

Deceptive choice architecture and behavioral audits: A principles-based approach

Stuart Mills 

Economics, University of Leeds, Leeds, UK; and Behavioural Science, London School of Economics and Political Science, London, UK

Abstract

Regulators are increasingly concerned about deceptive, online choice architecture, including dark patterns and behavioral sludge. From a behavioral science perspective, fostering a regulatory environment which reduces the economic harm caused by deceptive designs, while safeguarding the benefits of well-meaning behavioral insights, is essential. This article argues for a principles-based approach and proposes behavioral audits as a tool to support this approach.

Keywords: AI, behavioral audits, choice architecture, principles-based regulation, regulation.

1. Introduction

Regulators are increasingly concerned about the uses of behavioral science to create deceptive websites and online applications. The Biden Administration has committed to tackling “junk fees”—deceptive designs that hide additional fees—as have the Federal Trade Commission (FTC; Deese et al., 2022; FTC, 2022a). In the UK, the Competition and Markets Authority (CMA) and the Information Commissioner’s Office (ICO) have expressed concern about the harms of “behavioral design” on competition and consumer welfare (CMA, 2021; ICO & CMA, 2023; Sugg & Lesic, 2022; also see FCA, 2022). The European Union’s (EU) recent legislation on artificial intelligence (AI) includes discussion of AI for deceptive design practices within online markets (European Commission, 2021b). Their recent *Digital Services Act* (DSA; European Commission, 2022a) considers deceptive choice architecture and dark patterns (also see European Commission, 2021a, 2022c). The Indian government has also begun consulting on regulation in this area (Indian Department of Consumer Affairs, 2023). With such mounting scrutiny, applied behavioral science is entering a new era.

Applied behavioral science draws on insights from economic psychology to design choice environments to influence decision-makers and change behavior, often for the benefit of citizens and consumers (Sunstein et al., 2017). Behavioral science insights have been used to design effective product labels for goods, allowing consumers to make more informed decisions (Sunstein, 2021b; Sunstein & Reisch, 2014; Thunström, 2019). Various behavioral “nudges” (Thaler & Sunstein, 2008) have been used to tackle the deceptive effects of online dis- and misinformation (Pennycook et al., 2020; Roozenbeek et al., 2021). Regulators themselves have used behavioral science in areas such as gambling (e.g., Department of Culture, Media and Sport, 2023) and retail investing (FCA, 2021). The use of behavioral science to redesign processes is often referred to as choice architectural design (Thaler & Sunstein, 2008). While it is always wise to take a holistic view of policy and regulation, weighing the merits of one approach against another (Mills & Whittle, 2023), evidence suggests good intentioned behavioral science can support consumers and citizens, not harm them (Dudley & Xie, 2022; Sanders et al., 2018).

For as long as there have been markets, there have been tricks and other psychological manipulations which harm consumers (Akerlof & Shiller, 2015). So too, concerns about the potential harms of behavioral science have long existed (e.g., Rebonato, 2014). Current regulatory concerns derive from the growing use of online choice environments in everyday life (e.g., ICO & CMA, 2023). Entertainment, finance, government services, and more, are increasingly accessed via smartphone applications and websites. It is the potential for technology to make deceptive design practices endemic throughout online markets which is driving current regulatory concerns.

Correspondence: Stuart Mills, Economics, University of Leeds, Leeds, UK. Email: s.mills1@leeds.ac.uk

Accepted for publication 12 March 2024.

Deceptive design practices in online choice environments have been subject to discussion and debate within the user interface and user experience (UI and UX, respectively) domains for several years (Mills et al., 2023). Coined in 2010, *dark patterns* describe UI design features that lead users to outcomes which they do not want, but which benefit whatever service has introduced the design (Brignull, 2011). Substantial research into dark patterns has since emerged (e.g., Bösch et al., 2016; Conti & Sobieski, 2010; Gray et al., 2018; Kozyreva et al., 2020; Mathur et al., 2019; OECD, 2022; Sin et al., 2022; Waldman, 2020), though only in recent years has the link between dark patterns and behavioral science been explicitly investigated.

Some dark patterns are not based on behavioral science (Mills et al., 2023). For instance, forcing a person to accept a service's terms and conditions is not about influencing choices; *there is no choice*. But other dark patterns, such as time scarcity, default options, and social priming, share clear overlap with the behavioral science literature (Behavioural Insights Team, 2022; Kozyreva et al., 2020; Newall, 2022). Thus, many regulators concerned with dark patterns also link to behavioral science ideas (e.g., Sugg & Lesic, 2022).

This link is further compounded by behavioral science concept of *sludge*. Sludge describes transaction costs which arise because of choice architecture (Shahab & Lades, 2021). Sludge creates difficulties for consumers and citizens to achieve what they want (Thaler, 2018). For instance, unsubscribing from an online service, or accessing a government program (Sunstein, 2021a; Sunstein & Gosset, 2020). Sludge and dark patterns are not the same (Mills et al., 2023). For instance, dark patterns are exclusively *online* phenomena (Brignull, 2011); sludge can occur in "offline" spaces, such as filling out excessive (physical) paperwork (Sunstein, 2019b). Yet, because sludge, like dark patterns, can be used to benefit, say, a business at the expense of a consumer, they are increasingly discussed together (Mills et al., 2023; Newall, 2023). From the perspective of organizations such as the CMA, both dark patterns and sludge represent threats to consumer welfare in terms of the potential to cause economic harm (CMA, 2021).

Some regulatory responses are emerging. The European Commission's (2022a) DSA discourages practices such as making services harder for individuals to leave than they are to join (also see European Commission, 2022c, para. 34). The UK's Financial Conduct Authority (FCA, 2022) has recently introduced a new "Consumer Duty" standard requiring financial services to prioritize safeguarding consumer interests. This extends to online choice architecture. Despite these efforts, however, such regulation presents challenges from a behavioral science perspective. This article analyzes these challenges.

The first is the challenge of subjective experiences of choice architecture. What one person considers a sensible, useful arrangement of choice architecture, another may find wholly deceptive and unacceptable. Such opposition is not necessarily opposition to behavioral science, but rather, arises from heterogeneity within the population. People are different, have different preferences, and may experience choice architecture in different ways (Martin et al., 2023; Mills, 2022; Sunstein, 2022a). As a result, from a behavioral science perspective, it is difficult to determine "objectively" unacceptable choice architecture.

The second challenge is preserving the use of behavioral science to affect positive outcomes for individuals. If one accepts that choice architecture is inevitable (e.g., Sunstein, 2014, 2017), a program to reduce economic harms caused by deceptive online designs should also be married with a positive program to support the application of behavioral science to *improve* said designs, and thus individual outcomes (Thaler & Sunstein, 2008). This is a motivation for much of the modern, applied behavioral science industry, particularly in the United States and the UK (e.g., Hallsworth, 2023; Hallsworth & Kirkman, 2020; Halpern, 2015). Integrating positive uses of behavioral science into a regulatory approach which effectively counteracts deceptive choice architecture is thus an important consideration.

This article argues that, from a behavioral science perspective, a principles-based approach to regulating dark patterns and deceptive choice architecture provides flexibility given the above challenges. A key mechanism to support this regulatory approach will be *behavioral audits*. Behavioral audits should be used to ensure organizations comply with the *spirit* of regulatory principles, as well as offering insights for improving an organization's choice architecture for the benefit of consumers and citizens. Through a discussion of the nascent sludge auditing literature and current regulatory positions, this article outlines a loose program of research and development for behavioral auditing tools, given some suggested regulatory principles.

In presenting an analysis of regulatory positions from a behavioral science perspective, the ambition of this article is to add this perspective to existing regulatory and legal discussions about dark patterns and deceptive

choice architecture, and support the development of effective regulation in this area. The structure of this article is as follows. Section 2 reviews the regulatory environment, focusing on the United States, UK, and EU regulatory jurisdictions. Section 3 elaborates on the two challenges hereto thus far briefly discussed. Section 4 outlines a program for developing behavioral audits, focusing on several key questions. Section 5 concludes.

2. The current regulatory environment

Deceptive choice architecture is a concern for regulators around the globe. Regulators, such as the FTC in the United States, and the CMA in the UK, as well as legislative bodies such as the European Commission in the EU, have focused on two aspects of deceptive choice architecture. Namely, how these designs harm consumers, and how they undermine market competition.¹ The UK emphasizes the commonality between consumer outcomes and competitive outcomes (CMA, 2021), while the United States (FTC, 2022a para. 2, 2022b) emphasize consumer outcomes under the broad notion of “junk fees.”

In terms of deceptive choice architecture, the EU’s DSA focuses more on consumers, while their *Digital Markets Act* (DMA) focuses more on market competition (European Commission, 2022a).² The DSA itself is only the most recent of several contributions within the EU which interpretively form a regulatory position on dark patterns and deceptive choice architecture. Existing provisions, such as the General Data Protection Regulation (GDPR) and the Unfair Commercial Practices Directive (UCPD), have been suggested to partially address some concerns already,³ as do proposed acts, including the *Data Act* (e.g., European Commission, 2022c, para. 34) and the *AI Act* (e.g., European Commission, 2021b, para. 15–16). Within the European regulatory space, the DSA is thus not a lodestar, but one of several measures which capture a broad regulatory position (also see Luguri & Strahilevitz, 2021).

All are motivated by similar concerns. The CMA (2021) and the FTC (2022b) are specifically concerned with economic harms that arise through consumers being manipulated into choices they otherwise would not have made. The DSA, being broader, considers harms beyond economic harms.⁴ Regulators recognize that these harms are not new, and emphasize that it is technology which is now elevating regulatory concerns. The CMA (2021, p. 6) write that, “recent advances in technology mean that many algorithms today can enact changes at a scale that makes these harms more pronounced.” The FTC (2022b, p. 1) note that, “For decades, unscrupulous direct mail marketers and brick-and-mortar retailers have relied on design tricks and psychological tricks, such as pre-checked boxes, hard-to-find-and-read disclosures, and confusing cancellation policies, to get consumers to part with their money or data. As more and more commerce has moved online, so too have these manipulative design practices.”

Different regulators place different emphases on the role of technology within their approaches to deceptive choice architecture. The CMA (2021) argues that algorithms enable increasingly personalized online experiences, prompting two concerns. First, that some online personalization may disadvantage consumers and benefit vendors, such as personalized pricing. Second, that *claims* of personalization may deceive consumers into believing options are presented in their subjective interests, hiding alternative vendor motives.

The FTC (2022a, 2022b) focuses on “junk fees,” specifically, and on dark patterns generally. They suggest the internet and online intermediaries have increased the number of outlets where junk fees and psychological tricks analogous to dark patterns are found, and blurred the distinctions between domains, hampering regulatory responses. The FTC (2022b) thus shares several concerns with the CMA (2021).

The European Commission (2022a) focuses on targeting algorithms used by “digital intermediaries” such as social media websites. They contend that online advertising is a legitimate (and sometimes useful) activity for these vendors to carry out. The Commission is keen to not inadvertently attack this business practice while attempting to reduce dark patterns and deceptive choice architecture.⁵ Much like the CMA (2021), the European Commission through the DSA focuses on designs which could lead consumers to suffer material harms they would have otherwise avoided without the deceptive design. Yet, it also emphasizes that such designs may impinge on individual autonomy, and that algorithms may further undermine autonomy by reducing the transparency of vendor practices. The Commission thus emphasizes *understandability*, and the threat posed to consumer understanding by dark patterns.

Behavioral science, dark patterns, nudging, and sludge, all feature within regulator discussions. The CMA (2021, p. 8) is specifically concerned with “online choice architecture,” arguing that, “Although choice architecture can be used to benefit consumers, firms may instead exploit inherent weaknesses in consumers’ ability to engage in markets in ways which go against consumers’ best interests... These harmful user interface design choices are known as ‘dark patterns’.” The European Commission (2022a, p. 56) writes that, “Dark patterns... are practices that materially distort or impair, either purposefully or in effect, the ability of recipients of the service to make autonomous and informed choices or decisions... Providers of intermediary services should therefore be prohibited from deceiving or nudging recipients of the service.” They include techniques like “presenting choices in a non-neutral manner, such as giving more prominence to certain choices through visual, auditory, or other components.”⁶ The FTC (2022b, p. 2) single out “cognitive biases” as the mechanism through which consumer manipulation occurs (also see e.g., European Commission, 2021a, p. 101).

Writing on junk fees, the FTC (2022a, para. 3) note that, “These types of extra or redundant fees can mislead consumers, or prevent them from knowing the true cost of a purchase until they’ve already invested substantial time and energy. At that point, they may feel like it’s too late to walk away.” Writing on dark patterns, the FTC (2022b, p. 6, emphasis added) note that “Companies are on the hook for the *net impression conveyed* by the various design elements of their websites.” Both statements align with notions of behavioral sludge (Mills et al., 2023) and “nagging” dark patterns (Gray et al., 2018). Similarly, the CMA (2021, p. 4) express their concern at the “combined effect” of nudge-like designs, as does the DSA, which suggests services should not be “significantly more cumbersome” to leave than to join (European Commission, 2022a, p. 57). The notion of a “*death by a thousand nudges*” is a frequent theme among regulators.

Valuable insights regarding the powers sought by regulators can be found within these various discussion pieces. Both the CMA (2021)⁷ and the FTC (2022b) have advocated for an *ex-ante* regulatory environment. Both note that current powers only allow these bodies to respond to complaints after-the-fact (i.e., *ex-post*). The CMA (2021) suggests an *ex-ante* approach will reduce harms generally, though seems to favor this approach more so from the perspective of stymying anti-competitive practices.⁸ The FTC (2022b, p. 3) offer a more consumer-centric justification, emphasizing “special enforcement challenges.” They note that, “Because dark patterns are covert or otherwise deceptive, many consumers don’t realize they are being manipulated or misled... [E]ven when consumers do realize they have been deceived, many don’t report their experiences, some out of an unnecessary feeling of embarrassment at being tricked.” The FTC (2022b) advocates for an *ex-ante* approach to allow regulatory intervention before a critical mass of complaints arise because by the time this critical mass is achieved, many more consumers will have suffered *in silence*.

In summary, regulators are concerned with the economic and material harms of deceptive choice architecture on consumers. There is recognition that many deceptive practices have existed for decades, if not longer. However, new technologies—particularly recommendation algorithms, platform intermediaries, and AI—mean that such practices are growing in scale and sophistication. This has prompted current regulator attention. Regulators in the United States, the UK, and the EU all recognize that behavioral science ideas describe (and could *prescribe*) the mechanisms which underpin these deceptive practices.

3. Regulatory challenges

The above discussion highlights several challenges around regulating deceptive choice architecture. Some are found within the position pieces of regulators themselves, though benefit from being made explicit here. Others emerge from regulator perspectives and ambitions. From a behavioral scientist’s perspective, all are important influences on how online choice architecture should be regulated.

3.1. Subjective experiences of choice architecture

The first challenge comes from subjective experiences of choice architecture. Generally, choice architecture captures the context in which decisions are taken, including the ordering of options, use of color, and many other elements of a decision’s framing (Thaler et al., 2012). The meaningful redesign of choice architecture can lead to statistically significant shifts in a population’s behavior (e.g., Beshears & Kosowsky, 2021; Jachimowicz

et al., 2019). Nevertheless, individuals will respond differently to choice architecture—some may benefit from the redesign; others may not; others still may be indifferent (Mills, 2022; Thunström et al., 2018).

Choice architecture which benefits one individual, or indeed many, is still likely to be regarded as detrimental to a minority of individuals within a population (Sunstein, 2022a). For instance, some studies (e.g., Thunström et al., 2018) have found that choice architectural interventions to influence spending habits can disproportionately target those who stand to benefit the least, while offering much smaller benefits to those for whom the intervention is primarily designed. Alternatively, the reverse might also, occasionally, be true—choice architecture intended to exploit consumers may actually benefit *some* consumers, given their specific characteristics.

There is likely to be significant variation in which aspects of choice architecture affect the behavior of individuals, and individuals are likely to have a diversity of preferences (Mills, 2022). For instance, choice architecture might induce a choice which only *later* results in harm (Dolan & Galizzi, 2015). Not only does this complicate the assigning of blame to choice architecture; it extends the necessary foresight of a regulator (and well-meaning choice architect). There may also be differences in the mere *perception* of choice architecture. Some may simply perceive some designs as negative and reject them, regardless of how designs would hypothetically influence their behavior (e.g., Reisch & Sunstein, 2016). Some may even moderate their perceptions (positive or negative) of choice architecture by their attitudes toward those who are trying to influence them (e.g., Tannenbaum et al., 2017). Thus, it becomes exceedingly difficult to determine an example of choice architecture that is *objectively* deceptive, insofar as it *always* causes economic harm to consumers.

Regulators recognize the challenge of subjective experiences. The CMA (2021, p. 4), for instance, acknowledge that, “it is difficult to see the combined effect of smaller non-price ‘nudges’ [on consumer welfare].” They further acknowledge an absence of research which readily resolves the challenges of determining “harmful” choice architecture from “non-harmful.” Finally, they express concern that an increase in personalized choice architecture within online spaces, supported by technologies such as AI, further obfuscates the task of establishing clear classifications for unacceptable choice architecture. Likewise, the FTC (2022b, p. 3) recognizes, in addition to the challenge of *noticing* and *reporting* deceptive practices, that deceptive designs will be experienced differently depending on the medium through which one accesses a service (e.g., smartphone versus desktop) and the socio-economic circumstances of the individual being manipulated.⁹ For instance, someone who cannot afford a junk fee is likely to experience the deceptive design which leads to the fee differently than someone who can readily afford it (Rawls, 1971).¹⁰ Emerging research into the role of inequality and administrative burdens offers further support to this perspective (Martin et al., 2023).

Thus, as an intermediate conclusion, it is difficult to defend the notion that “objectively” deceptive choice architecture exists. Different individuals will experience the same design practices differently for an array of reasons.

3.2. Positive uses of choice architecture

The second challenge is safeguarding the positive uses of choice architecture and behavioral science, while reducing the harm caused by deceptive uses. Choice architecture can be used for well-meaning purposes. Behavioral science may be useful as a countermeasure against harmful dis- and misinformation online (Pennycook et al., 2020); as a means of protecting consumers within the gambling industry (Newall & Rockloff, 2022); and as a means of protecting nascent investors entering financial markets (Newall & Weiss-Cohen, 2022). Smart disclosures can help consumers understand critical information more easily (Sunstein, 2019a), and choice architecture can be used to support citizens in accessing government provisions, such as in education (Page et al., 2019). Regulators, in discussing the harms of deceptive choice architecture, recognize the positive applications of these techniques, too (e.g., CMA, 2021, p. 8).

Nevertheless, from a behavioral perspective, a twofold risk emerges when regulating choice architecture.

There is a risk of over-regulation, or at the least, naïve standard-setting. For instance, the European Commission (2022a, p. 56) associates “exploitative design choices,” with “presenting choices in a non-neutral manner.” A longstanding argument within behavioral science is the *impossibility* of neutral choice environments (e.g., Sunstein, 2014; Thaler & Sunstein, 2008). This argument holds that for a choice to be presented, some presentation of said choice must be made, and this presentation may influence the decision-maker

(Sunstein, 2017).¹¹ Efforts to design “neutral” choice architectures are likely only to encourage a retreat from *any* purposeful use of behavioral science, for good or for ill. While such a risk, at present, lacks evidence to move it beyond a simple hypothesis; as the regulatory environment develops, monitoring emerging evidence in support of such a hypothesis seems prudent.

There is also a risk, owing to ambiguities surrounding what is and is not deceptive choice architecture, that businesses and governments resist implementing pro-consumer and pro-citizen choice architecture for fear of *accidentally* causing harm. Such a concern seems substantial for the CMA (2021, p. 5), who specifically acknowledge the importance of considering the often-unforeseeable effects of choice architecture, particularly when designed in conjunction with complex, algorithmic systems. Yet, while acknowledging the possibility of choice architecture *accidentally* causing harm, neither they nor others offer substantial recourse to fears which may arise in the minds of business leaders and policymakers from using choice architecture when penalties hang over them.

A regulatory landscape which discourages the use of welfare-enhancing choice architecture would be deleterious to the goals of the regulators.

Broadly, regulators are concerned that new technologies create information asymmetries between consumers and vendors (e.g., Akerlof, 1970), and that vendors can profitably leverage these asymmetries through choice architecture. The FTC (2022a, 2022b) frequently emphasizes how deceptive choice architecture often hides or obscures information (such as additional fees) which would be worthwhile to a consumer when making an informed decision. The European Commission (2022a) are resolute in ensuring that users of a service can understand *how* algorithms within the service work, and *how* the specific outcomes experienced by the user are determined. The CMA (2021, p. 17–18) expresses concerns that a lack of transparency in algorithmically determined rank-orders may deceive consumers into believing high-ranked options are optimal and disguise the vested interests of vendors.

As such, *redesigning* choice architecture to promote useful information and reduce asymmetries—as has been done, for instance, with food labels (Sunstein, 2021b; Thunström, 2019) and fuel-usage disclosures (Sunstein & Reisch, 2014)—could be an effective way of supporting the welfare of individual consumers and citizens. This necessitates the use of behavioral science for welfare-enhancing objectives.

3.3. A principles-based regulatory approach

A principles-based regulatory approach would offer recourse to the challenges thus far discussed. Principles-based approaches establish guiding principles, coupled with various obligations such as mandatory disclosures, to convey regulatory expectations while affording practitioners flexibility to pursue innovative practices (Black et al., 2007). Principles-based regulation relies on an evaluation of the outcomes of regulatory efforts to adjust (and potentially tighten) standards as practices evolve (Gilad, 2010). In recent years, this outcome-orientated approach has led to the development of a type of principles-based regulation known as performance-based regulation. Here, regulators set desired outcomes, while allowing regulatory flexibility in how those outcomes can be achieved, adjusting regulation based on the performance of an industry in meeting the outcomes (Coglianese, 2017; Willis, 2014). However, the flexibility of principles-based approaches may lead to ambiguity in terms of interpreting principles, hindering practitioner decision-making. Furthermore, it may allow practices which subvert the spirit of regulatory principles to emerge, subsequently undermining the regulatory regime (Decker, 2018).

In contrast to a principles-based approach, a rules-based approach sets firmer rules which practitioners must follow and must demonstrate evidence of complying with. This approach can reduce uncertainty for practitioners by establishing clear rules. It may also be better suited to safeguarding vulnerable stakeholders (Decker, 2018). Yet, rules-based approaches can be criticized for a lack of flexibility, which could stifle innovation. Furthermore, determining effective regulatory rules requires much foresight on behalf of the regulator, which may not always be possible, and may not be a reasonable expectation (Sunstein, 1991). Even with established rules, there is still much license for practitioners to interpret and *misinterpret* rules (Sunstein, 1995).

A principles-based approach to regulating online choice architecture is likely preferable given the challenges of creating rules for subjective choice architectural designs. For instance, the DMA (European

Commission, 2022b) requires gatekeepers—those who provide services which enable other services—to provide users with the ability to change their default settings easily to prevent the gatekeeper from giving affiliate products preferential positioning. This presents challenges. First, what is easy to one individual may be tricky for another, such as someone lacking adequate digital literacy. Second, strategies such as personalized recommendations may mean, on occasion, that the affiliated product is predicted to be the preferential product for the consumer.

Approaches which lean more into rules-based approaches (e.g., banning preferential positioning for affiliate products) may be less useful given these nuances of human behavior. More flexible principles—for instance, that vendors should position products in the interests of consumers and be able to demonstrate how they arrive at said determinations—offers some recourse to the challenges of regulating choice architecture. Some elements of performance monitoring may also be integrated. For instance, the broad principle that “settings should be easy to change,” could be established alongside a required disclosure from vendors demonstrating evidence that user ease has factored into choice architectural design choices. Failure to meet this standard may be inferred as failure to perform to the principle.

Within the behavioral science literature, several insights have emerged which may be proposed as principles for regulating choice architecture. The following are not definitive, with further relevant details discussed in Section 4.

Some behavioral auditing studies (e.g., Mills et al., 2023) have proposed that a service’s features should have what might be called symmetrical burdens. This may include it being as easier to sign up for an online service as it is to leave said service—a principle which aligns with the DSA (European Commission, 2022a). However, just because a service may be difficult to join does not mean it should be use as difficult to leave. Gambling, for instance, is an industry where advantages come from it being easier to close one’s account than to open an account (Behavioural Insights Team, 2022; Newall & Weiss-Cohen, 2022). Alternatively, some finance products, such as basic savings accounts, may better serve consumers by being easier to set up than close (FCA, 2022). The concept of symmetrical burden may thus be a worthwhile guide for regulators, but benefits from being treated as a principle with flexibility for given industries, rather than as a rule which may lead to peculiar regulatory positions.

Some work on the role of behavioral science in the private sector has developed the notion of “Pareto” and “rent-seeking” choice architecture (Beggs, 2016). The former *may* benefit the “nudger” (e.g., an online vendor), but should *always* endeavor to benefit the “nudgee” (e.g., a consumer or citizen). The latter *only* benefits the “nudger,” or if some benefit is afforded to the “nudgee,” this benefit is substantially less than that received by the “nudger” (Mills, 2023). This distinction between Pareto and rent-seeking choice architecture has been supported by some within the behavioral science literature (e.g., Newall, 2023) as it allows discussions to move beyond the somewhat arbitrary labeling of choice architecture as “good” or “bad,” and instead encourages one to investigate the affordances of both those being steered toward different choices, and those doing the steering. The notion that vendors should only ever use Pareto choice architecture struggles as a regulatory rule. For instance, one would need to establish what is considered a significant difference in benefits, or indeed, what is a benefit. Treating Pareto choice architecture as a principle would require vendors to demonstrate (a) the rationale for their measurement of benefit; and (b) evidence of this rationale directing choice architectural design. This offers flexibility, which is helpful to regulators and vendors alike, and would broadly align with some regulatory ambitions to foster transparency from (CMA, 2021) and understandability of (European Commission, 2022a, p. 57) online services.

A third principle comes from considering a specific divide between some dark patterns and choice architecture. Choice architecture should steer individuals toward different choices, without forcing an individual to choose a particular option (Sunstein, 2014; Thaler & Sunstein, 2008). However, some perspectives within the dark patterns literature suggest website designs which force an individual to engage in some activity, when said activity may not be strictly necessary, can be considered a dark pattern (e.g., Brignull, 2011; Gray et al., 2018; Mathur et al., 2019). These “forced action” dark patterns do not typically fall within the purview of behavioral science but may still be of substantial interest to regulators (Mills et al., 2023). Thus, a third principle may be a principle of explainable mandates. Specifically, that vendors can implement forced action at their discretion, but any forced actions implemented should be justifiable to a regulator (or, if so desired, an individual) upon request (again, see European Commission, 2022a).

4. Behavioral audits

A principles-based approach is not perfect. While both rules-based and principles-based approaches necessitate various disclosures from regulated organizations, particularly when including measures of performance (Willis, 2014), principles-based approaches are likely to result in a greater variety and complexity of disclosure. The “cost” of flexibility in a principles-based approach is the challenge of ensuring compliance given the ambiguity and complexity which principles may engender.

A potential solution is developing behavioral audits.¹² Auditing methodologies are increasingly popular within behavioral science as a tool for investigating choice architecture. In particular, “sludge audits” analyze burdensome choice architectural designs which prevent individuals from achieving desired objectives (Sunstein, 2022b). Sludge audits have been undertaken in various domains, including in gambling (Behavioural Insights Team, 2022; Newall & Rockloff, 2022), health care (Rockwell et al., 2023), public administration (e.g., Executive Office of the President of the United States, 2023), online retail (Luguri & Strahilevitz, 2021; Mills et al., 2023), and finance (Ontario Securities Commission, 2024). Various policy-facing organizations (New South Wales Behavioural Insights Unit, 2023; Varazzani et al., 2023) have called for wider adoption of sludge auditing approaches, which has also been echoed by regulators (e.g., CMA, 2021).

Some audit studies, focusing on dark patterns, adapt the sludge audit approach (Behavioural Insights Team, 2022; Luguri & Strahilevitz, 2021; Mills et al., 2023), and in doing so, extend auditing methodologies. For instance, various dark patterns make options *easier* to choose, not harder, as is often seen in the gambling industry (Newall, 2022). These auditing approaches typically combine sludge audits with various qualitative data and behavior change frameworks to scrutinize a wider range of behavioral outcomes, including burdens, but also behavioral “shortcuts” which make potentially harmful options easier for consumers to choose (Behavioural Insights Team, 2022; Mills et al., 2023). It is from this perspective that the notion of a behavioral audit is derived.

Developed in conjunction with regulators, behavioral audits could be an important tool for investigating compliance with regulatory principles. For instance, Mills et al. (2023) undertake a sludge audit of various online services and use the principle of symmetrical burden (what they call the “equal clicks principle”) to guide subsequent analyses of qualitative data.

Within behavioral science, many auditing approaches are also undertaken to offer positive recommendations for improving services (e.g., Rockwell et al., 2023). Thus, behavioral audits should also offer organizations insights into how to *improve* outcomes for themselves and their clients. Insights could also be guided by regulatory principles (e.g., that choice architecture should be Pareto, not rent seeking), and should be encouraged by regulators, especially given some deceptive choice architecture will not be intentionally so. As acknowledged by regulators (e.g., CMA, 2021, p. 5) and behavioral scientists (e.g., Sunstein, 2019b) alike, many poor interactions between individuals and services arise because of *careless*, or *accidental*, designs. Consumer welfare is “left on the table,” if behavioral audits are *only* used for compliance, and not to improve individual experiences and organizational processes.

The remainder of this section proposes several questions which should guide the development of behavioral audits. These questions are offered as an initial foundation for developing a program of research into behavioral auditing, combining behavioral science perspectives with regulatory positions and demands. Table 1 summarizes these various research questions.

4.1. How should data be collected?

Of existing auditing approaches, one may distinguish between “auditor-led” (Mills et al., 2023) and “checklist-led” (Behavioural Insights Team, 2022) approaches. Auditor-led approaches encourage auditors to experience services *as if* they were an ordinary user. Auditors are empowered to make records of their experiences, highlighting what *they* determine to be positive or negative uses of choice architecture. Auditor-led approaches may be more adaptive to a variety of domains and may more often capture an experience comparable to a typical user (Mills et al., 2023). The disadvantage of this approach is it can lead to discrepancies between teams of auditors. For instance, in their review of several large, online services, Mills et al. (2023) report inconsistencies across two audit reports for some services examined.

Table 1 A research program for behavioral audits

Research question	Relevance
How should data be collected?	Audit results may vary based on who is undertaking the audit, and how the audits are conducted. Regulators and behavioral scientists should collaborate to develop auditing methodologies and determine effective practices.
How much data should be collected?	The depth of a behavioral audit will be important in determining the resources and outcomes of an audit report. Depending on regulator objectives and available resources, high-level metrics may be sufficient, or more detailed audits may be warranted.
Who should audit?	Different regulators may have different capacities to undertake behavioral audits. Independent behavioral auditing organizations may have a conflict of interest which must be addressed within any regulatory regime.
What should be the outputs?	Regulatory principles will determine the outputs of behavioral audits. Practical aspects of behavioral auditing, such as the frequency of audits and the planned use of any generated outputs, will also influence what a behavioral audit should achieve. Furthermore, the question of measuring welfare effects of deceptive choice architecture is central to understanding the outputs of a behavioral audit.
What are the alternatives to audits?	Some (if not all) dark patterns and deceptive choice architecture will already be prohibited by existing consumer protection legislation. It may thus be prudent to address enforcement challenges rather than developing auditing procedures. Though, some auditing procedures may be part of an enforcement strategy.

Checklist-led approaches use pre-determined checklists to engage in a more systematic review of choice architecture. For instance, the Behavioural Insights Team (2022) use a checklist-led approach in their sludge audit of gambling websites in the UK. This approach may be preferable when examining a single domain, and/or when established regulatory principles exist from which to build a coherent checklist. Checklists can reduce the variation in audits by standardizing metrics, allowing for more comparable results. However, this approach may be limited due to a lack of flexibility. For instance, personalization may lead to different choice architectural experiences for different auditors, and a checklist-led approach may miss this subtle difference in auditor experience (Morozovaite, 2022).

Skill- and experience-level of the auditor is an important consideration highlighted by regulators and the literature (e.g., CMA, 2021). The DSA, for instance, assumes an “average” user within its discussion (European Commission, 2022a, p. 56). The CMA (2021, p. 30) suggests one approach to auditing may be to “enlist consumers to act as digital ‘mystery shoppers’” as these experiences are likely to be representative of typical user experiences. Sludge audits undertaken by the New South Wales Behavioural Insights Unit (2023) draw on customer experience surveys. Such proposals may be promising but re-emphasize the methodological question of auditor-led versus checklist-led approaches. For instance, a checklist-led approach using digital mystery shoppers may prime participants in ways which do not reflect their typical experiences; an auditor-led approach may result in digital mystery shoppers missing elements of choice architecture which are influencing them (e.g., FTC, 2022b, p. 3).

The question of how data should be collected is an important opportunity for cross-collaboration between behavioral scientists and regulators. As research into auditing methodologies develops, behavioral scientists have opportunities to contribute further research insights, while regulators may have a guiding hand in establishing specific choice architectural design practices which require research attention. The merits and demerits of different approaches necessarily hinge on collaborative discussions of what is possible given what is to be accomplished.

4.2. How much data should be collected?

Audits can vary in depth of detail. “High-level” approaches use auditing methods to identify broad problem areas and assist regulators in targeting limited resources (Mills et al., 2023). A “high-level” approach typically uses metrics which are readily accessible to an external party (e.g., a consumer) to gain an overview of critical processes.

For instance, Mills et al. (2023) investigate the customer experience of creating and deleting user accounts for various online services, plotting user journeys throughout this process and annotating areas of concern. Through this approach, while auditors may gain limited understanding of specific dark patterns or other deceptive strategies, they do gain an overall picture of the process, which can be used to better target further resources and identify areas warranting more in-depth investigation.

Yet, “high-level” approaches have limitations. First, depending on the regulatory standards set, the insights gleaned from high-level metrics may be insufficient to validate compliance, or to justify penalties for non-compliance. High-level approaches may best serve as *initial* audits pending a more detailed investigation. Second, high-level approaches are limited in the “behaviors” which can be investigated. Both Mills et al. (2023) and the Behavioural Insights Team (2022) investigate the behaviors around creating and deleting online accounts. Yet, the former only adopt a high-level approach, and note this limits their ability to comment on some choice architectural phenomena observed. Third, high-level approaches may lack the detail needed to offer positive recommendations to vendors.

The Behavioural Insights Team (2022) examine multiple user behaviors beyond high-level metrics, and thus comment upon aspects of the user experience in greater detail, as do Newall et al. (2022). More detailed auditing approaches likely can overcome the drawbacks of only taking a high-level view, though at the cost of more time and resources. Depending on the regulatory environment, these requirements may exceed regulatory capacities.

As with the question of how data should be collected, how much data is collected will ultimately depend on available expertise and regulatory requirements. For instance, if the principle of symmetrical burden is adopted, any audit seeking to validate compliance will need some data about design choices relating to reversible decisions (e.g., ease of subscribing/unsubscribing to a service). High-level approaches are ample for identifying where burdens are asymmetrical (Mills et al., 2023), but lend few insights when there may be a legitimate reason for asymmetry.

4.3. Who should audit?

The question of “who should audit?” is not concerned with the above discussions of which individuals (e.g., behavioral scientists or average consumers) should collect data, but with which organizations should be permitted to undertake behavioral audits.

Broadly, this question hinges on the capacities of specific regulators, and the nature of the markets which they regulate. For instance, in any given country, there are likely to be only a handful of retail banks, but several thousand online traders. A finance regulator, therefore, may be well positioned to undertake behavioral audits of the firms it regulates because there are few of them, though each may account for a significant market share. A competition regulator, by contrast, may struggle to undertake audits of all the firms which fall under its remit as there are too many.

Various proposals could resolve the challenge of having to audit many firms. For instance, the DMA has criteria for determining whether an online service is a gatekeeper. These criteria include the number of users a service has, and whether “it has a significant impact on the internal market” of a given industry (European Commission, 2022b, para. 6). Under the DMA, firms designated gatekeepers are subject to specific regulatory requirements. In 2024, only six firms have been designated gatekeepers. Applying an approach like this may reduce the number of firms subject to behavioral audits within a given market to only a few firms which have an outsized impact on the market, and thus consumers.

Another approach may be to engage independent behavioral consultancies to undertake behavioral audits. A sizeable and global behavioral insights industry is already established (DellaVigna & Linos, 2022), and is likely well-positioned to conduct behavioral audits in some sectors. Treating behavioral audits in a manner like financial audits may, therefore, allow widespread compliance with a principles-based regulatory regime to be achieved.

Nevertheless, the use of independent auditing organizations raises substantial challenges which must be met within a behavioral auditing program. Centrally, the challenge of conflicting interests. Firms equipped to provide behavioral auditing services are already likely to be professional behavioral consultancies whose business model is selling behavioral insights to other organizations. Independent auditors would thus be incentivized to find “problems” with a vendor’s choice architecture, as this would create an opportunity for further business.

4.4. What should be the outputs?

An essential question to ask is: what should be the outputs of a behavioral audit? As above, much will depend on the regulatory regime which is settled upon, including the principles established and practical considerations, such as the frequency of audits and the intended uses of audit reports.

If Pareto choice architecture is adopted as a guiding principle, a robust behavioral audit should provide evidence (a) outlining the firm's belief about the mutual benefit the choice architectural design produces; and (b) evidence that such benefit is received by users of the service. Thus, a behavioral audit should at the least investigate these claims and the evidence offered to support them and offer conclusions as to the validity of these claims in comparison with the spirit of regulatory principles.

If the principle of symmetrical burden is adopted, as above, there may be some industries (e.g., gambling) where asymmetrical burdens can be justified (Newall, 2022). This being so, a behavioral audit should scrutinize any justifications for asymmetries in choice architecture, rejecting them, or offering correctives, where appropriate.

The frequency of behavioral audits is a practical consideration which is also likely to influence the outputs of any behavioral auditing process. For instance, a behavioral audit may follow accounting practices and be conducted annually, though if no substantial changes to a vendor's choice architecture, product range, or algorithmic infrastructure have been made, this may be excessive (especially given regulatory resource constraints). Audits may be warranted whenever a substantial change in choice architecture is undertaken or may be imposed by regulators on an *ad hoc* basis. Yet another approach may follow from industries such as the commercial food industry, where periodic inspections are undertaken, resulting in a simple inspection rating.

Each approach may be influenced by a further practical implication, namely, the intended use of a behavioral audit report. When audits are undertaken as data gathering exercises, or as part of an investigation where penalties may result, greater detail is likely necessary. Where audits are undertaken for the purposes of informing the public as to the quality (in terms of behavioral manipulation) of a service, simpler and more intuitive outputs may be warranted.

Finally, and perhaps most importantly, is the practical consideration of how to measure any harm caused by deceptive choice architecture. As above, regulators such as the CMA (2021) have a broad framework for *understanding* harms, but directly measuring the economic harms caused by dark patterns remains an area for further development. Some (Conti & Sobiesk, 2010) have used subjective wellbeing measures to evaluate the impacts of dark patterns. Others (Luguri & Strahilevitz, 2021) have used attitudinal measures of acceptability, while others still (Behavioural Insights Team, 2022) have proposed measuring time costs, which could feasibly be transformed into an economic cost using some time-value of money standard. Measurements, such as measures of economic harm, are likely to be critical outputs for any behavioral audit, and developing such measurements will be a key objective of any behavioral auditing program. Indeed, without any means of measurement, it is disputable to what extent any recommendations or penalties for deceptive practices can be justified.

4.5. What are the alternatives to audits?

Related to any program of behavioral auditing must be the question of whether alternative approaches for achieving regulatory compliance exist. As Luguri and Strahilevitz (2021) note, many dark patterns and deceptive uses of choice architecture probably already violate various consumer protection laws in the United States. This being so, the proliferation of these deceptive practices may demonstrate not the need for a new approach (e.g., behavioral auditing) but the mere enforcement of *existing* laws and the exercising of *existing* powers. For instance, if regulators such as the FTC already receive many complaints from consumers about deceptive practices in a particular industry, or by a particular vendor, a behavioral audit is unlikely to reveal substantially new information (though it may form an evidential basis for taking further action; CMA, 2021).

Related to this point, Luguri and Strahilevitz (2021) suggest that a broad consensus within the dark patterns literature exists as to the common types of dark pattern strategies used. While, from a behavioral science perspective, individuals may experience choice architecture differently and thus subjectively, from a more legal perspective, some deceptive designs can be clearly determined and broadly regulated against. For instance, emerging perspectives on dark patterns regulation in India generally focus on outlining specific dark patterns which

proposed legislation will ban (Indian Department of Consumer Affairs, 2023). The importance of subjective experiences may, to a behavioral scientist, be a substantial challenge for regulation, but, to a legal analyst, be a consideration which is surmountable without the development of behavioral audits, or an adherence to a principles-based regime.

Despite these criticisms of the importance of subjectivity, Luguri and Strahilevitz (2021, p. 43) still call for the development of “dark patterns audits.” They emphasize that even if existing legislation (in the United States) likely already protects consumers from dark patterns, these same protections often require an external party to be able to audit consumer-facing processes to demonstrate wrongdoing. Whether such audits essentially constitute regulatory probes into violations of existing legislation, or should be developed alongside emerging regulatory perspectives on deceptive choice architecture and support a more far-reaching program, is an outstanding question to consider, though one which again invokes questions surrounding the depth of audits, and the purposes and practicalities of auditing.

5. Conclusion

As concerns around deceptive choice architecture in online spaces mount, it is essential for the applied behavioral science community to engage with this changing regulatory environment. This article reviews the regulatory positions of three large and influential jurisdictions: the United States, the UK, and the EU. Regulators share concerns about deceptive choice architecture, and the role of new technologies in exacerbating such designs.

Two challenges for regulating deceptive choice architecture have been outlined in this article. First, that subjective experiences of choice architecture demand regulatory flexibility, favoring a principles-based approach. Second, that behavioral science can offer positive recommendations to vendors, and this should be afforded within any regulatory regime, once more encouraging the adoption of a principles-based approach.

A principles-based approach will require strategies to ensure compliance. Collaborative efforts between regulators, legal experts, and behavioral scientists should be directed toward developing regulatory principles. While some preliminary principles have been suggested in this article, effective principles will be those that account for the capabilities and demands of various stakeholders.

Once principles are established, ensuring compliance with said principles represents an important challenge. To this end, this article has advocated developing behavioral audits as a regulatory tool. Building from insights and examples found in behavioral science, where auditing methodologies are emerging, this article has incorporated regulatory ambitions and presented a program for further development of behavioral audits. This program is unlikely to be definitive, either in relation to behavioral auditing, or in relation to regulatory challenges. For instance, as technology progresses and online choice architecture evolves, auditing approaches may have to adapt, and whether adaptation will be sufficient remains an outstanding question. Nevertheless, given current regulatory ambitions and emerging methodologies in behavioral science, a principles-based regulatory approach supported by behavioral auditing methods represents a promising approach to the effective regulation of dark patterns and other deceptive online choice architecture.

Acknowledgments

I am grateful to Till Stowasser for the opportunity to present a working draft of this article at the “Behavioural Science and Digital Markets Regulation” workshop hosted by the University of Strirling, and to all participants of that event. I am also grateful to Alex Clark of the Behavioural Insights Team Asia-Pacific for helpful discussions and comments.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Endnotes

- ¹ Generally, this article will focus on the effects on consumers, which are understood as direct harms (CMA, 2021; FTC, 2022b). Anti-competitive practices are considered indirect harms. Briefly, competition concerns around deceptive choice architecture are that these practices may force rivals to also adopt them to remain competitive, and that heightened consumer switching costs undermine competition through effective consumer choice (CMA, 2021; FTC, 2022a, 2022b). Also see Akerlof and Shiller (2015).
- ² A potentially relevant inclusion within the *Digital Markets Act* (European Commission, 2022b, para. 37) is: “Gatekeepers [e.g., large, online services] should not design, organize or operate their online interfaces in a way that deceives, manipulates or otherwise materially distorts or impairs the ability of end users to freely give consent.”
- ³ The European Data Protection Board (2022, p. 2) state that, “Data protection authorities are responsible for sanctioning the use of dark patterns if these breach GDPR requirements.” EU guidance on interpreting the UCPD states that, “the Directive [UCPD] has a broad scope of application... For example, the Directive would... cover commercial practices such as capturing the consumer’s attention, which results in transactional decisions such as continue to use the service [sic] (e.g., scrolling through a feed), to view advertising content or to click on a link” (European Commission, 2021a, p. 99–100).
- ⁴ The CMA (2021, p. 7) *does* also recognize that there may be a broader perspective on harms than just economic harms: “The potential harms from the use of algorithms are wide ranging and raise issues in multiple overlapping areas of public policy and regulation... We focus primarily on economic harms... However, we also note that the use of algorithmic systems can cause harms to people in their role as citizens, and not just their role as consumers or business owners.” Broadly, this article takes “harm” or “consumer welfare” to be related to economic outcomes, and to have some degree of measurability in terms of monetary values. Also see Section 4.
- ⁵ For instance, within the European Commission’s proposed “Data Act” it is argued that “Common and legitimate commercial practices that are in compliance with [European] Union law should not in themselves be regarded as constituting dark patterns” (European Commission, 2022c, para. 34).
- ⁶ While not a legal opinion, the wording of the DSA could lead one to conclude that it is against *all* nudging, for good or for ill. Generally, though, the DSA seems concerned specifically with nudging that benefits the nudger, but not the nudgee—what Mills (2023, p. 321) has called “rent-seeking” nudges. This use of language, however, is a common feature in some EU perspectives on dark patterns and deceptive choice architecture. For instance, the “Data Act” states “[dark patterns] deceive users by nudging them into decisions on data disclosure transactions or to unreasonably bias the decision-making of the users of the service, in a way that subverts or impairs their autonomy, decision-making and choice” (European Commission, 2022c, para. 34). Also see European Commission (2021a, p. 101): “Within the category of manipulative practices, the term ‘dark patterns’ is used to refer to a type of malicious nudging, generally incorporated into digital design interfaces. Dark patterns could be data-driven and personalized, or implemented on a more general basis, tapping into heuristics and behavioral biases, such as default effects or scarcity biases.”
- ⁷ The CMA (2021) discussion of new powers relates to the establishing of a new Digital Markets Unit (DMU) within the CMA, with the CMA having been asked by the UK Government to suggest what powers this new unit should have.
- ⁸ Specifically, the CMA (2021, p. 6) is concerned that the pace of AI development, coupled with its broad applications within markets, could allow the technology to cause irreparable market harm without *ex-ante* powers.
- ⁹ FTC (2022b, p. 3): “The specific types of dark patterns consumers are most likely to face differ depending on the types of websites or apps they frequently use. The medium through which consumers access online information also affects the number and types of dark patterns they may encounter... [D]ark patterns may have a differential impact on lower-income consumers or other vulnerable populations who are more likely to rely on a mobile device as their sole or primary access to the internet.”
- ¹⁰ Regulators emphasize the need for regulation to accommodate particularly vulnerable groups, including but not limited to children. See, for instance, the European Commission (2022a, p. 57).
- ¹¹ A common retort is to use, wherever possible, active choices (e.g., not offering defaults). This may, in some instances, be worthwhile (e.g., Sunstein, 2012). However, all regulators emphasize that the *rank-order* of choices should also be considered choice architecture (this is because recommendation algorithms purposely change rank orders). Thus, even adopting an active choice approach does not produce “neutral” choice architecture. One could go further still and *randomize* the order of choice each time. But at this point, one is designing an especially unnavigable choice environment, which *in itself* will influence decision-makers. Perhaps the only truly “neutral” choice architecture is that which relies on clairvoyance.

- ¹² Following the publication of a working draft of this article, though prior to final submission, the OECD's *Observatory of Public Sector Innovation* also launched a petition for the development of "behavioral audits" (Varazzani et al., 2023). No reference to the working draft is offered, and the petition focuses centrally on sludge audits in the public sector.

References

- Akerlof, G. A. (1970). The market for "Lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500.
- Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception*. Princeton University Press.
- Beggs, J. (2016). Private-sector nudging: The good, the bad, and the uncertain. In S. Abdulkadirov (Ed.), *Nudge theory in action*. Palgrave Macmillan.
- Behavioural Insights Team. (2022). *Behavioural risk audit of gambling operator platforms* [Online]. <https://www.bi.team/wp-content/uploads/2022/07/Behavioural-Risk-Audit-of-Gambling-OperatorPlatformsfindings-report-July-2022.pdf>
- Beshears, J., & Kosowsky, H. (2021). Nudging: Progress to date and future directions. *Organizational Behavior and Human Decision Processes*, 161, 3–19.
- Black, J., Hopper, M., & Band, C. (2007). Making a success of principles-based regulation. *Law and Financial Markets Review*, 1(3), 191–206.
- Bösch, C., Erb, B., Kargl, F., Kopp, H., & Pfatteicher, S. (2016). Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies*, 4, 237–254.
- Brignull, H. (2011). 'Dark patterns: Deception vs. honesty in UI design' A list apart [Online]. <https://alistapart.com/article/dark-patterns-deception-vs.-honesty-in-ui-design/>
- CMA. (2021). *Algorithms: How they can reduce competition and harm consumers*. Competition and Markets Authority. [Online]. <https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers>
- Coglianesse, C. (2017). The limits of performance-based regulation. *University of Michigan Journal of Law Reform*, 50, 525–563.
- Conti, G., & Sobieski, E. (2010). Malicious interface design: Exploiting the user. *Proceedings of the 19th International Conference on World Wide Web*, 4, 271–280. <https://doi.org/10.1145/1772690.1772719>
- Decker, C. (2018). 'Goals-based and rules-based approaches to regulation' BEIS Research Paper Number 8 [Online]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/714185/regulation-goals-rules-based-approaches.pdf
- Deese, B., Mahoney, N., & Wu, T. (2022). 'The president's initiative on junk fees and related pricing practices' The White House [Online]. <https://www.whitehouse.gov/briefing-room/blog/2022/10/26/the-presidents-initiative-on-junk-fees-and-related-pricing-practices/>
- DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1), 81–116.
- Department of Culture, Media and Sport. (2023). *High stakes: Gambling reform for the digital age* [Online]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1153228/1286-HH-E02769112-Gambling_White_Paper_Book_Accessible1.pdf
- Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, 47, 1–16.
- Dudley, S. E., & Xie, Z. (2022). Nudging the nudger: Toward a choice architecture for regulators. *Regulation and Governance*, 16, 261–273.
- European Commission. (2021a). *Guidance on the interpretation of directive 2005/29/EC*. The European Commission. [Online]. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021XC1229\(05\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021XC1229(05)&from=EN)
- European Commission. (2021b). *Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*. The European Commission. [Online]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>
- European Commission. (2022a). *Regulation of the European Parliament and of the council of... on a single market for digital services (Digital Services Act) and amending Directive 2000/31/EC*. The European Commission. [Online]. https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AL%3A2022%3A277%3ATOC&uri=uriserv%3AOL.L_2022.277.01.0001.01.ENG
- European Commission. (2022b). *Regulation (EU) 2022/1925 of the European Parliament and of the Council of... on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)*. The European Commission. [Online]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925the>
- European Commission. (2022c). *Proposal for a regulation of the European Parliament and of the council on harmonised rules on fair access to and use of data (Data Act)*. The European Commission. [Online]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN>
- European Data Protection Board. (2022). *Guidelines 3/2022 on dark patterns in social media platform interfaces: How to recognize and avoid them*. The European Data Protection Board. [Online]. https://edpb.europa.eu/system/files/2022-03/edpb_03-2022_guidelines_on_dark_patterns_in_social_media_platform_interfaces_en.pdf
- Executive Office of the President of the United States. (2023). *Tackling the time tax*. [Online]. <https://www.whitehouse.gov/wp-content/uploads/2023/07/OIRA-2023-Burden-Reduction-Report.pdf>

- FCA. (2021). *About the InvestSmart campaign*. The Financial Conduct Authority. [Online]. <https://www.fca.org.uk/investsmart/about-campaign>
- FCA. (2022). *A new consumer duty: Feedback to CP21/36 and final rules*. The Financial Conduct Authority. [Online]. <https://www.fca.org.uk/publication/policy/ps22-9.pdf>
- FTC. (2022a). *Statement of Chair Lina M. Khan regarding the advanced notice of proposed rulemaking on unfair or deceptive “junk” fees omission file No. R207011*. The Federal Trade Commission. [Online]. <https://www.ftc.gov/legal-library/browse/cases-proceedings/public-statements/statement-chair-lina-m-khan-regarding-advanced-notice-proposed-rulemaking-unfair-or-deceptive-junk>
- FTC. (2022b). *Bringing dark patterns to light*. The Federal Trade Commission. [Online]. https://www.ftc.gov/system/files/ftc_gov/pdf/P214800%20Dark%20Patterns%20Report%209.14.2022%20-%20FINAL.pdf
- Gilad, S. (2010). It runs in the family: Meta-regulation and its siblings. *Regulation and Governance*, 4(4), 485–506.
- Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. (2018). The dark (patterns) side of UX design. *CHI*, 2018, 1–14.
- Hallsworth, M. (2023). A manifesto for applying behavioral science. *Nature Human Behaviour*, 7, 310–322.
- Hallsworth, M., & Kirkman, E. (2020). *Behavioral insights*. MIT Press.
- Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. W. H. Allen, UK.
- ICO, & CMA. (2023). *Harmful design in digital markets: How Online Choice Architecture practices can undermine consumer choice and control over personal information*. Information Commissioner’s Office and the Competition and Markets Authority. [Online]. https://www.drcc.org.uk/__data/assets/pdf_file/0024/266226/Harmful-Design-in-Digital-Markets-ICO-CMA-joint-position-paper.pdf
- Indian Department of Consumer Affairs. (2023). *Proposed issue of guidelines on prevention and regulation of dark patterns*. Department of Consumer Affairs. [Online]. <https://consumeraffairs.nic.in/sites/default/files/file-uploads/latestnews/Draft%20Guidelines%20for%20Prevention%20and%20Regulation%20of%20Dark%20Patterns%202023.pdf>
- Jachimowicz, J. M., Duncan, S., Weber, E. U., & Johnson, E. J. (2019). When and why defaults influence decisions: A meta-analysis of default effects. *Behavioural Public Policy*, 3(2), 159–186.
- Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizen versus the internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest*, 21(3), 103–156.
- Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, 13(1), 43–109.
- Martin, L., Delaney, L., & Doyle, O. (2023). Everyday administrative burdens and inequality. *Public Administration Review*, 1–14. <https://doi.org/10.1111/puar.13709>
- Mathur, A., Acar, G., Friedman, M. J., Lucherini, E., Mayer, J., Chetty, M., & Narayanan, A. (2019). Dark patterns at scale: Findings from a crawl of 11K shopping websites. *Proceeding on the ACM Human-Computer Interactions*, 2(81), 1–32.
- Mills, S. (2022). Personalized nudging. *Behavioural Public Policy*, 6(1), 150–159.
- Mills, S. (2023). Nudge/sludge symmetry: On the relationship between nudge and sludge and the resulting ontological, normative and transparency implications. *Behavioural Public Policy*, 7(2), 309–332.
- Mills, S., & Whittle, R. (2023). Seeing the nudge from the trees: The 4S framework for evaluating nudges. *Public Administration*, 1–21. <https://doi.org/10.1111/padm.12941>
- Mills, S., Whittle, R., Ahmed, R., Walsh, T., & Wessel, M. (2023). Dark patterns and sludge audits: An integrated approach. *Behavioural Public Policy*, 1–27. <https://doi.org/10.1017/bpp.2023.24>
- Morozovait, V. (2022). Hypernudging in the changing European regulatory landscape for digital markets. *Policy and Internet*, 15(1), 78–99.
- New South Wales Behavioural Insights Unit. (2023). *Making government easier by eliminating ‘sludge’*. [Online]. <https://www.nsw.gov.au/behavioural-insights-unit/blog/making-government-easier-by-eliminating-sludge>
- Newall, P. W. S. (2022). Reduce the speed and ease of online gambling in order to prevent harm. *Addiction*, 118(2), 204–205.
- Newall, P. W. S. (2023). What is sludge? Comparing Sunstein’s definition to others. *Behavioural Public Policy*, 7(3), 851–857.
- Newall, P. W. S., & Rockloff, M. J. (2022). Promoting safer gambling via the removal of harmful sludge: A view on how behavioural science’s ‘nudge’ concept relates to online gambling. *Addiction*, 117(3), 838–839.
- Newall, P. W. S., Walasek, L., Ludvig, E. A., & Rockloff, M. J. (2022). Nudge versus sludge in gambling warning labels: How the effectiveness of a consumer protection measure can be undermined. *Behavioral Science and Policy*, 8(1), 17–23.
- Newall, P. W. S., & Weiss-Cohen, L. (2022). The Gambification of investing: How a new generation of investors is being born to lose. *International Journal of Environmental Research and Public Health*, 19(9), e5391.
- OECD. (2022). *Dark commercial patterns*. OECD Digital Economy Papers no. 336. [Online]. <https://www.oecdilibrary.org/docserver/44f5e846-en.pdf?expires=1666811516&id=id&accname=guest&checksum=141DA43B1CA0CFA24217B28948A1B004>
- Ontario Securities Commission. (2024). *Digital engagement practices: Dark patterns in retail investing*. [Online]. https://www.osc.ca/sites/default/files/2024-02/inv-research_20240223_dark-patterns.pdf
- Page, L. C., Castleman, B. L., & Meyer, K. (2019). Customized nudging to improve FAFSA completion and income verification. *Educational Evaluation and Policy Analysis*, 42(1), 3–21.
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770–780.
- Rawls, J. (1971). *A theory of justice*. Belknap Press.
- Rebonato, R. (2014). A critical assessment of libertarian paternalism. *Journal of Consumer Policy*, 37, 357–396.
- Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? *Judgment and Decision Making*, 11(4), 310–325.
- Rockwell, M. S., Frazier, M. C., Stein, J. S., Dulaney, K. A., Parker, S. H., Davis, G. C., Rockwell, J. A., Castleman, B. L., Sunstein, C. R., & Epling, J. W. (2023). A ‘Sludge Audit’ for health system colorectal cancer screening services. *The American Journal of Managed Care*, 29(7), 1–30.

- Roozenbeek, J., Freeman, A. L. J., & van der Linden, S. (2021). How accurate are accuracy-nudge interventions? A preregistered direct replication of Pennycook et al. (2020). *Psychological Science*, 32(7), 1169–1178.
- Sanders, M., Snijders, V., & Hallsworth, M. (2018). Behavioural science and policy: Where are we now and where are we going? *Behavioural Public Policy*, 2(2), 144–167.
- Shahab, S., & Lades, L. K. (2021). Sludge and transaction costs. *Behavioural Public Policy*, 1–22. <https://doi.org/10.1017/bpp.2021.12>
- Sin, R., Harris, T., Nilsson, S., & Beck, T. (2022). Dark patterns in online shopping: Do they work and can nudges help mitigate impulse buying? *Behavioural Public Policy*, 1–27. <https://doi.org/10.1017/bpp.2022.11>
- Sugg, O., & Lesic, V. (2022). *Online choice architecture – How do we end up making decisions we don't want to make?* Competition and Markets Authority. [Online]. <https://competitionandmarkets.blog.gov.uk/2022/04/07/online-choice-architecture-how-do-we-end-up-making-decisions-we-dont-want/>
- Sunstein, C. R. (1991). Administrative substance. *Duke Law Journal*, 1991(3), 607–646.
- Sunstein, C. R. (1995). Problems with rules. *California Law Review*, 83(4), 953–1026.
- Sunstein, C. R. (2012). *Impersonal default rules vs. active choices vs. personalized default rules: A triptych*. [Online]. https://dash.harvard.edu/bitstream/handle/1/9876090/decidingbydefault11_5.pdf?sequence=1
- Sunstein, C. R. (2014). *Why nudge? The politics of libertarian paternalism*. Yale University Press.
- Sunstein, C. R. (2017). *Misconceptions about nudges*. SSRN. [Online]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3033101
- Sunstein, C. R. (2019a). Nudging: A very short guide. *Business Economics*, 54, 127–129.
- Sunstein, C. R. (2019b). Sludge and ordeals. *Duke Law Journal*, 68, 1843–1883.
- Sunstein, C. R. (2021a). *Sludge: Bureaucratic burdens and why we should eliminate them*. MIT Press.
- Sunstein, C. R. (2021b). Viewpoint: Are food labels good? *Food Policy*, 99, e101984.
- Sunstein, C. R. (2022a). The distributional effects of nudges. *Nature Human Behaviour*, 6, 9–10.
- Sunstein, C. R. (2022b). Sludge audits. *Behavioural Public Policy*, 6(4), 654–673.
- Sunstein, C. R., & Gosset, J. L. (2020). Optimal sludge? The price of program integrity. *Duke Law Journal Online*, 70, 74–90.
- Sunstein, C. R., & Reisch, L. A. (2014). Automatically green: Behavioral economics and environmental protection. *Harvard Environmental Law Review*, 38(1), 127–158.
- Sunstein, C. R., Reisch, L. A., & Rauber, J. (2017). A worldwide consensus on nudging? Not quite, but almost. *Regulation and Governance*, 12(1), 3–22.
- Tannenbaum, D., Fox, C. R., & Rogers, T. (2017). On the misplaced politics of behavioural policy interventions. *Nature Human Behaviour*, 1, 1–7.
- Thaler, R. H. (2018). Nudge, not sludge. *Science*, 361, 431–432.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth and happiness*. UK.
- Thaler, R. H., Sunstein, C. R., & Balz, J. P. (2012). Choice architecture. In E. Shafir (Ed.), *The behavioral foundations of public policy*. Princeton University Press.
- Thunström, L. (2019). Welfare effects of nudges: The emotional tax of calorie menu labeling. *Judgment and Decision making*, 14(1), 11–25.
- Thunström, L., Gilbert, B., & Jones Ritten, C. (2018). Nudges that hurt those already hurting – Distributional and unintended effects of salience nudges. *Journal of Economic Behavior and Organization*, 153, 267–282.
- Varazzani, C., Hubble, C., Galassi, A., Koromilas, E., Chan, S., & Trudinger, D. (2023). *Spotting the sludge: Behavioural audits to improve public services across countries*. OECD Observatory of Public Sector Innovation. [Online]. <https://oecd-opsi.org/blog/behavioural-science-spotting-the-sludge/>
- Waldman, A. E. (2020). Cognitive biases, dark patterns, and the ‘privacy paradox’. *Current Opinions in Psychology*, 31, 105–109.
- Willis, L. E. (2014). Performance-based consumer law. *University of Chicago Law Review*, 82, 1309–1409.