REGULATING PRIVACY DARK PATTERNS IN PRACTICE—DRAWING INSPIRATION FROM CALIFORNIA PRIVACY RIGHTS ACT

Jennifer King[*] & Adriana Stephan[**]

CITE AS: 5 GEO. L. TECH. REV. 250 (2021)

ABSTRACT

Scholars define dark patterns as user interface design choices that benefit an online service by coercing, manipulating, or deceiving users into making unintended and potentially harmful decisions. Although dark patterns are a subject of interest of policymakers around the globe, to date, little legislation has been enacted to regulate them. This Article analyzes the definition of dark patterns introduced by the California Privacy Rights Act (CPRA), the first legislation explicitly regulating dark patterns in the United States. We discuss the factors that make defining and regulating privacy-focused dark patterns challenging, review current regulatory approaches, consider the challenges of measuring and evaluating dark patterns, and provide recommendations for policymakers. We argue that California's model offers the opportunity for the state to take a leadership role in regulating dark patterns generally and privacy dark patterns specifically, and that the CPRA's definition of dark patterns, which relies on outcomes and avoids targeting issues of designer intent, presents a potential model for others to follow. Because many dark patterns

[*] Jennifer King, Ph.D., Stanford Institute for Human-Centered AI, kingjen@stanford.edu.

[**] Adriana Stephan, Independent Researcher, adebolt@stanford.edu. The authors wish to thank Alessandro Acquisti, Katelyn Ringrose, Richmond Wong, and Lauren Willis for their comments on drafts on this Article, and Chris Hoofnagle, Arunesh Mathur, and Mihir Kshirsagar for sharing their time with us to discuss their work. Dr. King wishes to acknowledge the support of her husband and the network of caregivers that made completing this Article possible in the midst of a global pandemic while sharing the homeschooling of two children. Dr. King's work was supported in part by unrestricted funding from Accenture and Facebook to the Center for Internet and Society (CIS) at Stanford Law School. CIS adheres to a strict non-interference policy for research funding, and neither funder was provided pre-publication drafts of this Article, nor any opportunity to shape this research.

do not explicitly rely upon deception, we argue that regulating dark patterns necessitates expanding the Federal Trade Commission's regulatory authority to include coercion and manipulation, as well as an embrace of performance-based standards, a change that state authorities may also wish to adopt. We conclude by suggesting that in regulating dark patterns, policymakers may also have the opportunity to introduce a paradigm shift in how to measure, evaluate, and enforce consumer privacy protections by drawing on human-centered design as a model.

TABLE OF CONTENTS

I. INTRODUCTION

2020 was a notable year for privacy legislation in the United States. The California Consumer Privacy Act (CCPA) took effect on January 1, 2020, and provided California residents one of the first broad consumer privacy laws with a set of new rights and protections for their personal information. The CCPA was followed in November by the passage of the California Privacy Rights Act (CPRA), a ballot initiative submitted directly to the voters that

passed with a 56% majority that will take effect on January 1, 2023.[1] Taken together, the two laws provide the most comprehensive legislative protections for personal data in the United States. Given California's size and stature, these two laws will likely have both a catalyzing effect on state and federal responses in the near-term, as well as a spillover effect with companies that opt to extend their provisions to customers outside of California's borders. Furthermore, the CPRA authorized the creation of a new California Privacy Protection Agency to "administer, implement, and enforce [the new laws] through administrative actions."[2] The Agency will be the first of its kind in the United States.

Both the CCPA and the CPRA are also notable for their attempt to regulate a phenomenon of growing interest to policymakers—dark patterns.[3] The CPRA 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."[4] While there have been a few attempts in the U.S. Congress to regulate dark patterns, California accomplished it first, both with the passage of the CPRA and with modifications in March of 2021 to the CCPA that expressly address attempts to subvert or impair Californians' ability to opt-out of sales of their personal information.[5] A closer look at these new regulations illustrates some of the problems of concern to policymakers, as well as the challenges of regulating this phenomenon.

In this Article, we discuss the key factors that policymakers should consider when contemplating how to prohibit dark patterns with privacy impacts, though aspects of this discussion could apply to dark patterns in other contexts. First, we start with the difficult task of defining dark patterns, relying on recent scholarship that seeks to utilize normative definitions rather than

---

[1] Matt Dumiak, *California Privacy Rights Act Has Passed*, COMPLIANCE POINT (Dec. 10, 2020), https://www.jdsupra.com/legalnews/the-california-privacy-rights-act-has-23645/ [https://perma.cc/WYG9-Z4H5].

[2] California Privacy Rights Act, CAL. CIV. CODE § 1798.199.40(a) (West 2020) (effective Jan. 1, 2023).

[3] The definition of dark patterns that we utilize for our analysis throughout this Article is attributed to: Arunesh Mathur, Jonathan Mayer & Mihir Kshirsagar, *What Makes a Dark Pattern... Dark?: Design Attributes, Normative Considerations, and Measurement Methods*, PROC. 2021 CHI CONF. ON HUM. FACTORS COMPUTING SYS. (2021). We acknowledge that some stakeholders are calling for the term dark patterns to be abandoned in favor of language that doesn't rely on contrasts between light and dark due to concerns about possible racialized implications of these concepts, suggesting terms such as manipulative, deceptive, or coercive design instead. Because our argument in this paper relies upon unpacking these specific terms, we could not single out one of these descriptors given each carries different implications, but we are ultimately supportive of whatever term the community of stakeholders adopts.

[4] California Privacy Rights Act § 1798.140(l).

[5] California Consumer Privacy Act, CAL. CODE REGS. tit. 11, § 999.315(h) (2021).

descriptive ones. Using the CPRA as a case study, we then evaluate how the CPRA could provide a model for both national and state policymakers considering regulating privacy dark patterns. We suggest that the CPRA's focus on outcomes rather than intent is well suited for using performance-based standards to evaluate manipulative privacy dark patterns. The CPRA also includes conceptual language to address dark patterns that may not fit under existing definitions of deception. We conclude with recommendations for policymakers that urge defining dark patterns in terms of their effects on users, as well as a paradigm shift away from placing rational choice theory at the center of dark pattern regulation.

II. DARK PATTERNS: DEFINITIONS AND CHALLENGES

Dark patterns are a focus of the "design community" writ large—the visual designers, interaction designers, interface designers, human-computer interaction scholars, computer scientists, and information scientists for whom the designing or evaluation of computer interfaces is their expertise. More recently, policymakers and advocates have joined this community of stakeholders because of the growing awareness of the power that interface design has over our online behavior and decision-making. In this Part we review several definitions of dark patterns, identify the harms associated with their use, and take note of complexities that make regulating dark patterns especially challenging.

A. Definitions

In the visual design universe, design patterns are "reusable/recurring components which designers use to solve common problems in user interface design,"[6] such as navigation menus for webpages or mobile devices. The term "dark patterns" was popularly coined in 2010 by designer Harry Brignull, who described them as "tricks used in websites and apps that make you do things that you didn't mean to, like buying or signing up for something."[7] "Normal" design patterns typically become de facto best practices over time based on results from user research and practitioner experience, but with a general expectation that they represent a *benefit* to the user, in terms of efficiency, minimal cognitive burden, or simplicity. Dark patterns are similar sets of recurring design components, but that are instead optimized for the benefit of

---

[6] *User Interface (UI) Design Patterns*, INTERACTION DESIGN FOUND., https://www.interaction-design.org/literature/topics/ui-design-patterns [https://perma.cc/3PJ3-GH6H] (last visited Apr. 10, 2021).

[7] Harry Brignull, *What are Dark Patterns?*, DARK PATTERNS, https://darkpatterns.org/ [https://perma.cc/J2FF-M2Q9] (last visited Apr. 10, 2021).

the designer,[8] usually at the direct expense of the user. An emerging literature has grown around cataloging deceptive retail practices in e-commerce and documenting dark patterns across additional contexts, such as privacy and gaming.[9] In this Article, we are concerned primarily with dark patterns with privacy effects,[10] namely those that coerce consumers into disclosing more personal information than they otherwise might by coercing or manipulating consent, obscuring opt-out options, pre-selecting privacy-invasive defaults,[11] or "confirmshaming"[12] consumers into making disclosure choices they may not otherwise make.

[8] While "designers" (e.g., visual designers or engineers) may be the individuals who actually implement dark patterns in practice, the term is a proxy for whomever or whatever is benefitting from their use. For a discussion of the user's best interest *see* Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt & Austin L. Toombs, *The Dark (Patterns) Side of UX Design*, 2018 PROC. 2018 CHI CONF. ON HUM. FACTORS COMPUTING SYS., 9–10.

[9] *See, e.g.*, Mathur et al., *supra* note 3.

[10] Christoph Bösch, Benjamin Erb, Frank Kargl, Henning Kopp & Stefan Pfattheicher, *Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns*, 2016 PROC. ON PRIVACY ENHANCING TECH. 237, 248–49.

[11] Norwegian Consumer Council, *Deceived by Design*, FORBRUKER RADET 13–18 (June 27, 2018), https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf [https://perma.cc/NRA6-TM8L]; Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger & Lalana Kagal, *Dark Patterns after GDPR: Scraping Consent Pop-ups and Demonstrating their Influence*, 2020 PROC. 2020 CHI CONF ON HUM. FACTORS COMPUTING SYS., 3.

[12] "Confirmshaming" refers to the practice of "guilting the user into opting in to something." *Types of Dark Pattern: Confirmshaming*, DARK PATTERNS, https://www.darkpatterns.org/types-of-dark-pattern/confirmshaming [https://perma.cc/E2BJ-FTT3] (last visited Apr. 10, 2021).

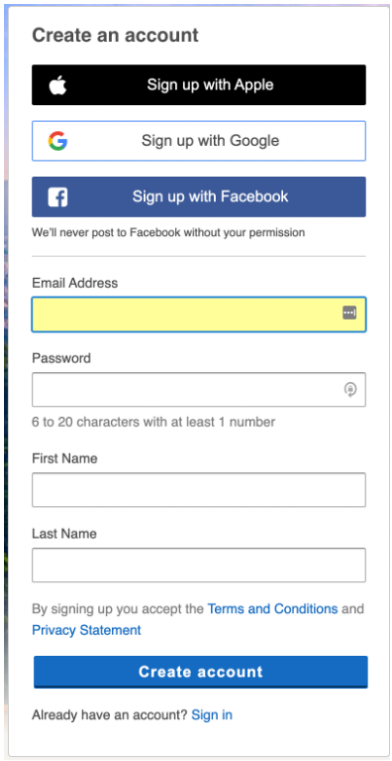

Figure 1: Example of bundled consent.

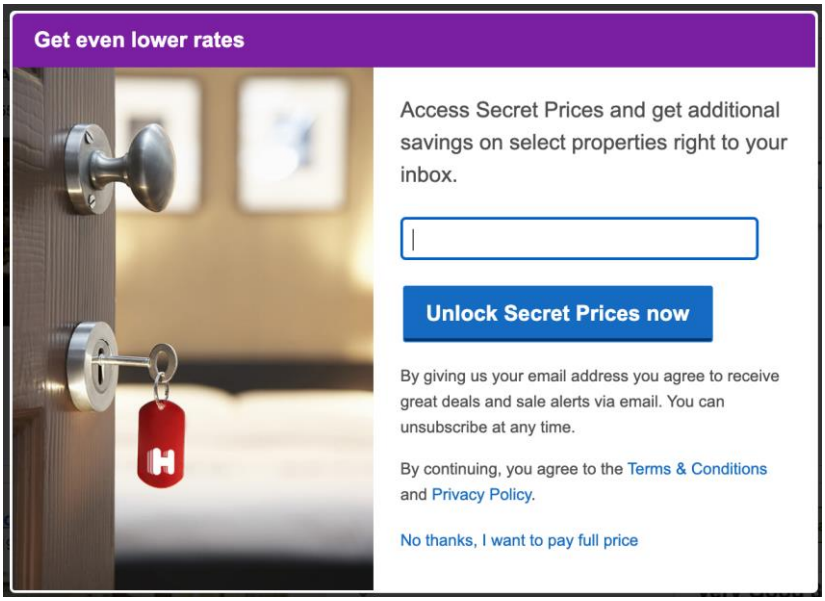

Figure 2: Example of confirmshaming from Hotels.com

The expanding use of dark patterns coincides with the rise of *persuasive technology*,[13] the practice of using digital devices to change people's attitudes and behavior, as well as with the popularization of *choice architectures*, the organization of "the context in which people make decisions."[14] In behavioral economics, choice architectures can be manipulated using *nudges*, changing the presentation of choices in ways that encourage people to pick a particular option,[15] to make consumer behavior more predictable. One example of a nudge is the practice of automatically enrolling employees in pension saving plans so that contributions are deducted from their paychecks unless they formally request to opt out, rather than relying on them to opt into such programs. As the public increasingly understands that digital interfaces are designed with specific outcomes in mind, some have begun calling out dark patterns as examples of anti-nudges, or 'sludge.'[16] In the e-commerce context, this includes design patterns that add unwanted items to shopping carts, push consumers toward more expensive options, or make it difficult for consumers to cancel purchases or subscriptions.[17]

Recently, scholarship on dark patterns has evolved from identifying and describing them toward coalescing around systematic definitions. Princeton researchers Arunesh Mathur, Jonathan Mayer, and Mihir Kshirsagar surveyed the published literature and grouped the many definitions and descriptions they observed based on how they affect choice architectures: whether the pattern "modifies the decision space" (e.g., eliminates choices or poses asymmetric burdens) or "manipulates the flow of information" to the user (e.g., uses deception or hides or obscures information).[18] From this analysis, the authors identified four normative lenses—individual welfare, collective welfare, regulatory objectives, and individual autonomy—to

---

[13] *See* B.J. FOGG, PERSUASIVE TECHNOLOGY: USING COMPUTERS TO CHANGE WHAT WE THINK AND DO 1 (2003).

[14] Richard H. Thaler, Cass R. Sunstein & John P. Balz, *Choice Architecture*, *in* THE BEHAVIORAL FOUNDATIONS OF PUBLIC POLICY 428, 428 (Eldar Shafir ed., 2013).

[15] *See* RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 6 (Penguin Books 2009) (defining a nudge as "any aspect of the choice architecture that alters people's behavior in a predictable way").

[16] *See generally* Cass R. Sunstein, *Sludge Audits*, 5 BEHAV. PUB. POL'Y 1, 1–20 (2020), https://www.cambridge.org/core/journals/behavioural-public-policy/article/sludgeaudits/12A7E338984CE8807CC1E078EC4F13A7 [https://perma.cc/WB8E-9SFD] (referring to nudges that "cost time or money" or generally "make life difficult to navigate" as "sludge").

[17] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, & Austin L. Toombs, *The Dark (Patterns) Side of UX Design*, 2018 PROC. 2018 CHI CONF. ON HUM. FACTORS COMPUTING SYS., 1–14.

[18] Mathur et al., *supra* note 3, at 9.

"attempt to explain *why* dark patterns should concern us,"[19] as well as to develop a common language to discuss them. Their scholarship highlights the fact that there is not a singular, one-size-fits-all definition of dark patterns, but rather that one's definition should be informed by the issue one seeks to solve. This argument is not merely academic; as the authors discuss, without any standards clarifying what constitutes a dark pattern (or even a borderline "gray" pattern), policymakers will continue to struggle with how best to legislate them.[20] For policymakers, the authors' framing may prove especially useful, as it emphasizes the need to be clear about both the problem and the specific context one wishes to address when regulating dark patterns, while at the same time suggesting that overly broad definitions may end up being meaningless or unenforceable in practice.

Another important aspect of defining dark patterns in the policymaking context is specifying the differences between deception and manipulation. This concern can be traced directly to the Federal Trade Commission's (FTC) authority in the United States to investigate deceptive and unfair trade practices, which we discuss in more detail in the Performance-Based Standards, Not Intent Section below. In short, while some argue that the FTC has authority sufficient to regulate dark patterns under Section 5 of the FTC Act,[21]others believe that the agency must expand its mandate to expressly include manipulative or "abusive" practices (similar to the Consumer Financial Protection Bureau's authority to prohibit "abusive" practices).[22] Daniel Susser, Beate Roessler, and Helen Nissenbaum define manipulation as hidden influence, "intentionally and covertly influencing [one's] decision-making, by targeting and exploiting their decision-making vulnerabilities."[23] In their view, "deception is a special case of manipulation[,]" specifically, the "plant[ing of] false beliefs."[24] Given that many dark patterns do not deceive, but instead coerce, confuse, subvert autonomy, and otherwise steer users towards behaviors that directly benefit a company or a designer at the expense

[19] *Id.* at 13.
[20] *Id.* at 2.
[21] *See* Federal Trade Commission Act, 15 U.S.C. § 45(a)(1) (2018).
[22] *See* Chris Jay Hoofnagle, Woodrow Hartzog & Daniel J. Solove, *The FTC can rise to the privacy challenge, but not without help from Congress*, BROOKINGS: TECHTANK BLOG (Aug. 8, 2019), https://www.brookings.edu/blog/techtank/2019/08/08/the-ftc-can-rise-to-the-privacy-challenge-but-not-without-help-from-congress/ [https://perma.cc/ZCF8-MH2M]; *see also* WOODROW HARTZOG, PRIVACY'S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES 143–47 (2018) (explaining the difference between deceptive design and abusive design: while deceptive design "misrepresents reality and subverts expectations," abusive design "unreasonably exploits cognitive vulnerabilities").
[23] Daniel Susser, Beate Roessler & Helen Nissenbaum, *Technology, Autonomy, and Manipulation*, 8 INTERNET POL'Y REV. 1, 4 (2019).
[24] *Id.*

of the individual, we believe the FTC's current deception authority may be insufficient in cases where deception is not the core issue.

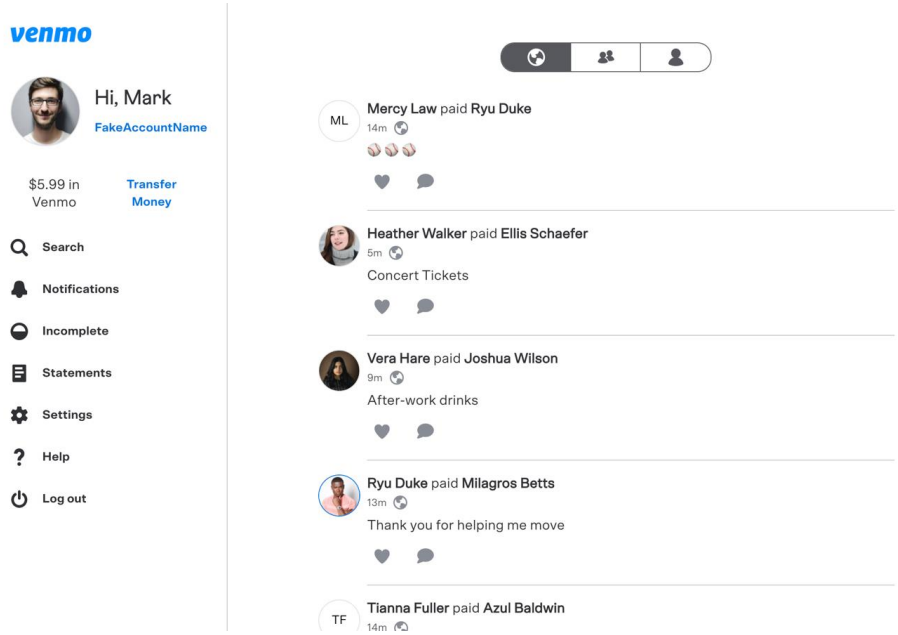

Figure 3: An example of a privacy dark pattern that exposes extraneous information by default. Venmo's user page (as of May 2021) showed all other Venmo users' transactions unless individuals opted-out to this default setting. All images, names, and transaction dates have been altered from their original to protect privacy.

B. Harms and Contexts

A core concern with dark patterns is that they undermine individual autonomy through coercion and manipulation. They mock the open market by snagging consumers into purchases or disclosures of personal information they would otherwise not elect to make. As Maya MacGuineas, president of the Committee for a Responsible Federal Budget, wrote in *The Atlantic* in April 2020,

> [I]n a well-functioning market, consumers have the freedom to act in their own self-interest and to maximize their own well-being. Prices are transparent, and people have a basic level of trust that exchanges of goods, services, and money benefit all parties. Consumers, it is

> assumed, are discerning and rational in the face of the market's blandishments—an assumption that is crucial to the whole system's ability to produce social good . . . [b]ut the new powers in the digital age have built their business models on strategies—enabled and turbocharged by self-improving algorithms—that actively undermine the principles that make capitalism a good deal for most people.[25]

There is ample empirical evidence from both human-computer interaction (HCI) research and behavioral economics that consumers are not "discerning and rational," but are instead subject to cognitive biases and heuristics in their decision-making.[26] With respect to privacy specifically, researchers in communications have identified at least twelve privacy heuristics that predict information disclosure,[27] including the "paradox of control" first identified by Brandimarte, Acquisti, and Loewenstein, where they discovered that one's perception of greater control over personal information led to increased rates of personal disclosure.[28] It is these cognitive biases and heuristics, in concert with the control that designers exert over both the decision space and information flows, that dark patterns exploit.

Dark patterns can appear in nearly any online context. However, based on proposed legislation and government reports, the current areas of greatest concern to policymakers are shopping, gaming, information collection (privacy), and consent. In the most extreme cases, policymakers are concerned that dark patterns foster addiction, whether it be to online gaming or to social media feeds. For example, in 2018, U.S. Senators Mark R. Warner (D-VA) and Deb Fischer (R-NE) introduced the DETOUR Act (reintroduced as part

---

[25] Maya MacGuineas, *Capitalism's Addiction Problem*, ATLANTIC (Apr. 2020), https://www.theatlantic.com/magazine/archive/2020/04/capitalisms-addiction-problem/606769/ [https://perma.cc/7P9Z-UP4P].

[26] *See generally* DAN ARIELY, PREDICTABLY IRRATIONAL: THE HIDDEN FORCES THAT SHAPE OUR DECISIONS (2008) (providing an overview of cognitive biases); DANIEL KAHNEMAN, THINKING, FAST AND SLOW (paperback ed. 2013) (providing an in-depth explanation of cognitive biases and heuristics); Alessandro Acquisti, Laura Brandimarte & George Loewenstein, *Privacy and Human Behavior in the Age of Information*, 347 SCIENCE 509 (2015) (illustrating an application of cognitive biases and heuristics in behavioral economics to privacy).

[27] *See* S. Shyam Sundar, Jinyoung Kim, Mary Beth Rosson & Maria D. Molina, *Online Privacy Heuristics that Predict Information Disclosure*, PROC. 2020 CHI CONF. ON HUM. FACTORS COMPUTING SYS. 1, 3 (2020).

[28] Laura Brandimarte, Alessandro Acquisti & George Loewenstein, *Misplaced Confidences: Privacy and the Control Paradox*, 4 SOC. PSYCHOL. PERSONALITY SCI. 340, 340–47 (2012).

of the SAFE DATA Act in 2020). In addition to taking aim at privacy dark patterns, both proposals focused on personal disclosure as well as specifically aiming to curb user interfaces designed to foster "compulsive usage" by children under age thirteen.[29] There are, unsurprisingly, unique concerns about children and other vulnerable Internet users who may be especially susceptible to dark patterns or do not have the online experience or savvy to identify them. Jamie Luguri and Lior Strahilevitz found in their experimental research with dark patterns that education levels were predictive of susceptibility to dark patterns, with participants who reported lower education levels being more vulnerable to their effects.[30]

C. Practical Challenges

Legislating to reduce the incidence of dark patterns is not as simple as identifying the worst offenders and prohibiting their use. First, attempting to prohibit specific forms of interaction design (e.g., autoscrolling) may be, ironically, both too narrow and too broad an approach. Singling out specific design patterns will motivate designers to adapt their designs to thwart prohibitions while also potentially limiting legitimate applications. Further, zeroing in on specific patterns may identify the worst dark patterns while leaving a broad gray area of harmful patterns untouched by regulation. Finally, existing definitions too often focus on designers' (and companies') intent,[31] which can be extremely difficult to prove. According to privacy scholar Chris Hoofnagle, "[I]t is impossible to prove a subjective intent to mislead, because so many different people are involved in a message's formulation. This is one reason why Congress did not require the FTC to prove intent in its Section 5 actions."[32] As we explain below, using a performance-based standard that focuses on outcomes and outputs, rather than intent, is a more effective means of assessment.

---

[29] Müge Fazlioglu, *Consolidating US privacy legislation: The SAFE DATA Act*, INT'L ASS'N PRIVACY PROFS. (Sept. 21, 2020), https://iapp.org/news/a/consolidating-u-s-privacy-legislation-the-safe-data-act/ [https://perma.cc/L7D6-RUUT].

[30] Jamie Luguri & Lior Jacob Strahilevitz, *Shining a Light on Dark Patterns*, 13 J. LEGAL ANALYSIS 43, 70–71 (2021).

[31] Several definitions center dark patterns around designer intent, most frequently by relying on the word "trick." *See* Brignull, *supra* note 7; Bösch et al., *supra* note 10, at 239; Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier & Igor Santos, *Can I Opt Out Yet?: GDPR and the Global Illusion of Cookie Control*, 2019 PROC. 2019 ACM ASIA CONF. ON COMPUT. & COMMC'NS SEC. 349; Nouwens et al., *supra* note 11, at 3 (using the term "malicious interaction flows").

[32] CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND PUBLIC POLICY 346 (2016).

D. Performance-Based Standards, Not Intent

As part of the FTC's enforcement authority, the agency is responsible for overseeing consumer protection in the United States. Section 5 of the FTC Act explicitly prohibits "unfair or deceptive acts or practices in or affecting commerce." The agency also enforces a number of other consumer protection statutes that prohibit specific practices, such as third-party sellers charging consumers without first providing a description of the goods or services being offered.[33] Though curbing the use of dark patterns remains a priority for the FTC, the agency presently does not explicitly ban them. Any enforcement action the FTC takes going forward must be tied to the deception or unfairness principles stipulated under its Section 5 authority.

The enactment of the CCPA and the passage of the CPRA are particularly timely as the FTC seeks to more aggressively curb the use of dark patterns. In September 2020, the agency filed charges against Age of Learning, a children's education company for use of deceptive practices in its subscription service. In a statement regarding the $10 million complaint against the company, former FTC Commissioner Rohit Chopra wrote, "Digital deception should not be a viable American business model. If the FTC aspires to be a credible watchdog of digital markets, the agency must deploy these tools to go after large firms that make millions, or even billions, through tricking and trapping users through dark patterns."[34] In spite of these strong statements, it is unclear if the FTC's current mandate prohibiting unfair or deceptive practices is sufficient to guard against dark patterns that undermine user privacy, especially those that straddle the line between dark and gray patterns.

Not all privacy dark patterns are based on deception, which raises the question as to whether the FTC's unfairness authority is sufficient to bring an enforcement action. While the FTC frequently brings privacy cases based on deceptive practices, it rarely brings cases based solely on unfairness.[35] This phenomenon is the result of the FTC being more reluctant to apply the legal standard for establishing unfairness after the agency was accused in the past of applying the unfairness standard too broadly.[36] Demonstrating a practice is

---

[33] *See* 15 U.S.C. §§ 8401–02.

[34] Statement of Rohit Chopra, FTC Commissioner, *Regarding Dark Patterns in the Matter of Age of Learning, Inc.*, Commission File No. 1723186 (Sept. 2, 2020), https://www.ftc.gov/system/files/documents/public_statements/1579927/172_3086_abcmouse_-_rchopra_statement.pdf [https://perma.cc/UKF4-5KVY].

[35] *See* Luguri & Strahilevitz, *supra* note 30, at 47; Hoofnagle, *supra* note 32, at 160.

[36] J. Howard Beales, III, *The Federal Trade Commission's Use of Unfairness Authority: Its Rise, Fall, and Resurrection*, 22 J. PUB. POL'Y & MKTG. 192, 193 (2003).

unfair rests on proving substantial injury to consumers, that the practice could not be reasonably avoided, and that the injury is not outweighed by countervailing benefits to consumers or to competition.[37]

The FTC Act standard for deception, while less limited than the unfairness standard, still requires proof that a practice is likely to mislead a reasonable consumer and that the misleading representation is "material" or "likely to affect a consumer's decision to purchase or use a product or service."[38] Materiality often refers to explicit claims made by a company about their product or service, which may not extend to their data collection practices. Furthermore, while proving an act or practice is deceptive does not require demonstrating an intent to deceive, "deception cases often rely heavily on evidence of the defendant's intent to deceive or knowledge of the deceptiveness of its practices."[39]

Thus far, FTC enforcement against the use of dark patterns has focused on the most egregious cases of consumer harm. While dark pattern violations in e-commerce are often tied to monetary harm, privacy harms are more difficult to demonstrate because they often involve the use of personal data by free online services.[40] Interference with individual autonomy is also challenging to measure. To make matters worse, many privacy dark patterns are deployed by free services where the goal of their use is obtaining one's personal information, not a traceable payment.[41]

Prescriptive rules are ill-equipped to protect consumers in an era where large firms can rapidly develop techniques to influence consumer behavior. Law professor Lauren Willis argues for an alternative approach, *performance-based regulations*, which instead "sets a measurable standard closer to the regulator's ultimate goal and allows the regulated entity to choose how to meet that standard."[42] Such a standard shifts the burden of proof from individual consumers to large firms, while still providing enough flexibility for companies to determine how to best meet performance goals. Furthermore, it is a forward-thinking approach, one that can anticipate changes in dark

---

[37] *See* Int'l Harvester Co., 104 F.T.C. 949, 1070–76 (1984) (*FTC Policy Statement on Unfairness*).

[38] *See in re* Cliffdale Assocs., Inc., 103 F.T.C. 110, 174-84 (1984) (*FTC Policy Statement on Deception*).

[39] Lauren E. Willis, *Deception by Design*, 34 HARV. J. L. & TECH. 115, 158 (2020).

[40] *See generally* Danielle K. Citron & Daniel J. Solove, *Privacy Harms*, GWU Legal Studies Research Paper No. 2021-11 (2021) (detailing the current issues with describing privacy harms and the various forms of privacy harms that exist currently).

[41] *See* SEBASTIAN RIEGER & CAROLINE SINDERS, DARK PATTERNS: REGULATING DIGITAL DESIGN 17 (Stiftung Neue Verantwortung, 2020), https://www.stiftung-nv.de/sites/default/files/dark.patterns.english.pdf [https://perma.cc/7C4D-MMA4].

[42] *See* Lauren E. Willis, *Performance-Based Consumer Law*, 82 U. CHI. L. REV. 1309, 1330 (2015) [hereinafter *Performance-Based*].

patterns over time. For example, Willis anticipates the use of machine learning by businesses to produce ads or even user interfaces optimized for a business's advantage that are organically, rather than intentionally, deceptive, where assigning intent could be functionally impossible.[43]

While performance-based standards avoid the challenge of proving intent, what remains is the task of measuring a pattern's effects and clearly determining its "darkness." As Mathur *et al.* argue, the data-driven optimization of user interfaces via widespread A/B testing,[44] particularly at scale, helped to facilitate the proliferation of dark patterns.[45] When different subsets of users view variants of the same web page, companies can determine which design choices optimize metrics like highest click-through rate without considering how those choices may undermine user autonomy or informed consent.[46] This process can be used to entrench dark patterns in user interface design as companies aim to meet specific benchmarks.

Although A/B testing contributed to the widespread use of dark patterns in web design, changing the metrics for which companies use A/B testing may be a part of the solution. Measuring the "darkness" of privacy dark patterns might require a number of empirical research methods;[47] one approach may involve showing variants of a web page to assess how different interfaces affect consumer expectations about the collection of their personal information, or how well consumers understand the terms of specific offers or even the content of privacy policies.[48]

---

[43] *See* Willis, *supra* note 39, at 156.

[44] *See generally* ROCHELLE KING, ELIZABETH CHURCHILL, & CAITLIN TAN, DESIGNING WITH DATA ii (O'Reilly Media 2017) ("A/B testing is a methodology to compare two or more versions of an experience to see which one performs the best relative to some objective measure.").

[45] *See* Arvind Narayanan, Arunesh Mathur, Marshini Chetty & Mihir Kshirsagar, *Dark Patterns: Past, Present, and Future*, ACM QUEUE Mar.–Apr. 2020, at 10.

[46] *See, e.g.*, Ed Felten, *On the Ethics of A/B* Testing, FREEDOM TO TINKER (Jul. 8, 2014), https://freedom-to-tinker.com/2014/07/08/on-the-ethics-of-ab-testing/ [https://perma.cc/M2L5-Q6JS]; KATHARINE SCHWAB, *We're All Being Manipulated by A/B Testing All The Time*, FAST CO. (Feb. 14, 2019), https://www.fastcompany.com/90306916/were-all-being-manipulated-by-a-b-testing-all-the-time [https://perma.cc/QU8H-PPTM].

[47] *See generally* Mathur et al. *supra* note 3. (suggesting user observation, in the lab as well as field studies, as the primary means by which to evaluate how consumers interact with a particular design). Specific applications, in addition to A/B testing, include surveys, interviews, focus groups, ethnographic research, gaze tracking systems, and controlled experiments. *Id.*

[48] *See* Narayanan et al., *supra* note 45, at 9–10; *see also* Luguri & Strahilevitz, *supra* note 30, at 47.

III. TEST CASES: THE CCPA AND THE CPRA

What is the best way forward for regulating dark patterns? Conveniently, California provides two new laws that we can use as a test case to evaluate the strengths and weaknesses of their particular approach. First, however, we believe regulations must avoid descriptive catalogs of dark patterns and embrace a definition that can apply equally to different types of dark patterns. Next, a definition should focus on outcomes rather than designer intent. To these ends, we apply Mathur, Mayer, and Kshirsagar's definitional framework,[49] discussed in Section II.A. above, to the CPRA's definition of dark patterns as a test case to help illuminate which approaches might be the most effective and where the challenges lie.

We draw on the two of the four normative lenses of Mathur *et al.* that we believe are the most relevant to the CPRA's definition of dark patterns: individual autonomy and regulatory objectives. Individual autonomy, according to the authors, "is the normative value that users have the right to act on their own reasons when making decisions,"[50] and is undermined by the vast majority of dark patterns. However, applying this lens does not automatically lead to a clear assessment of a dark pattern. For example, there is an inherent assumption that individuals can make decisions in their own best interests, and yet that is not always the case—there can be legitimate reasons to interfere in a consumer's decision-making process, as the debate on nudges illuminates. It is nevertheless possible that a balancing test could be developed in this area to identify to what extent the outcome of a dark pattern produces a result that benefits the designer over consumers, especially if that result leads to an objective harm.

The regulatory objective lens "uses democratically created rules and standards to view when dark patterns cause individual and collective harms such as diminishing the individual's financial welfare and undermining fair market competition, respectively."[51] Instead of relying on a moral, ethical, or societal norms-based foundation, this approach "takes the operating legal regime as a given in making this judgment about an interface."[52] The aforementioned proposed DETOUR Act of 2018 is an example of this approach, which sought not only to prohibit dark patterns (though it did not use the term) used to "obtain consent or user data", but also to ban behavioral or psychological experiments or studies without explicit user consent, as well

[49] Mathur et al. identify four normative lenses—individual welfare, collective welfare, regulatory objectives, and individual autonomy—to "attempt to explain why dark patterns should concern us." Mathur et al., *supra* note 3, at 13.
[50] *Id.* at 18–19.
[51] *Id.* at 17.
[52] *Id.*

as addictive interfaces or services specifically targeted at children under age 13.[53] Similarly, the General Data Protection Regulation (GDPR) in the European Union provides explicit interface guidelines for invalid consent, noting that pre-ticked boxes or a lack of affirmative activity do not constitute online consent for data processing.[54] The regulatory objective lens, the authors note, "does not inherently advance a normative argument about why we should care [about] financial losses, privacy harms, or cognitive burdens, beyond noting whether the law directs us to care about those values."[55] However, it is subject to the whims of lawmakers, meaning that misunderstandings of dark patterns or poorly written definitions with unintended consequences can slip into law.

A. CCPA Dark Pattern Revisions

Prior to the passage of the California Consumer Privacy Act (CPRA), the CCPA was amended in March 2021 with language that addressed dark patterns specifically related to the statute's "do not sell my personal information" opt-out provision. A research study conducted by Consumer Reports of the CCPA's personal data sales opt-opt mechanisms, as well as similar empirical evidence gathered by this paper's authors, demonstrated that some businesses were using design elements and mechanisms to interfere with consumers' exercising of their opt-out rights.[56] These dark patterns included excessively long or complex opt-out processes, confusing or

---

[53] Deceptive Experiences To Online Users Resolution Act, S. 1084, 116th Cong. (2019).

[54] Specifically, Recital 32: "Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. This could include ticking a box when visiting an internet website, choosing technical settings for information society services or another statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity should not therefore constitute consent." Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1 art. 32.

[55] Mathur et al., *supra* note 3, at 18.

[56] MAUREEN MAHONEY, CONSUMER REPORTS, CALIFORNIA CONSUMER PRIVACY ACT: ARE CONSUMERS' DIGITAL RIGHTS PROTECTED? 4–5, 23 (2020) https://advocacy.consumerreports.org/wp-content/uploads/2020/09/CR_CCPA-Are-Consumers-Digital-Rights-Protected_092020_vf.pdf [https://perma.cc/7MVK-F6RA]. *See also* Jennifer King, *Dark Patterns and the CCPA*, STAN. L. SCH.: CTR. INTERNET & SOC'Y (Oct. 30, 2020), https://cyberlaw.stanford.edu/blog/2020/10/dark-patterns-and-ccpa [https://perma.cc/3TF3-QM5H].

"confirmshaming" language, requiring unnecessary personal information to complete an opt-out request, and forcing consumers to search for opt-out links. Given that most businesses across a wide range of sizes and types were able to provide opt-out mechanisms that were simple and direct, we and other researchers concluded these design choices had the effect of making it challenging for consumers to exercise their rights, regardless of the intent.

Notably, the CCPA did not rely on a definition of dark patterns, nor even used the term. Instead, it invoked the following language to create both a baseline allowable condition as well as narrowly targeted guidelines of what to avoid:

> A business's methods for submitting requests to opt-out shall be easy for consumers to execute and shall require minimal steps to allow the consumer to opt-out. A business shall not use a method that is designed with the purpose or has the substantial effect of subverting or impairing a consumer's choice to opt-out.[57]

This step was appropriate for addressing a problem specific to the CCPA. The CPRA, however, expands both laws' reach by providing a definition of dark patterns that can apply to broader contexts.

B. The CPRA's Definition of Dark Patterns

The CPRA, which will go into effect in 2023 after a public comment and rulemaking process, introduces two elements related to dark patterns. One, it provides the first definition of dark patterns in U.S. law: "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."[58] Per Mathur *et al.*, this definition falls under their individual autonomy lens: "a dark pattern is a user interface that undermines individual decision-making."[59] Two, the CPRA includes a provision that explicitly forbids obtaining consent related to the processing of personal information by means of dark patterns. This maps to the regulatory objectives lens of Mathur *et al.*, where a dark pattern is defined as "any interface that modifies the choice

[57] California Consumer Privacy Act, CAL. CODE REGS. tit. 11, § 999.315(h) (2021).
[58] California Privacy Rights Act, CAL. CIV. CODE § 1798.140(l) (West 2020) (effective Jan. 1, 2023). We note that this definition appears to have been adapted from the definition included in the DETOUR Act.
[59] Mathur et al., *supra* note 3, at 18.

architecture to interfere with or undermine specific regulatory objectives."[60] While this provision sounds fairly narrow, in practice its application will likely be very broad, covering nearly every collection or exchange of personal information for digital processing in California. The combination of an autonomy-based definition with regulatory objectives allows the CPRA to take a normative stance on dark patterns while directing enforcement towards specific outcomes.

C. Unpacking the Language

One of the strongest aspects of the CPRA's definition is that it focuses on outcomes, also understood as *effects*, rather than intent. As mentioned above, many non-legal definitions of dark patterns are centered on the intent of the designer, which even in the best cases may be difficult to prove. Mathur *et al.* note that with a regulatory objective lens, "the underlying regulations that define the standards that any interface needs to meet do not have to be described with a great deal of specificity." That said, the specified applicability to "user interfaces" could require further definition or expansion. This term could be construed as applying narrowly to visual digital user interfaces, which could make it a challenge to include dark patterns in other types of user interfaces, such as voice-based interfaces, or even physical (non-digital) interfaces, like Internet-enabled devices. The definition also explicitly refers to manipulation,[61] emphasizing that this aspect is distinct from unfairness or deception, which we will discuss in more detail below.

D. Identifying and Evaluating Consent-Based Dark Patterns

In the specific case of obtaining consent for the disclosure of personal information, when a suspected dark pattern does not rise to the level of deception or unfairness, the CPRA specifies two principles by which consent may be invalidated:[62]

- Consent must not be *coerced;*
- Consent must not be *manipulated*.

While this Section examines each consent principle, our analysis is not exhaustive. In fact, there are undoubtedly other ways by which to make online

[60] *Id.* at 17.
[61] California Privacy Rights Act § 1798.140(l).
[62] *See id.*

consent mechanisms both non-coercive and non-manipulative than what we describe here.

E. Coercive Consent

According to Susser *et al.*, coercion influences someone "by constraining their options, such that their only rational course of action is the one the coercer intends."[63] Although coercion robs an individual of choice, the individual is still ultimately able to make a decision—just perhaps not the one they might have arrived at of their own accord absent the coercive influence. In practice, this manifests as control by the designer over the range of possible decisions, or the "decision space."[64] In this space, the easiest or most logical decision by the consumer is the one that results in what the designer wants or intends. For example, the typical online consent process consists of a lengthy set of terms or a privacy policy that consumers rarely read. Capitalizing on this well-known behavior, designers usually present a consent screen that allows consumers to avoid reading the terms or policy entirely, where the consent option (*e.g.*, "I Accept", "I Agree", or "I Consent") is often a highlighted or starkly contrasted button or link, while the non-acceptance option is often presented in a low contrast color and smaller size relative to the consent option.

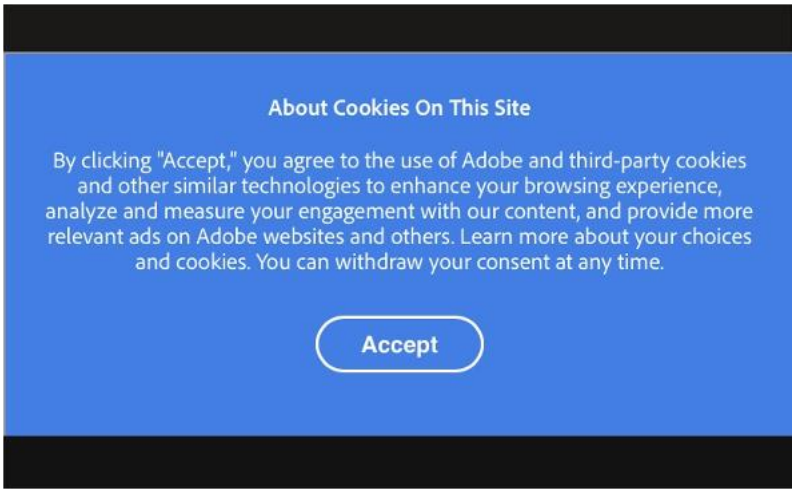

Figure 4: Example of coercive cookie consent.

A "light pattern"—or non-coercive—design principle could instead require consent options that place one's competing options at equal

[63] Susser et al., *supra* note 23, at 4.
[64] *See* Mathur et al., *supra* note 3, at 9.

prominence without favoring one choice over the other. If this light pattern were presently mandated, it would invalidate the vast majority of consent mechanisms we see online today, where the "I Accept" button or link is preselected, highly color-contrasted, or much larger in size than the decline button or link. This interpretation also suggests that not only must decline options be given equal weight, *they must actually exist in the first place*. A sleight of hand formerly widespread in EU cookie consent dialogs was the presentation of a consent option with no corresponding option to decline.[65]

Similar to the CCPA, the CPRA also contains language requiring that the process for opting-out of data sales must be as straightforward as the act of accepting or opting-in.[66] Enforcement of this standard could result in the creation of a measurable and quantifiable baseline requirement that declining consent shall not create a burden on consumers. Trying to opt-out of data sales would no longer deliver users into a pit of interaction despair, where they are forced to wander through flows with excessive steps or jump through unreasonable hoops—such as being required to call a company representative by phone despite having signed up online—in order to opt-out. For an example of how prevalent this issue has become, readers can refer to the report issued by the Norwegian Consumer Council documenting the difficulty for Amazon.com customers when unsubscribing from Amazon Prime. The report is quite critical, noting that, "Consumers who want to leave the service are faced with a large number of hurdles, including complicated navigation menus, skewed wording, confusing choices, and repeated nudging." [67]

[65] For example, the examination of EU consent management platforms by Nouwens *et al.* found that 32.5% of the 680 sites they scraped relied upon implicit consent in defiance of the GDPR. Nouwens *et al.*, *supra* note 11, at 5.

[66] California Privacy Rights Act § 1798.135(b)(2)(A) ("The consent web page also allows the consumer or a person authorized by the consumer to revoke the consent as easily as it is affirmatively provided."). *See also* the amended language of California Consumer Privacy Act § 999.315(h)(1) ("The business's process for submitting a request to opt-out shall not require more steps than that business's process for a consumer to opt-in to the sale of personal information after having previously opted out. The number of steps for submitting a request to opt-out is measured from when the consumer clicks on the "Do Not Sell My Personal Information" link to completion of the request. The number of steps for submitting a request to opt-in to the sale of personal information is measured from the first indication by the consumer to the business of their interest to opt-in to completion of the request.").

[67] NOR. CONSUMER COUNCIL (FORBRUKERRÅDET), YOU CAN LOG OUT, BUT YOU CAN NEVER LEAVE 3 (2021), https://fil.forbrukerradet.no/wp-content/uploads/2021/01/2021-01-14-you-can-log-out-but-you-can-never-leave-final.pdf [https://perma.cc/78V8-2RRK].

F. Manipulative Consent

Susser *et al.* differentiate manipulation from deception or coercion; by definition, manipulation is *hidden* influence, where pattern creators drive an individual into taking an action by displacing their decision-making authority. Thus, any interaction where the choice of consent is made for a user, or implied through some other action they take, is a manifestation of manipulative consent. The CPRA's consent definition includes specific language that provides examples of manipulative interaction: "[a]cceptance of a general or broad terms of use or similar document that contains descriptions of personal information processing along with other, unrelated information, does not constitute consent. Hovering over, muting, pausing, or closing a given piece of content does not constitute consent." In short, any mechanism that presupposes the user's choice and affirms her consent without her taking a direct action is manipulative.

Viewed from a human-computer interaction perspective, the CRPA's definition and examples have the potential to upend many current approaches to consent. However, this potential lies in the hands of the new California Privacy Protection Agency, as mandated by the passage of the CPRA, and the Agency's willingness to adopt a more expansive approach to regulation. For example, many companies design their initial account creation process in such a way that users must consent to both the Terms of Service (TOS) and the privacy policy by clicking a single button that simultaneously initiates the creation of a new account and signifies their consent. Some researchers have suggested that this approach in the EU, referred to as bundled consent, violates the GDPR.[68] A crucial question is whether asking consumers to consent to both a TOS and a privacy policy with a single interaction would meet the definition of a dark pattern under the CPRA, especially if there are no options given to reject either one separately. Furthermore, there is ample research demonstrating that not only does the vast majority of the public not read privacy policies, they may not even understand them if they tried.[69] Given this,

---

[68] *See* Nouwens et al., *supra* note 11, at 2; *see also* Délibération n° 2019-093 du 4 juillet 2019 portant adoption de lignes directrices relatives à l'application de l'article 82 de la loi du 6 janvier 1978 modifiée aux opérations de lecture ou écriture dans le terminal d'un utilisateur (notamment aux cookies et autres traceurs) (rectificatif) [Deliberation no. 2019-093 of July 4, 2019 adopting guidelines relating to the application of article 82 of the law of January 6, 1978 as amended to read or write operations in a user's terminal (in particular cookies and other tracers) (corrigendum)] JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE [J.O.] [OFFICIAL GAZETTE OF FRANCE], Jul. 19, 2019, p. 284; Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, & Thorsten Holz, *(Un)informed Consent: Studying GDPR Consent Notices in the Field*. ACM 973, 973–90 (2019).

[69] *See* Kevin Litman-Navarro, *We Read 150 Privacy Policies. They Were an Incomprehensible Disaster*, N.Y. TIMES (Jun. 12, 2019),

one may ask whether embedding multiple forms of information collected into a long-form privacy policy with a single "I Agree" button affixed to it is in fact manipulative consent. If consumers are not provided with the actionable and meaningful ability to make an independent decision about all aspects of information collection, but instead are forced to reckon with all of them at a single point in time with a single action, it raises the question as to whether the consent is obtained in an inherently manipulative manner. If this were found to be the case by regulators, it would have profound consequences for online consent as we currently experience it.

G. Shifting Paradigms

To be clear, the optimal outcome is not one where consumers are given more checkboxes to check and buttons to click in the name of "compliance." If we are not careful about how we interpret coercion and manipulation, consent mechanisms will merely be fragmented into more rote and meaningless actions rather than transformed into new mechanisms that are more substantive, meaningful, and informative.[70] In prohibiting dark patterns, the CPRA creates an opportunity for California to lead by example and develop standards that demonstrate best practices—or light patterns—for consent.

A focus on regulatory objectives emphasizes the necessity of using a performance-based standard. As Mathur *et al.* also note, this perspective has a significant advantage: "fashioning regulation into measurable metrics for empirical research is usually much easier than adapting normative principles to research."[71] But grounding a definition of dark patterns on interference with individual autonomy provides a basis for evaluation—one based on human-centered design—that emphasizes the need to incorporate an understanding of human behavior when evaluating them. Doing so opens a door for regulators to avail themselves of a rich set of methods, tools, and research literature on which to evaluate consent mechanisms. It also creates an opportunity to rethink the mechanisms themselves by broadening the standards by which we evaluate them. For example, according to law professor Nancy Kim, the current contract law paradigm by which we traditionally evaluate notice and consent mechanisms "does not require actual (subjective) knowledge" for

---

https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html [https://perma.cc/3TNQ-LGFV].

[70] *See* JENNIFER KING, ANNE JOSEPHINE FLANAGAN, AND SHEILA WARREN, WORLD ECON. FORUM, REDESIGNING DATA PRIVACY: REIMAGINING NOTICE AND CONSENT FOR HUMAN-TECHNOLOGY INTERACTION 13–33 (2020) (for a discussion on transforming notice and consent).

[71] Mathur et al., *supra* note 3, at 17.

consent.[72] Instead, contract law substitutes capacity and access to information, or notice, for knowledge.[73] A human-centered design approach to consent, in contrast, could place the individual's ability to understand what they are consenting to squarely at the core of what it means to consent, forcing us to fundamentally rethink how to provide non-coercive and non-manipulative digital consent.

H. Applying the CPRA Definition in Practice—Challenges and Limitations

As with most complex issues, the challenge with regulating dark patterns lies in the details. Even with a workable definition in place, the most egregious violations will be the easiest to single out for enforcement. The real challenge will come with evaluating the designs that are at the border between dark and gray. At this border is where performance-based standards and agreements on measurement are crucial.

Unfortunately, measuring how dark a pattern might be is neither simple nor well-established. Academics have only very recently begun publishing user studies of dark patterns; at this time, we are only able to identify two studies that do so.[74] Mathur *et al.* make specific suggestions as to what to measure for each of their normative lenses. As an example for regulatory objectives, Mathur *et al.* suggest that "the goal of any measurement should be to assess whether a dark pattern complies with relevant regulation,"[75] using the work of Nouwens *et al.*, Utz *et al.*, and Machuletz *et al.*, all of whom have studied cookie consents in the EU, as examples. Furthermore, there are questions regarding what the appropriate thresholds for influence, as well as who is affected by dark patterns, that are unsettled. There is some related prior work in this area, much of it based on research in the usable security and privacy fields[76] on privacy policies, though scholars of communications, consumer psychology, and even addiction may also have useful research to contribute to these questions. The FTC has employed HCI academics and professionals as experts in their investigations,[77] but those

---

[72] NANCY S. KIM, CONSENTABILITY: CONSENT AND ITS LIMITS 83 (Cambridge Univ. Press 2019).

[73] *Id.*

[74] *See, e.g.*, Dominique Machuletz & Rainer Böhme, *Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR*, 2 PROC. ON PRIVACY ENHANCING TECH. 481, 481–98 (2020); Luguri & Strahilevitz, *supra* note 30, at 3.

[75] Mathur et al., *supra* note 3, at 21.

[76] *See generally* Florian Schaub, Rebecca Balebako, Adam L. Durity, & Lorrie Faith Cranor, *A Design Space for Effective Privacy Notices*, USENIX ASS'N 1, 1–17 (2015) (overviewing this topic and providing additional citations).

[77] *E.g.*, Testimony of Plaintiff's Expert Witness, Jennifer King, FTC v. Commerce Planet, Inc., No. SACV 09-1324, 2012 WL 12533790 (C.D. Cal. July 17, 2012); Expert Report of

investigations are often confidential and very few of the evaluations are in the public record. As of yet, the FTC has not issued any guidance on either design patterns that are inherently dark, or best design practices for areas where dark patterns are of concern.

Regulating dark patterns also runs up against several difficult questions that remain unsettled. For example, what are the limits of permissible persuasion in online environments, where the line between persuading and coercing or manipulating can be so easily crossed? Can the design community create "neutral" interfaces, and should regulators mandate them at key online decision-points, such as at sign-up, disclosure, or payment?[78] These are open questions that researchers should investigate—ideally with public funding—to build a corpus of findings free of industry influence.

Finally, there are undoubtedly limitations with the approach we've discussed here. For one, our analysis is prospective—California regulators will not even begin enforcing the CPRA until 2023. At the time of writing there were no enforcement actions made public with regard to the CCPA revisions by which to evaluate the effectiveness of the new regulations. Stacey Schesser of the California Attorney General's Office noted in January of 2021 that companies had been responsive to notices of alleged violations, though at the time none of those were related to dark patterns.[79] Furthermore, the CPRA itself may change as part of the regulatory process before it is finalized into law. And, regardless of the careful thought underlying the approaches we discuss, policymakers may still write regulations that are too narrow to identify gray areas, create unintended consequences, such as limiting legitimate design patterns, or spur a race among designers to exploit loopholes in interpreting definitions or measurement of outcomes. Notwithstanding these limitations, our aim is to highlight the most promising elements of the CPRA definition and to demonstrate how those elements might be meaningfully enforced.

Jennifer King, FTC v. Amazon.com, Inc., No. 2:14-cv-01038, 2015 WL 11252957 (W.D. Wash. Nov. 10, 2016); *see also* Luguri & Strahilevitz, *supra* note 30, at 40 (discussing the case law in this area).

[78] We flag this issue as one that needs additional specification; for example, is a neutral interface or design pattern simply free from persuasion, coercion, or manipulation? Or does the HCI community categorize it as benefitting the user through providing an "optimal" user experience, efficiencies, or having been tested widely? The question has not yet been answered by the HCI field but implicates what norms we are considering as either neutral or beneficial. For discussions about using neutral interfaces or design patterns for measurement purposes, *see* Mathur et al., *supra* note 3; Luguri & Strahilevitz, *supra* note 30.

[79] *See* Srivats Shankar, *FPF Hosted a CPDP 2021 Panel on US Privacy Law: The Beginning of a New Era*, FUTURE OF PRIVACY F. (Mar. 25, 2021), https://fpf.org/blog/fpf-hosted-a-cpdp-2021-panel-on-us-privacy-law-the-beginning-of-a-new-era/ [https://perma.cc/W6UQ-7LYJ].

I. The Future of Dark Patterns Enforcement in California

California has the opportunity to creatively imagine how to tackle privacy enforcement with the creation of the California Privacy Protection Agency. This new Agency will inherit the responsibility of overseeing the state's data privacy laws from the attorney general, and in so doing will create a new enforcement infrastructure from scratch. While hiring enforcement attorneys will likely be the Agency's top priority, the obligation to enforce the dark patterns provision will require staff not only with knowledge and expertise of privacy law, but also with education and experience in human-computer interaction and related fields. Employing technologists of many stripes will be crucial to building an enforcement agency for the 21st century. Tapping the expertise of HCI professionals to evaluate and enforce dark pattern policies is an important first step.

Regardless of staffing decisions, the Agency must, potentially in partnership with the FTC, convene a panel of outside experts and stakeholders to create uniform guidance on dark patterns across multiple contexts. In addition, it should recommend performance-based standards for identifying dark patterns, measuring their impact, and posit methods for evaluating them.

IV. CONCLUSION

As we write this Article, legislation on dark patterns is proliferating. Washington State debated a privacy law for the third time in three years, which includes the CPRA's definition of dark patterns verbatim.[80] This increased focus on regulating dark patterns could provide an opportunity to fundamentally reevaluate online consent mechanisms entirely, a topic often ignored in proposed privacy legislation at both the state and federal levels. There is also a clear need for more enforcement of dark pattern violations generally. This could be achieved either through a state-by-state approach or an expansion of the FTC's authority to regulate "manipulative" or "abusive" practices. The CPRA offers guidance on how to think about the regulation of dark patterns and related problems.

Regulating dark patterns requires more than writing legislation that attempts to prohibit the worst of the worst. As we discussed, it starts first with a definition that focuses on outcomes rather than intent. As the CPRA demonstrates, while the core definition may be based on a broad-ranging concept like individual autonomy, the regulation itself should coalesce around specific outcomes to avoid being overly broad. Further, it likely requires expanding consumer protection statutes beyond deceptive and unfair practices

[80] *See* Washington Privacy Act, S.B. 5062, 67th Leg. § 101(10) (Wash. 2021).

to explicitly include manipulation. Finally, regulation also requires agreement on how to measure and evaluate its effects, which we argue should utilize performance-based standards.

But in addition to merely charting a path forward in this space, confronting the design problems that dark patterns present also creates an opportunity to shift enforcement and regulatory paradigms. Instead of relying on the myth of the rational actor—which dominates thinking in both law and economics—we may move towards a human-centered approach that attempts to meet people where they are, rather than as we imagine them to be. Where privacy and personal disclosure intersect with dark patterns, we see the possibility to embrace a fundamental reckoning with the assumptions upon which mechanisms that enable dark patterns are built.