funded by the Office of the Privacy Commissioner of Canada (OPC); the views expressed herein are those of the author(s) and do not necessarily reflect those of the OPC.

References

[1] H. Brignull, Dark patterns, 2010. URL: https://www.darkpatterns.org/

[2] L. Di Geronimo, L. Braz, E. Fregnan, F. Palomba, A. Bacchelli, UI dark patterns and where to find them: A study on mobile applications and user perception, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1-14. URL: dl.acm.org/doi/pdf/10.1145/3313831.3376600

[3] J. Luguri, L. J. Strahilevitz, Shining a light on dark patterns, Journal of Legal Analysis 13, 1 (2021) 43-109. URL: https://doi.org/10.1093/jla/laaa006

[4] M. Nouwens, I. Liccardi, M. Veale, D. Karger, L. Kagal, Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–13. URL: https://doi.org/10.1145/3313831.3376321

[5] C. Bösch, B. Erb, F. Kargl, H. Kopp, S. Pfattheicher, Tales from the dark side: Privacy dark strategies and privacy dark patterns, in: Proceedings on Privacy Enhancing Technologies 4, 2016, pp. 237-254. URL: https://doi.org/10.1515/popets-2016-0038

[6] L. Fritsch, Privacy dark patterns in identity management, in: L. Fritsch, H. Roßnagel, D. Hühnlein (Eds.), Open Identity Summit 2017: Proceedings, 2017, pp. 93-104. URL: http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-63722

[7] T. Mildner, G.-L. Savino, P. R. Doyle, B. R. Cowan, and R. Malaka, About engaging and governing strategies: A thematic analysis of dark patterns in social networking services, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1-15. URL: https://doi.org/10.1145/3544548.3580695

[8] Forbrukerrådet, Deceived by design: How tech companies use dark patterns to discourage us from exercising our rights to privacy, 2018. URL: fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf

[9] European Data Protection Board, Dark patterns in social media platforms: How to recognise and avoid them, 2022. URL: https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-32022-dark-patterns-social-media_en

[10] K. Raynes-Goldie, Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook, First Monday 15, 1 (2010). URL: https://doi.org/10.5210/fm.v15i1.2775

[11] R. Gross, A. Acquisti, Information revelation and privacy in online social networks, in: Proceedings of the 2005 ACM workshop on Privacy in the Electronic Society, 2005, pp. 71-80. URL: https://doi.org/10.1145/1102199.1102214

[12] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, L. F. Cranor, "I regretted the minute I pressed share": A qualitative study of regrets on Facebook, in: Proceedings of

the Seventh Symposium on Usable Privacy and Security, 2011, pp. 1-16. URL: https://doi.org/10.1145/2078827.2078841
[13] D. Solove, Understanding privacy, Harvard University Press, Cambridge, MA, 2008.
[14] M. Burke, C. Marlow, T. Lento, Feed me: Motivating newcomer contribution in social network sites, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2009, pp. 945-954. URL: https://doi.org/10.1145/1518701.1518847
[15] D. Chang, E. L. Krupka, E. Adar, A. Acquisti, Engineering information disclosure: Norm shaping designs, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016, pp. 587-597. URL: https://doi.org/10.1145/2858036.2858346
[16] C. Cheung, Z. W. Y. Lee, T. K. H. Chan, Self-disclosure in social networking sites: The role of perceived cost, perceived benefits and social influence, Internet Research 25, 2 (2015) 279-299. URL: https://doi.org/10.1108/IntR-09-2013-0192
[17] K. Lewis, J. Kaufman, N. Christakis, The taste for privacy: An analysis of college student privacy settings in an online social network, Journal of Computer-Mediated Communication 14, 1 (2008) 79-100. URL: https://doi.org/10.1111/j.1083-6101.2008.01432.x
[18] P. K. Masur, D. DiFranzo, N. N. Bazarova, Behavioral contagion on social media: Effects of social norms, design interventions, and critical media literacy on self-disclosure, PLOS ONE 16, 7 (2021) 1-20. URL: https://doi.org/10.1371/journal.pone.0254670
[19] P. K. Masur, N. N. Bazarova, D. DiFranzo, The impact of what others do, approve of, and expect you to do: An in-depth analysis of social norms and self-disclosure on social media. Social Media + Society 9, 1 (2023) 1-14. URL: https://doi.org/10.1177/20563051231156401
[20] A. Acquisti, C. Taylor, L. Wagman, The economics of privacy. Journal of Economic Literature 54, 2 (2016) 442-492. URL: https://doi.org/http://dx.doi.org/10.1257/jel.54.2.442
[21] Discord, Discord's terms of service, 2023. URL: https://discord.com/terms#5
[22] Instagram, Privacy policy, 2023. URL: https://privacycenter.instagram.com/policy/version/5601719079934335/
[23] Snapchat, Privacy policy, 2023. URL: https://values.snap.com/privacy/prior-privacy-policy-08-15-2023
[24] TikTok, Terms of service, 2021. URL: https://www.tiktok.com/legal/page/row/terms-of-service/en
[25] Twitter, Privacy policy, 2023. URL: https://twitter.com/en/privacy/previous/version_19
[26] D. Fitton, J. C. Read, Creating a framework to support the critical consideration of dark design aspects in free-to-play apps, in: Proceedings of the 18th ACM International Conference on Interaction Design and Children, 2019, pp. 407-418. URL: https://doi.org/10.1145/3311927.3323136
[27] Statista, Most popular social networks of teenagers in the United States from fall 2012 to fall 2020, 2021. URL: https://www.statista.com/statistics/250172/social-network-usage-of-us-teensand-young-adults/

[28] d. boyd, E. Hargittai, Facebook privacy settings: Who cares? First Monday 15, 8 (2010). URL: https://firstmonday.org/ojs/index.php/fm/article/download/3086/2589
[29] I. S. Rubinstein, N. Good, Privacy by design: A counterfactual analysis of Google and Facebook privacy incidents, Berkeley Technology Law Journal 28 (2012) 1333-1414. URL: https://www.jstor.org/stable/24119897
[30] J. Ronson, So you've been publicly shamed, Riverhead Books, 2015.
[31] F. D. Stutzman, R. Gross, A. Acquisti, Silent listeners: The evolution of privacy and disclosure on Facebook, Journal of Privacy and Confidentiality 4, 2 (2013) 7-41. URL: https://doi.org/10.29012/jpc.v4i2.620
[32] A. E. Marwick, d. boyd, I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience, New Media & Society 13, 1 (2010) 114-133. URL: https://doi.org/10.1177/1461444810365313
[33] T. Kroll, S. Stieglitz, Digital nudging and privacy: Improving decisions about self-disclosure in social networks, Behaviour & Information Technology 40, 1 (2019) 1-19. URL: https://doi.org/10.1080/0144929X.2019.1584644
[34] E. A. Vogels, R. Gelles-Watnick, N. Massarat, Teens, social media and technology 2022, 2022. URL: https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/
[35] G. Schurgin O'Keeffe, K. Clarke-Pearson, Council on Communications and Media, The impact of social media on children, adolescents, and families, Pediatrics 127, 4 (2011) 800-804. URL: https://doi.org/10.1542/peds.2011-0054
[36] J. Gunawan, A. Pradeep, D. Choffnes, W. Hartzog, C. Wilson, A comparative study of dark patterns across web and mobile modalities, in: Proceedings of the ACM on Human-Computer Interaction 5, CSCW2, 2021, pp. 1-29. URL: https://doi.org/10.1145/3479521
[37] K. Krippendorff, Content analysis: An introduction to its methodology, 2nd ed., Vol. 1, Sage, 2004.
[38] D. Kelly, V. L. Rubin, Dark pattern typology: How do social networking sites deter disabling of user accounts? The 12th International Conference on Social Media and Society, 2022. URL: https://easychair.org/publications/preprint/GD6S
[39] D. Kelly, J. Burkell, Documenting privacy dark patterns: How social networking sites influence users' privacy choices, Working paper, 2023. URL: https://ir.lib.uwo.ca/fimspub/376/
[40] Federal Trade Commission, Fortnite video game maker Epic Games to pay more than half a billion dollars over FTC allegations of privacy violations and unwanted charges, 2022. URL: https://www.ftc.gov/news-events/news/press-releases/2022/12/fortnite-video-game-maker-epic-games-pay-more-half-billion-dollars-over-ftc-allegations
[41] J. Merkel, Dark patterns come to light in California data privacy laws, The National Law Review 12, 230 (2021). URL: https://www.natlawreview.com/article/dark-patterns-come-to-light-california-data-privacy-laws
[42] K. Bongard-Blanchy, A. Rossi, S. Rivas, S. Doublet, V. Koenig, G. Lenzini. "I am definitely manipulated, even when I am aware of it. It's ridiculous!" - Dark patterns from the end-

user perspective, in: Designing Interactive Systems Conference 2021, 2021, pp. 763-776. URL: https://doi.org/10.1145/3461778.3462086