2020

Cognitive biases, dark patterns, and the 'privacy paradox'

Ari Ezra Waldman

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Psychology

Cognitive biases, dark patterns, and the 'privacy paradox'

Ari Ezra Waldman[1,2]

Scholars and commentators often argue that individuals do not care about their privacy, and that users routinely trade privacy for convenience. This ignores the cognitive biases and design tactics platforms use to manipulate users into disclosing information. This essay highlights some of those cognitive biases – from hyperbolic discounting to the problem of overchoice – and discusses the ways in which platform design can manipulate disclosure. It then explains how current law allows this manipulative and anti-consumer behavior to continue and proposes a new approach to reign in the phenomenon.

Addresses
[1] Princeton University, Center for Information Technology Policy, Princeton, NJ 08540, United States
[2] New York Law School, 185 West Broadway, New York, NY 10013, United States

Corresponding author: Waldman, Ari Ezra (ari.waldman@nyls.edu)

Current Opinion in Psychology 2020, **31**:105–109

This review comes from a themed issue on **Privacy and disclosure, online and in social interactions**

Edited by **Leslie John**, **Diana Tamir**, and **Michael Slepian**

https://doi.org/10.1016/j.copsyc.2019.08.025

Introduction

Privacy scholars have long argued that most individuals make rational disclosure decisions. Westin [46] used the phrase 'privacy pragmatists' to describe this majority: pragmatists are forward-looking, utility-maximizing, and base their decisions to share on how the information in front of them compares to their privacy preferences. Privacy pragmatists are rational actors in the classical model. This rational choice model of disclosure decision-making informs the dominant approach to privacy governance in the United States: notice-and-consent [33•]. Notice and consent is, at bottom, an informed consent framework that requires websites and other data collectors to be transparent about the ways in which they collect, analyze, and distribute user data, allowing users to rational privacy decisions for themselves.

In 2007, however, Norberg *et al.* [31] found inconsistencies between our stated privacy preferences and our actual disclosure behavior. They called these inconsistencies the 'privacy paradox.' That is, internet users assert strong interest in privacy while simultaneously disclosing substantial personal information for meager rewards. Rationalists try to explain the paradox with nods to the contextual nature of disclosure. Huberman *et al.* [20], for example, suggest that individuals demand a greater price for disclosing stigmatized, less desirable, or embarrassing data, but are quite willing to disclose information they perceive as harmless or innocuous for little to no rewards.

But disclosure choices are not made in vacuums. Sharing is contextual [4], and contingent on both mental capacity and constraints placed on us by designers. In this essay, I offer an alternative explanation for the yawning gap between individuals' disclosure behavior and stated privacy preferences: Any supposed paradox does not reflect users' disinterest in privacy; rather, it reflects users responding in predictable ways to the ways in which platforms leverage design to take advantage of our cognitive limitations.

Moreover, social scientists have debunked many of the assumptions of human decision-making on which the rational-actor disclosure model is based [38,23,6]. Recent literature shows that individuals do not make rational disclosure decisions online [3]. Cognitive biases make rationality difficult and so-called 'dark patterns', or design tricks platforms use to manipulate users into taking actions they might otherwise have not, weaponize the design of built online environments to harm consumers and their privacy [26••,9]. What's more, a rational-actor regime is a largely ineffective way of giving individuals control over the dissemination of their data. It is, in fact, designed to fail [35••,19•]. A rational choice model leverages metacognitive processes that encourage users to give up, to become nihilists about their privacy, and to cede what little control they do have back to technology companies [30••,39••].

Privacy decision-making

The 'privacy paradox' and the correlative rational actor model behind the notice-and-consent regime is based on the myth of rational disclosure. The myth is practically dangerous and systematically unsound.

Practical problems

Today, we have too much data, too many data collection pathways, and too much opacity about those pathways. In that context, notice-and-consent is ill-equipped to inform

users of corporate data use practices. The regime's chief tools—privacy policies—are long [28] and inscrutable; even experts find them misleading [33•]. Cranor [12] estimates that it would take a user an average of 244 hours per year to read the privacy policies of every website she visits, or 54 billion hours per year for every United States consumer to read every privacy policy she encountered [27]. Therefore, even if users were capable of making rational disclosure decisions, privacy policies' inability to adequately convey information means users are unable to do so in practice [34].

There is also an entire industry of data brokers that collects vast amounts of data on individuals in secret and without consent. The Federal Trade Commission [16], the United States' consumer and de facto privacy watchdog, found that one data broker's "databases contain information about 700 million consumers worldwide with over 3000 data segments for nearly every U.S. consumer" and "[m]uch of this activity takes place without consumers' knowledge." As a practical matter, consent cannot operate in a world of passive, secret data collection.

Cognitive biases

The rational choice model is ineffective. It also fails to describe privacy decision-making. Individuals have bounded rationality, which limits their ability to acquire all relevant information and translate it into an evidence-based decision [37]. Recent research has identified myriad cognitive and behavioral barriers to rational privacy and disclosure decision-making [2]. I will discuss the five most pervasive ones here.

The first is what psychologists call anchoring, or the disproportionate reliance on the information first available when we make decisions. For example, Ariely [7] asked experiment participants to provide the last two digits of their Social Security Numbers and then estimate the price of a consumer good. Participants' estimates were close to the two digits they first provided, even though there should be no rational connection between random identity numbers and consumer prices. More recently, Chang *et al.* [10•] showed anchoring effects when he showed that individuals were more likely to disclose personal information after seeing examples of increasingly salacious selfie images. The pictures anchored the participants' perception of what is appropriate to disclose. Anchoring, therefore, can skew individuals' disclosure behavior based on what they see others have shared.

Framing is a second form of bias that technology companies manipulate regularly. Framing concerns the way in which an opportunity is presented to consumers—namely, either as a good thing or a bad thing. Positively framing a privacy policy or a product as more protective of consumer privacy than a competitor's results in a higher propensity to disclose personal information [5]. This is why technology companies explain their data use practices with leading language: "if you don't allow cookies, website functionality will be diminished" or "opting in to data collection will enable new and easier functionality". This has the effect of establishing the positives of data collection while glossing over or ignoring the negatives.

Third, hyperbolic discounting, or the tendency to overweight the immediate consequences of a decision and to underweight those that will occur in the future, makes it difficult for consumers to make rational disclosure decisions. Disclosure often carries with it certain immediate benefits—convenience, access, or social engagement, to name just a few. But the risks of disclosure are usually only felt much later. As such, our tendency to overvalue current rewards while inadequately discounting the cost of future risks makes us more willing to share now. For example, Jentzsch *et al.* [21] found that people preferred barely less expensive movie tickets even though the cheaper ticket required more extensive personal information. Yet, consumer choices changed when tickets were offered at the same price—the privacy protective movie company won more customers. The authors concluded that consumers were heavily discounting the risks associated with disclosing personal information, even far below small differences in price. Other studies have shown that consumers make disclosure decisions without fully appreciating time inconsistent preferences. Wang *et al.* [45] found that users of social networks may gain some immediate pleasure from posting a salacious selfie, but often end up regretting it later and wish they had never posted the picture in the first place. And Acquisti and Fong [1] found that users do not appreciate that posting religious, sexual, or marital status information could result in employment discrimination or ostracism in the future.

A fourth common cognitive barrier to rational disclosure decision-making is overchoice. Overchoice is the problem of having too many choices, which can overwhelm and paralyze consumers [36]. A form of overchoice affects internet and mobile app users trying to navigate their privacy. When making disclosure choices, Hartzog [19•] has shown that users are overwhelmed not with the choices they have, but with the number of choices they have to make. Most apps, websites, and platforms require us to make yes/no choices with respect to cookies, location tracking, and behavioral targeting, among others. Indeed, as Olmstead and Atkinson [32] have shown, mobile apps often ask users for more than 200 permissions, with the average app asking for about five. It is hard to see how ordinary users, without any particular technological expertise, can navigate it all.

Finally, metacognitive processes in decision-making impair individuals' ability to make choices that accurately reflect their preferences. As Mourey [30••] has shown,

when some individuals are confronted with difficult choices, they perceive difficulty as a signal of importance, encouraging them to deploy strong cognitive processes to a meaningful or weighty decision. But when individuals perceive difficulty as a cue of impossibility, they tend to give up, ceding their power and autonomy to choose to the default. When applied to the content of privacy navigation and online disclosure, it stands to reason that the more users feel it is difficult to maintain their privacy online, as many do, the more likely many of them are to nihilistically decline to manage their disclosure.

Design and 'dark patterns'

Even if none of these cognitive hurdles to rational disclosure decision-making existed, internet users would still face the limitations imposed on them by design. By 'design', I am following Hartzog's [19•] broad definition, which embraces the "processes that create consumer technologies and the results of their creative processes instantiated in hardware and software." Science and Technology Studies has long recognized that the design of built environments constrains human behavior [47]. The same is true online [13,19•], and even more so when millions, if not billions, of people with potentially different preferences are using the same service. As Hartzog [18•] has noted, "[t]he realities of technology at scale mean that the services we use must necessarily be built in a way that constraints our choices." Users can only click on the buttons or select the options presented to them; we can only opt-out of the options from which a website allows us to opt-out. Harris [17] likens the power of design to manipulate user choices to a magician's misdirection: "we ignore how . . . [our] choices are manipulated upstream by menus we didn't choose in the first place. . . . This is exactly what magicians do. They give people the illusion of free choice while architecting the menu so that they win, no matter what you choose. . . . By shaping the menus we pick from, technology hijacks the way we perceive our choices and replaces them with new ones." We see this throughout the digital ecosystem. Facebook tells us when our friends have 'liked' a page, encouraging us to do the same; dark patterns trigger our preference for shiny buttons over grey ones; platforms nudge us to buy products others have bought before us; and apps gamify sharing by encouraging us to continue a 'streak' with our friends. The list goes on.

At a minimum, the power of design means that our choices do not always reflect our real personal preferences. At worst, online platforms manipulate us into keeping the data flowing, fueling an information-hungry business model. That manipulation is often the result of so-called 'dark patterns' in platform design. Mathur *et al.* [26••] define dark patterns as "interface design choices that benefit an online service by coercing, steering, or deceiving users into making decisions that, if fully informed and capable of selecting alternatives, they might not make." And they are increasingly common. Designers use dark patterns to hide, deceive, and goad users into disclosure. They confuse users by asking questions in ways nonexperts cannot understand, they obfuscate by hiding interface elements that could help users protect their privacy, they require registration and associated disclosures in order to access functionality, and hide malicious behavior in the abyss of legalese privacy policies. Dark patterns also make disclosure 'irresistible' by connecting information sharing to in-app benefits. In these and other ways, designers intentionally make it difficult for users to effectuate their privacy preferences.

Online platforms are also socially constructed, designed by real people with biases both implicit and explicit [14•,24•]. Although there are, in fact, several social groups involved in the creation and design of websites, apps, and other data collection platforms—from executives and lawyers to marketers and users—the mostly white male engineers and technologists on the ground play a critical role in channeling ideas into design [24•,40••]. As such, theirs is the vision most likely integrated into code [40••]. The biases of these designers not only damage our privacy [40••,44••], but may also constrain one group of consumers more than others, whether it is Asians who cannot use a 'smart' camera because the camera's artificial intelligence thinks all Asians have their eyes closed when posing for pictures [25•] or an app in which location tracking cannot be turned off regardless of the potential for harm to women and stalking victims [11].

Trust and privacy

The design of built online environments and our cognitive biases make rational disclosure decision-making difficult. Therefore, the disconnect between users' stated preferences and disclosure behavior has been misunderstood. The evidence suggests that individuals care about their privacy. They are, however, dissuaded from acting effectively on those preferences by cognitive limitations leveraged by the digital platforms themselves.

Nor is disclosure arbitrary. As we have seen, it can be manipulated in specific, multidirectional ways. The contextual nature of disclosure and privacy [4] means that we can explain some sharing of personal data through principles of trust and corporate attempts to manipulate that trust.

Trust, a resource of social capital between or among two or more parties concerning the expectation that others will behave according to accepted norms, is a powerful predictor of a willingness to share personal information online [43••]. Trust is the 'favorable expectation regarding other people's actions and intentions,' or the belief that others will behave in a predictable manner [29]. For example, when an individual speaks with relative strangers in a support group like Alcoholics Anonymous, she trusts that they will not divulge her secrets. Trust,

therefore, includes a willingness to accept some risk and vulnerability toward others and steps in to grease the wheels of social activity.

We share our information with others in contexts of trust [22,43••]. When we learn that others have shared personal information, we share, as well [42••]. When we know our close friends participate in an online social network, we are more likely to participate, as well. When someone shares a stigmatizing social identity, we are more likely to share personal information with them [42••]. We develop privacy management techniques based on indicia of trust.

This means that trust is a target of design and manipulation. Dark patterns may hide, confuse, and obfuscate, but they do so without users knowing [26••]. The only thing users see are nudges to behave in certain ways. Dark patterns can hide disclosure dangers while simultaneously highlighting the powerful social cues to share. Facebook, for example, introduces and follows every News Feed post with information about which friends and how many of them have 'liked' or shared the content. Websites cue trust through professional design while hiding their invasive data collection practices in inscrutable privacy policies [43••].

As such, privacy law should reflect the fact that we share data with others in contexts of trust, that rational choice inadequately describes our disclosure decision-making, and that users must be protected from unfair and manipulated disclosure. We entrust our information to digital platforms much like we entrust our financial information to estate planners or our medical information to doctors or our legal situation with lawyers [8]. The law requires those parties to act in a trustworthy manner. They are, in fact, trustees, or fiduciaries, of our data: we are vulnerable to them, we depend on them, and they hold themselves out as experts and trustworthy [8]. As such, they should also be responsible for acting in our benefit or, at least, not acting in ways that benefit them at our expense. In particular, this notion of 'information fiduciaries' requires three things: duties of care, duties of confidentiality, and duties of loyalty. Duties of care would require technology companies to take reasonable steps to secure our data, with 'reasonable' defined by centuries of common law. Duties of confidentiality would require those who collect our data to collect only so much as necessary to achieve a particular purpose and limit the use of collected data to specific purposes to which users consent. And duties of loyalty would ensure that companies do not profit by harming us. To put it another way, treating data collectors as fiduciaries of our data would ban their use of dark patterns to manipulate and coerce disclosure of our personal information. Recent legislation proposed by U.S. Senator Brian Schatz (D-Hawaii) proposes to integrate these fiduciary principles of care, confidentiality, and loyalty into a comprehensive United States federal privacy law [15,41]. It deserves consideration particularly because it reflects the latest research on how disclosure can be manipulated by cognitive biases and coercive design.

Conclusion

Many internet users care about their privacy. And yet, technology companies have made design choices that make it difficult for users to realize those preferences. Online environments are built not only to constrain users, but to coerce disclosure and trigger cognitive biases that encourage us to give up and cede control over our privacy. Online platforms can behave in these predatory ways because the law, based on the myth of rational disclosure, allows them. A different approach, one based on the connection between trust and sharing, would hold online platforms to a higher standard of loyalty, confidentiality, and care.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Acquisti A, Fong CM: *An Experiment in Hiring Discrimination Via Online Social Networks*. . Available at SSRN 2031979 2014:1-81.
2. Acquisti A, Brandimarte L, Loewenstein G: **Privacy and human behavior in the age of information**. *Science* 2015:509-514.
3. Acquisti A, Grossklags J: **What can behavioral economics teach us about privacy**. In *Digital Privacy: Theory, Technologies, and Practices*. Edited by Acquisti A, Gritzalis S, Lambrinoudakis C, di Vimercati S. 2007:363-377.
4. Acquisti A, John L, Loewenstein G: **The impact of relative standards on the propensity to disclose**. *J Mark Res* 2012, **49**:160-174.
5. Adjerid I, Acquisti A, Brandimarte L, Loewenstein G: **Sleights of privacy: framing, disclosures, and the limits of transparency**. *Proceedings of the Symposium on Usable Privacy and Security (SOUPS'13) ACM* 2013:1-11.
6. Ariely D: *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. HarperCollins Publishers; 2008.
7. Ariely D, Loewenstein G, Prelec D: **"Coherent arbitrariness": stable demand curves without stable preferences**. *Q J Econ* 2003, **118**:73-106.
8. Balkin J: **Information fiduciaries and the first amendment**. *Univ Calif Davis Law Rev* 2016, **49**:1183-1234.
9. Calo R: **Digital market manipulation**. *George Wash Law Rev* 2014, **82**:995-1051.
10. • Chang D, Krupka EL, Adar E, Acquisti A: **Engineering information disclosure: norm shaping designs**. *Proceedings of the Conference on Human Factors in Computing Systems (CHI'16), ACM* 2016:587-597.

Study showing the ways in which design patterns can affect disclosure behavior. In particular, showing that individuals are more willing to share personal information as they are shown more revealing images, and notably more likely to disclose if being shown revealing images changes their perception of what is appropriate to share.

11. Citron DK: **Spying, Inc**. *Wash Lee Law Rev* 2015, **72**:1243-1282.

12. Cranor LF: **Necessary but not sufficient: standardized methods for privacy notice an choice**. *J Telecommun High Technol Law* 2012, **10**:273-307.

13. Cohen JE: **DRM and privacy**. *Berkeley Technol Law J* 2003, **18**:575-617.

14. • Crawford K, Whittaker M, Elish MC, Barocas S, Plasek A, Ferryman K: *The AI Now Report: The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term*. 2016:1-25.
Demonstrating, through field research, the human biases translated into the design of new technologies. In particular, noting the biases in data used as inputs for artificial intelligence and the biases of the male-dominated design workforce.

15. Data Care Act, https://www.schatz.senate.gov/imo/media/doc/Data%20Care%20Act%20of%202018.pdf.

16. Federal Trade Commission: *Data Brokers: A Call for Transparency and Accountability*. 2014.

17. Harris T: **How technology is hijacking your mind—from a magician and Google design ethicist**. *Medium*. 2016. November 2018.

18. • Hartzog W: **The case against idealizing control**. *Eur Data Prot Law Rev* 2018, **4**:423-432.
Arguing, that traditional approaches to data protection—giving users 'control' over their information—fail for several reasons, including the overuse of consents, cognitive hurdles, and design strategies.

19. • Hartzog W: *Privacy's Blueprint: The Battle to Control the Design of New Technologies*. Harvard University Press; 2018.
Calling for a design agenda at the regulatory level to protect individuals from data privacy harms and documenting the many ways in which the design of online environments nudges and manipulates user behavior.

20. Huberman BA, Adar E, Fine LR: **Valuating privacy**. *IEEE Secur Priv* 2005, **3**:22-25.

21. Jentzsch N, Preibusch S, Harasser A: *Study on Monetising Privacy: An Economic Model for Pricing Personal Information*. European Union Agency for Network and Inf. Sec. (ENISA); 2012.

22. Jourard SM: *The Transparent Self*. Van Nostrand Reinhold Inc.; 1971.

23. Kahneman D: *Thinking, Fast and Slow*. Farrar, Straus and Giroux; 2013.

24. • Katyal SK: **Private accountability in the age of artificial intelligence**. *Univ Calif Los Angel Law Rev* 2019, **66**:54-141.
Identifying the many ways in which artificial intelligence tools reflect and entrench racial and sexual biases and calling for soft law proposals, from impact assessments to ethics protocols, to ensure AI can be used without bias problems.

25. • Levendowski A: **How copyright law can fix artificial intelligence's implicit bias problem**. *Wash Law Rev* 2018, **2**:579-630.
Arguing that one of the reasons for bias in machine learning is the limited data available to train machine learning algorithms because so much is protected by copyright, and calling for a fair use exception for training data.

26. •• Mathur A, Acar G, Friedman M, Lucherini E, Mayer J, Chetty M, Narayanan A: **Dark patterns at scale: findings from a crawl of 11K shopping websites**. *ACM Conf. Comp.-Supported Cooperative Work* 2019.
Presenting automated research techniques to help experts identify dark patterns on a large scale and leveraging those techniques on thousands of websites and platforms to identify 1818 dark patterns and 183 websites that engage in deceptive practices.

27. McDonald AM, Cranor LF: **The cost of reading privacy policies**. *I/S: J Law Policy Inf Soc* 2008:543-568.

28. Milne GR, Culnan MJ, Greene H: **A longitudinal assessment of online privacy notice readability**. *J Public Policy Mark* 2006, **25**:238-249.

29. Möllering G: **The nature of trust: from Georg Simmel to a theory of expectation, interpretation and suspension**. *Sociology* 2001, **35**:403-420.

30. •• Mourey JA: *Prime Time: A Mental Resource Opportunity Cost Approach to When and How Primes Influence Choice*. 2019. [In press].
Discussing several studies showing that those who interpret difficult decisions as cuing importance approach those decisions differently than those who interpret difficulty as cuing impossibility.

31. Norberg PA, Horne DR, Horne DA: **The privacy paradox: personal information disclosure intentions versus behaviors**. *J Consum Aff* 2007, **41**:100-126.

32. Olmstead K, Atkinson M: *Apps Permissions in the Google Play Store*. Pew Research Center; 2015.

33. • Reidenberg J, Breaux T, Cranor LF, French B: **Disagreeable privacy policies: mismatches between meaning and users' understanding**. *Berkeley Technol Law J* 2015, **30**:39-68.
Analyzing experiments showing that ordinary users do not understand the text of privacy policies and that even privacy experts either do not understand or, at least, disagree as to the meaning of privacy policy terms.

34. Reidenberg J, Russell NC, Callen A, Qasir S, Norton T: **Privacy harms and the effectiveness of the notice and choice framework**. *I/S: J Law Policy Inf Soc* 2015:485-524.

35. •• Richards N, Hartzog W: *The Pathologies of Consent*. 2019. [In press].
Discussing the ways in which technology companies overuse the tools of consent to manipulate users into disclosing personal information.

36. Scheibehenne B, Greifeneder R, Todd PM: **Can there ever be too many options? A meta-analytic review of choice overload**. *J Consum Res* 2010, **37**:409-425.

37. Simon HA: *Models of Bounded Rationality*. MIT Press; 1982.

38. Thaler R: *Misbehaving: The Making of Behavioral Economics*. W. W. Norton & Co.; 2016.

39. •• Turow J, Hennessy M, Draper N, Akanbi O, Virgilio D: *Divided We Feel: Partisan Politics Drive Americans' Emotions Regarding Surveillance of Low-Income Populations*. Annenberg School for Communication, University of Pennsylvania; 2018:1-30.
Finding, among other things, that most Americans feel resigned to their inability to protect their privacy online rather than not caring about their privacy.

40. •• Waldman A: **Designing without privacy**. *Houst Law Rev* 2018, **55**:659-727.
Discussing qualitative interviews with engineers in the high technology sector showing their narrow view of privacy and their biased approach to design are the approaches most likely to be integrated into design, and recommending micro and macro social and legal approaches to better design for privacy.

41. Waldman A: **Duties of loyalty and care: Sen. Brian Schatz offers a new approach to data privacy**. *Medium*. 2018.

42. •• Waldman A: **Law, privacy, and online dating: "revenge porn" in gay online communities**. *Law Soc Inq* 2019.
In accordance with a survey of gay and bisexual male users of geosocial dating apps, showing a high rate of nonconsensual pornography on the platforms and identifying trust-based privacy navigation techniques among users.

43. •• Waldman A: *Privacy As Trust: Information Privacy for an Information Age*. Cambridge University Press; 2018.
Arguing that because we share when we trust, privacy law should be oriented around protecting as private information shared in contexts of trust.

44. •• Waldman A: **Privacy law's false promise**. *Wash Univ Law Rev* 2019. [In press].
Showing that privacy technology vendors, and the engineers designing their tools, are exerting increasing control over what privacy laws mean in practice, threatening the protections those laws offer.

45. Wang Y, Norcie G, Komanduri S, Acquisti A, Leon PG, Cranor LF: **I regretted the minute I pressed share: a qualitative study of regrets on Facebook**. In *Proceedings of the Symposium on Usable Privacy and Security ACM*. 2011:1-16.

46. Westin A: *Privacy and Freedom*. Bodley Head; 1970.

47. Woolgar S: **Configuring the user: the case of usability trials**. *Sociol Rev* 1990, **38**:58-99.