Nudging and Social Media:
The Choice Architecture of Online Life

Douglas R. Campbell
(Alma College, Michigan, USA)
campbelldr@alma.edu

Title: Nudging and Social Media: The Choice Architecture of Online Life

Abstract: In this paper, I consider the way that social-media companies nudge us to spend more time on their platforms, and I argue that, in principle, these nudges are morally permissible: they are not manipulative and do not violate any obvious moral rules. The moral problem, I argue, is not with nudging in principle but is instead with the fact that users are being nudged towards something bad for them. In practice, this often involves being nudged to spend an unhealthy amount of time using a social-media app or being nudged towards content that is bad for us, such as by promoting eating-disorder content to young girls. Since nudging is morally permissible, it is open to these companies to use the same technologies to nudge us towards the good.

Key words: social media; nudging; choice architecture; libertarianism; paternalism.

1. *Nudging and choice architecture, online and offline*

To nudge is to change the context in which a choice is made in order to promote the likelihood of a certain outcome. The context in which the choice is made is usually called the *choice architecture*, and the one doing the nudging is usually called the *choice architect*. Nudges are intended to be a gentle form of influencing someone's decision, so gentle, in fact, that it is consistent with the nudged person's freedom of choice. Accordingly, a nudge is a tool whose use is defended by so-called *libertarian paternalists*. These thinkers aim to promote good outcomes for others, a mission that we tend to associate with conventional paternalists, but with an eye on preserving freedom of choice, a consideration that we do not typically associate with paternalists but instead with libertarians. The goal of this paper is not to defend libertarian paternalism or to illuminate the nature of nudging or manipulation.[1]

---

[1] There is not enough space in this paper to consider high-level objections to libertarian paternalism. Mitchell 2005 and Veetil 2011 argue that libertarian paternalism is an oxymoron, rebutting Sunstein and Thaler 2003 who maintain the opposite; Gigerenzer 2015 doubts that the psychologi-

Douglas R. Campbell

This paper aims at understanding choice architecture of online life from within a libertarian-paternalist framework: specifically, the way that social-media companies such as Facebook, Instagram, Twitter, and YouTube nudge users to spend more time on their platforms and nudge users towards certain kinds of content, mostly with the aim of ensuring that users enjoy the content and then spend more time on their platforms. Social-media companies nudge users for the sake of increasing their own revenue from advertising. For instance, if YouTube can nudge a user to watch more videos, then the user has watched more advertisements. Along similar lines, a user might be nudged to reveal more information about themselves, which the company might then use to show them more-appropriate advertisements, thus increasing the probability that the user will click on the advertisement. The result is that the company could sell advertising space for a more lucrative fee.

In this paper, I shall argue that there is no compelling reason to think that it is wrong for social-media companies to employ nudges in principle. However, I shall argue that it is wrong for social-media companies to nudge users when users are being nudged towards something that is bad for them. Indeed, as we shall see, advocates of nudging build into their conception of it that it is an influence for *improving* people's lives and decisions, while making it a sufficiently gentle form of influence that people who disagree about whether their lives are being improved can easily defy the choice architect. I shall maintain that social-media companies are generally being sufficiently gentle, but they are not promoting good outcomes, nor are they even sufficiently well-intentioned.

I shall begin by illustrating nudges in offline contexts. One form of nudging is to make the default choice be whatever the choice architect wants as the promoted outcome.[2] A powerful example is to make organ donation opt-*out* instead of opt-*in*.[3] This small change has profound effects on how many people donate their organs. Another form of nudging is to change the environment. For

---

cal evidence for the view, in fact, supports it. At least as far as the philosophical support goes, see Sunstein 2012 and 2015 for the development of the theoretical framework against these objections. This paper is more about applying the view to help us understand moral issues online than it is defending it against high-level objections. By the same token, I do not intend to rebut the views that compete with libertarian paternalism (e.g., coercive paternalism). Only at this end of this paper do I maintain that it is better, all things being equal, to nudge when coercion can be avoided, but even then, I leave it open that if nudges do not appreciably promote well-being in the cases being considered below, then it might be appropriate to consider coercion.

[2] In fact, this might be the most popular form of nudging. See Halpern et al. 2007 for a discussion of this in health-care and policy-making contexts. It is also the form of nudging most frequently recommended by Sunstein and Thaler 2008, probably because it most obviously preserves freedom of choice and is remarkably effective, especially when the one being nudged lacks strong preferences (cf. Halpern et al. 2007: 1341).

[3] This example comes from Sunstein and Thaler 2008: 177ff. They argue that by presuming consent for cadaveric organ donation, two similar countries, Austria and Germany, achieved starkly different levels of consent rates: Austria presumed consent and achieved a near-universal level of consent (i.e., 99 percent), whereas Germany did not presume consent and achieved only a 12-percent consent rate.

instance, consider the example of a cafeteria, where the healthy food is placed at eye level and at the front of the set-up, where it is most visible and accessible; the unhealthy food is virtually just as accessible, but perhaps cafeteria-goers would have to reach down to get it. In this case, one is being nudged towards eating healthily, but someone who prefers to eat unhealthily is hardly even being inconvenienced. The last form of nudging we should consider is that of merely informing people. Consider an experiment done by Montana as it tried to get undergraduate students to drink less. Behavioral psychologists had done research showing that the amount of alcohol that people tend to consume is a function of how much alcohol they tend to think that other people consume. Montana rolled out an education program, informing undergraduate students of how much their peers, in fact, drink. Since the amount was lower than what many had previously thought, the amount that they drank decreased accordingly.[4]

It is crucial for the libertarian paternalist that people are being nudged towards what they already think is good for them, at least by and large. The libertarian paternalist prefers to nudge whereas others might prefer to coerce.[5] Those who disagree can easily opt out, for instance, of organ donation. Much research has been done on people's preferences to donate their organs; it is easy to see from the research that people have struggled to live in accordance with their preferences.[6] Similarly, there is no plausible way that it could be construed that undergraduate students in Montana were being treated objectionably by being informed of something true.

The example of the cafeteria illustrates something profound about nudging: specifically, that it is often unavoidable. The choice architecture of the cafeteria must be designed in *some* way. There has to be some arrangement of the foods, and whatever arrangement there is will nudge hungry people towards this or that option. The nudge might be towards what is healthy for them or what is most

---

[4] This example comes from Sunstein and Thaler 2008: 67-68. See Perkins 2003:7-9, Linkenbach 2003, Linkenbach and Perkins 2003, and Weschsler et al. 2000 for more details. The gist is that Harvard School of Public Health found that 44 percent of college students engaged in binge-drinking in the two-week period before the study. Lurking behind this behavior is the belief that alcohol abuse is common, more common than it really is. Montana tried to solve this problem by alerting students to the reality that people drink alcohol far less than the undergraduate students perceived. The educational campaign also included information about the popularity of smoking (e.g., "most (70 percent) of Montana teens are tobacco free"). By informing people, Montana successfully nudged many away from smoking and drinking.

[5] See, e.g., Conly 2013, who uses the same argumentative moves as libertarian paternalists to defend coercive paternalism as a means of helping people live in accordance with their own values.

[6] This is called the *transplant paradox*. See Kurtz and Saks 1996 for their illustrative study of the situation in Iowa at the time. They found that despite overwhelming support for organ donation in Iowa, not many people had, in fact, signed their driver's license to indicate that they would donate their organs. 97% of participants supported organ donation, but of that 97%, not even half had signed. Participants were also asked how strongly they wanted to be organ donors; of those who responded favourably in the strongest possible terms, only 64% had signed their driver's license; only about one-third had signed the organ-donation card that is required for making an anatomical gift in Iowa.

profitable for the cafeteria-owners, but, either way, there has to be some nudging. However, this is not the case with every instance of nudge. The Montana state government did not have to offer any information about alcohol consumption, and it could hardly be said that by withholding this information, it was nudging people to drink more. Nudges are *positive contributions*; they are influences. We might say that someone's ignorance of alcohol-consumption statistics contributed to their decision to drink a certain amount, but if Montana not informing someone about these statistics counts as a nudge, then, by the same token, parents and neighbors similarly nudged the drinkers by not informing them. In cases where nudges are not unavoidable, the libertarian paternalist thinks that they are justified on the grounds that they promote a person's already-existing values. Even if they do not promote those values, the nudges can be defied with no meaningful cost.

Nudging is a familiar part of the online landscape, just as it is of the offline one. We know, for instance, that Facebook changed the color of its notification that a user ought to check out some new post or comment from blue to red on account of how much more effective red is at ensuring that users click on the notification.[7] Users are being nudged to see what new piece of content Facebook suspects that they would like. Notifications themselves are nudges, in the same sense that simply being informed about something true can count as a nudge. Colors are a powerful feature of the choice architecture in online life: we also know that if we change the color of our smartphone screens such that the display is only in black and white, we spend less time on them; we find it attractive to look at multicolored displays.[8] The way that a social-media company recommends us content that we would like counts as a nudge. The promotion of some content as popular or "trending" similarly nudges us towards, analogously to the way that doctors might nudge a patient towards accepting the course of treatment that they list first.[9]

We might also think of the way that social-media sites such as Reddit and Twitter are designed these days with the "endless scroll" in mind. Reddit, for instance, was not always this way. It used to be that, on Reddit, one would scroll to the bottom of the page, run out of content on that page, and then manually

---

[7] See Newport 2019: 18-20 for an interesting discussion of this. The information comes from whistleblower Tristan Harris, who worked in the industry as a so-called design ethicist before leaving and drawing attention to the harms that he perceived in social-media landscape: see Paul Lewis, "'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia," *The Guardian* October 6, 2017 https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia Accessed October 19, 2021.

[8] This technique of reducing screentime has been popular for some time now, but Holte and Ferraro 2020's research confirms that this technique significantly reduces time spent on social media and Internet browsing, though not time spent on video-watching. An interesting similar study by Holte, Giesen and Ferraro 2021 found that the same technique also reduced anxiety and the problematic smartphone use that is correlated with a decline in productivity, sleep difficulties, risky driving, and more.

[9] See Cohen 2013: 5-6 and Simon 1956 for more on this.

click to move to the next page.[10] Nowadays, users never experience that feeling of running out of content on any given page, because one can scroll down endlessly as the app or site loads more content for the user.[11]

The goal is for someone to spend as much time on the platform as possible, while respecting a person's freedom of choice.[12] When Netflix began as a streaming service, the term 'binge-watch' did not exist, for all intents and purposes, in English, and when a user finished watching an episode or movie, the default option was that the user would not continue watching; he or she had to opt in to watching another. However, in 2012, Netflix changed the choice architecture such that the default option was that the user would watch more content; this is Netflix's *auto-loading* feature, called *post-play*. The term 'binge-watch' sprung into existence not long after.[13] Such was the power of the choice architecture.

This strategy is replicated on social-media sites. For instance, on YouTube, the default option is for videos to play when one is completed, and YouTube has constructed an endless supply of videos for users to move through. YouTube, unlike Netflix, has a much more immediate financial incentive for this: YouTube wants users to watch as many videos as possible because the more videos, the more advertisements one watches. Generally, the same reason explains why many non-social-media sites prefer to create their content as videos or slideshows instead of the written word: they take longer to get through than a page that a reader could quickly skim; the longer one is on a page, the more time that one is exposed to ads, and the slideshow format itself allows for new ads on each slide of the slideshow.

The use of the default-option nudge is powerful, as we can observe from the behavior of Netflix users. We encounter this type of nudge most frequently on

---

[10] This is the same behavior that leads people to not go beyond the first page of Google's search results. We cannot be sure how often people do this because that knowledge is proprietary, but van Deursen and van Dijk 2009 found that 91% of participants did not go beyond the first page of search results, and 36% did not go beyond the first three results. This is the online equivalent of a patient privileging the option that a doctor lists first.

[11] Pagination is a powerful *stopping rule.* Stopping rules guide our behavior by prompting us to stop doing what we are doing. It is easy to use technology to disrupt these stopping rules: Alter 2017: 188 points out how the use of credit cards instead of cash disrupts the natural rule to stop buying when the wad of cash in our wallet visibly dwindles. (See Prelec and Simester 2001 for a study finding that shoppers will pay up to *double* for the same item with a credit card than how much they would with cash.) Alter 2017: 184-190 discusses this problem widely with references to video games, new gambling technology, and so on.

[12] Consider what Sean Parker, the founding president of Facebook, said in 2017 at an event discussing Facebook's attention-engineering: "the thought process that went into building these applications, Facebook being the first of them, […] was all about: 'how do we consume as much of your time and conscious attention as possible?' And that means that we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo or a post of whatever" (Newport 2019:19).

[13] See Alter 2017: 209ff for the data and a discussion of it. Alter is aware that the popularity of Google searches is only an indirect measurement, but he also cites some interesting research that Netflix conducted internally, showing how influential the post-play feature is.

the Internet when we are signing up for an account at some site, and we are asked whether we would like to receive advertisements and information about promotions from that site. Sometimes we are asked to opt *into* these emails, whereas other times we are asked to opt out of these emails. Sometimes we are asked to opt into them by clicking on a box, but the box is already automatically clicked, and we instead to click to un-fill the box. One might wonder how it could be otherwise: there might not seem to be an alternative to either approach; either we opt in or opt out of such emails. However, sites could, and sometimes do, implement the mandated-choice nudge. Here, users are *required* to make a choice. One cannot move on from that stage of account-creation until one has selected 'yes' or 'no' in response to the question about receiving emails from the site. There would be no default option. This counts as a nudge because it promotes whatever outcome tends to be likeliest among users upon reflection.

When signing up for an account on a site that features an online store, opting in to receiving emails allows the site to nudge us regularly. Advertising promotions certainly counts as a nudge, and so do emails that remind us that we were looking at some product, and maybe even added it to our online shopping cart, but then left the site without buying anything. There is another incentive for such e-commerce sites to do the same thing, too: they want to leverage the so-called *buy-in effect* of signing up for a site.[14] The clearest illustration of this effect is when we sign up for an account in order to receive some discount at the site. The site could very well have simply applied the discount automatically at check-out instead. One might think that it amounts to the same, but it does not. Marketers are sufficiently aware that consumers are more likely to *do more* once they have already done *a little*. This is the buy-in effect: once we have already bought into something, we are more likely to buy into more later on. Once we have signed up for an account, we are more likely to spend money at that store.[15]

These are some of the clearest examples of nudging in online life. Most of these examples concern specifically social-media platforms, but others are more general, relating to the way that we shop online.

---

[14] Another example from e-commerce: seven-day free trials that do not end after seven days but renew and automatically become paid memberships. Since users are free to end their membership before it becomes paid, freedom of choice is preserved, but e-commerce businesses know that users are more likely to become paying customers this way than if their memberships automatically ended after the seven-day free trial.

[15] Similarly, there is the so-called endowment effect. The idea is that people are more likely to value something more when they own than if they did not own it. For instance, if Dropbox gives you two gigabytes of extra space for free, but then three months later starts charging you for it, you are more likely to spend the money to retain the two gigabytes than you would be to have paid for them in the first place. You will pay more money to retain something than to obtain it because the mere ownership makes it more valuable in your eyes. This can be used handily to nudge someone: give them something for free with the intention of charging them later on to keep it. See Morewedge and Gilbin 2015 for a study of why.

2. *Online nudges are generally morally permissible, in principle*

The ethical question is whether there is something wrong with using or designing the choice architecture this way. The question is whether it is wrong to nudge. In what follows, I shall consider some possible objections to the nudges and argue that, for the most part, these objections fail. I shall then argue that there is nothing wrong with these nudges in principle and, instead, the problem consists in what we are being nudged towards.

There are a couple of central constraints on what a choice architect can do, according to libertarian paternalists. The first of these constraints is that the nudge must comply with the *publicity principle*.[16] This principle requires that the choice architect do only what he or she would be willing to defend publicly to those nudged. This principle has practical value: it saves the choice architect from possible embarrassment, for instance. Secondly, it embodies the respect that people are meant to have for their fellow citizens; we avoid using people as tools and manipulating them.

It seems to me that most nudges in online contexts comply with this principle. There are a couple of exceptions that are better discussed later on. For the most part, though, few people, I suspect, would find themselves put off by learning that Facebook's notifications are meant to nudge them to spend more time on the platform. In fact, these nudges are generally so transparent already that it is almost as if social-media companies have already successfully defended them to the public: they have engineered the choice architecture right in front of our eyes, not in secrecy. We might find it annoying when a site opts us in to receiving emails by default, but it is implausible that this is so egregious that the choice architect would be unwilling to defend it publicly; indeed, what is going on here is already transparent.

The second constraint is that nudges must be sufficiently libertarian. That is, they need to preserve freedom of choice. There are, in most cases, *some* costs to defying the choice architect and to saying 'no' the outcome that one is being nudged towards. Perhaps the only exception to this is the mandated-choice nudge, where one is required to make a choice: here, either decision has the same cost. With respect to other nudges, it is important for the nudge to involve a cost of defiance as small as possible in order to still register as *libertarian*. In other words, a cafeteria where the healthy foods are at eye level and the unhealthy foods are just below eye level is designed unobjectionably. However, if the unhealthy foods have had their prices increased by ten dollars, that is less libertarian and more coercive. It is hard to specify the exact point at which a nudge stops being gentle and starts being coercive. Consider the cafeteria again, this time with the unhealthy food placed in separate room down the hallway. The cost of defying the choice architect is clearly higher,

---

[16] See Sunstein and Thaler 2008: 243ff. They owe this principle to Rawls 1971.

but this is not qualitatively different from simply putting them lower on the shelf, where one has to reach down. The cost is non-zero in both cases, but it is perhaps still trivial when reaching down and perhaps non-trivial when moving to a separate room.

That being said, I maintain that there is no need to pursue this line of thinking with respect to nudges online. That is because none of the nudges that we have surveyed so far, as a matter of fact, have an obviously high cost of defiance at all. It is easy to *not* click on a Facebook notification. It is easy, even if annoying, to opt out of receiving emails. It might be a more pleasant online experience to have YouTube give us more time between finishing a video and the automated start of the next one, but it is not difficult to defy this nudge, and it is rather easy to turn off this feature altogether. There are times when social-media sites might make things objectionably difficult. Much has made been of how difficult it is to delete one's Facebook account, for instance.[17] This is a kind of nudge, a desperate nudge to keep us on the site.[18] This might be equivalent of making a citizen have to fill out a pile of paperwork to opt out of an organ-donation program: the cost of saying 'no' might be objectionably high. This point is well-taken: certainly, some nudges can and often do cross the line and cease being libertarian. I do not, however, think that this is the case with most nudges online, where our freedom of choice is preserved.

Nudging is often perceived to be manipulative. When David Cameron, Prime Minister of the United Kingdom at the time, made a change to the British behavioral-insights team of psychologists and economists, *The Guardian* ran the headline "Nudge nudge, say no more. Brits' minds will be controlled without us knowing it," reflecting this fear.[19] The intuition that nudges are manipulative is powerful, and this intuition might be strengthened when we consider how effective, say, Netflix has been at getting viewers to watch more content by re-designing the choice architecture of its site. However, there is reason for thinking that this intuition is not trust-worthy. An interesting psychological study researching people's attitudes towards nudges in political contexts revealed how one's acceptance of a nudge depends on whether the nudge is explained using

[17] Specifically, the fact that Facebook makes users wait weeks before the account can be deleted and makes it hard to ensure that *all* of one's data is deleted (e.g., including conversations with friends). Deleting one's account correctly is unwieldy enough to warrant the creation of sites dedicated to guide one through it; see Deletefacebook.com.
[18] I encountered something like this first-hand while researching for this paper. I had to subscribe to the *Wall Street Journal* to read some reporting about Facebook's effects on young girls, and I discovered that I could not cancel my subscription by any means other than calling the *Wall Street Journal* (unless I lived in one of four states that require them to provide some other means). During the phone call, the customer-service agent tried to keep me as a paying customer. This might not be outright coercion, but it is not sufficiently easy to defy the choice architect that this still counts as an acceptable nudge.
[19] See Ian Dunt, "Nudge nudge, say no more. Brits' minds will be controlled without us knowing it," *The Guardian* February 5, 2014 https://www.theguardian.com/commentisfree/2014/feb/05/nudge-say-no-more-behavioural-insights-team Accessed October 20, 2021.

a policy that one supports or does not support.[20] For instance, conservatives disapproved of nudges when their usefulness was illustrated with a liberal policy objective, such as encouraging eligible people to enrol in food-stamp programs; liberals disapproved when nudges were illustrated with a conservative policy objective, such as encouraging eligible rich people to take advantage of some relief from the capital-gains tax. The same effect was observed when the participants were not voters but were sitting United States mayors.[21] This might explain why, whereas left-leaning *The Guardian* ran a headline worrying about nudging when a Conservative prime minister was in charge, right-leaning *The Daily Caller* ran a similarly mistrustful headline when Barack Obama was president and considering nudging: "President Obama Orders Behavioral Experiments on American Public."[22] This is some evidence that our intuitions about nudging are not reliable and can easily change with how nudges are framed.

Let us consider the charge that nudges are manipulative. One immediate problem is the difficulty of clarifying what manipulation is. For instance, manipulation might be characterized in the following way: "manipulation, unlike coercion, does not interfere with a person's options. Instead it perverts the way that person reaches decisions, forms preferences or adopts goals."[23] Consider also: "manipulation, broadly conceived, can perhaps be understood as intentionally causing or encouraging people to make the decisions one wants them to make by actively promoting their making the decisions in ways that rational persons would not want to make their decisions."[24] One common denominator is interference with our reasoning, and this interference appears to be gentler than coercion but also objectionable in ways that persuasion usually or always is not.[25] I shall argue that what we see in the choice architecture of social-media sites cannot be characterized as manipulation.

Firstly, many of the nudges that we see online and offline are simply inevitable in a morally relevant sense. Consider the cafeteria case. Cafeterias need to be organized in *some* way; this kind of nudge is *unavoidable*, as I explained earlier.

---

[20] For the research, see Tannenbaum et al. 2017. Two of the co-authors also discuss their findings in the *New York Times*; see Fox and Tannenbaum, "The Curious Politics of the 'Nudge'", *The New York Times* September 26, 2015 https://www.nytimes.com/2015/09/27/opinion/sunday/the-curious-politics-of-the-nudge.html Accessed October 18, 2021.

[21] Tannenbaum et al. 2017 found that libertarians tended to be more generally resistant to nudges, which is unfortunate for Sunstein and Thaler 2008's explicit attempts to appeal to libertarians.

[22] See Chuck Ross, "President Obama Orders Behavioral Experiments on American Public," *The Daily Caller* September 15, 2015 https://dailycaller.com/2015/09/15/president-obama-orders-behavioral-experiments-on-american-public/ Accessed October 20, 2021.

[23] This is from Raz 1986: 377-378.

[24] This is from Hill 1999: 33.

[25] Faden et al. 1986: 355 argue that there is a kind of manipulation, which they call 'manipulation of options', that does alter one's options: "manipulation of options involves the direct modification of the options available to a person with the intent, which is successful, of modifying the person's behavior or beliefs." However, this does not seem relevant to the current discussion at any rate, since social-media sites do not change our options.

The healthy foods need to be put somewhere: maybe at eye level; maybe not. If the choice architect puts them at eye level, then the cafeteria-goers are being nudged towards them; if they are put elsewhere, then the cafeteria-goers are being nudged towards whatever is most visible and accessible. There is no escaping a nudge like this, to the extent that cafeterias need to be designed at all. Something similar is going on with at least some of the nudges online: for instance, the interface surrounding notifications needs to be designed in *some* way. If Facebook wants to nudge people away from using their site, then it is open to them to design notifications such that we are not prompted to engage with them. Still, we are being nudged. If Facebook wants to nudge people towards using their site more, then they can display notifications prominently and in certain attention-grabbing colors. There are all sorts of options in between, too: notifications displayed in boring and drab colors; notifications that one opts into; notifications about only those posts made by pre-selected users; and so on. The larger point is that we are always being nudged.

The same goes for some other nudges, too. If a site wants to send emails to users, then they have to ask for their permission. How they ask is part of the choice architecture: the option on the account-creation screen can have 'yes, I would like to receive emails' already selected, or not already selected, or it could be the user has to choose between a 'yes' and a 'no' option, or perhaps the question can be phrased in any number of ways. Still, no matter which option is taken, we are being nudged. However, this does not mean that all nudges are inevitable or unavoidable; I will call these other nudges the 'optional nudges'.[26] It is hard to see how the design of a cafeteria could be manipulative when the relevant features of the design were unavoidable, but this is not obviously so for the optional nudges. Yet, it does not seem that the optional nudges meet any of the definitions of manipulation presented earlier. For instance, it does not pervert decision-making processes or promote irrationality to load a YouTube video automatically after a person has finished one.[27] Neither does informing users that their comments have been liked

---

[26] By 'unavoidable', I do not mean that the nudges themselves cannot be defied; I mean that one cannot avoid designing the choice architecture such that people are being nudged.

[27] However, there is a perhaps stronger and more interesting way of arguing that the design of this technology, considered broadly, promotes (or takes advantage of, at least) some irrationality. For instance, consider Ward et al. 2016's finding that smartphones significantly reduce one's cognitive capacity; even when we resist the temptation to check the phone, the mere presence of it measurably reduces cognitive capacity. This might make us bad at reasoning in a different way than anti-nudging ethicists usually consider. This might be similar to the way that consumers at, e.g., malls are more likely to buy products when the building is laid out confusingly and so are more inclined towards less-deliberate, more-impulsive purchases (i.e., the Gruen effect). I see no reason for thinking that social-media companies try to disorient us so that we spend more time on their sites (i.e., that they nudge us towards using their platforms by cognitively stunning us), although perhaps Facebook's modification of the notification color could be likened to the way that malls are laid out. That being said, it might be more plausible or precise to say that Facebook changed the color to *grab our attention*, not to disorient us.

and that the users should click to see who liked their comments.[28] The latter might actually promote rationality and good decision-making by informing the user of something true.[29]

If we think of the offering of rewards or the threat of punishments as paradigmatic of manipulation, then the nudges that we see online are not manipulative. For instance, perhaps an offer of freedom to a prisoner in exchange for participation in a medical trial is manipulative since the offer is abnormally attractive; threatening that same prisoner with indefinite confinement if he or she does not participation could similarly be manipulative or outright coercive.[30] The fact that there are no increased rewards – or, really, any rewards at all – for doing what the choice architect online wants us to do obviates the concern that people might be manipulated by an unusually attractive offer. Similarly, the fact that there are no negative consequences for defying the choice architect – and there is not even the perception of a negative consequence – obviates the concern that users might be manipulated by a threat.

I recognize that some readers might still wonder about the manipulativeness of some of the inevitable nudges and might even doubt their inevitability at all. However, I think that none of them obviously meet the definition of manipulation, either. The mandated-choice nudge is particularly innocent: forcing users to choose whether they want to receive emails from a site can hardly be said to promote irrationality or perverting their decision-making processes. It might be helpful to think about the choice architecture in terms of *escape clauses*: one necessary condition for being manipulated might be that there was no way to escape the influence of the manipulator.[31] It might be objectionably difficult to delete one's Facebook account, but it is not so difficult at all to *escape* nudges of

---

[28] Something similar occurs when a platform recommends content that users might like: users are being informed of the existence of something; it might feel like a relentless insistence from, e.g., YouTube or Instagram, but being informed is hardly the promotion of irrationality. The same goes for the endless-scroll nudge: it is just more content or information, made more accessible.

[29] Faden et al. 1986: 362 identify what they call 'manipulation of information': "The manipulation of information is a deliberate act that successfully influences a person by nonpersuasively altering the person's understanding of the situation, thereby modifying perceptions of the available options. The manipulator does not change the person's actual options; only the person's perception is modified as a result of the manipulation. Thus, informational manipulation affects what a person believes." It might be that Facebook's notifications do meet this definition, but Faden et al. do not think that being so manipulated is incompatible with our autonomy, and they do argue that the manipulated person needs to be kept confused or in partial ignorance, and it is hard to see how Facebook is doing either of those things by notifying us; there does not seem to be much else that Facebook ought to inform us about. Other forms of manipulation of information, according to Faden et al. 1986: 363 concern straightforward deception, such as lying, withholding information, true assertion but with the omission of a crucial qualification, so on. Social-media companies might well do all these things to their users when it comes to the selling of data or privacy violations, but they do not seem to deceive us when it comes to nudging us.

[30] This articulation of the problem is indebted to Faden et al. 1986: 340.

[31] This analysis is coming from Wilkinson 2013, who develops an account of when nudges are manipulative.

social-media sites in other ways (e.g., by simply disabling notifications on one's phone or by disabling the feature by which YouTube automatically loads up the next video). It is impossible to say with certainty whether social-media sites sincerely *want* their users to defy the choice architect; their business interests, after all, depending on promoting engagement with their platforms.[32] It might be that their choice architects have roughly the same intention as those who try to manipulate us, and intention is a key component of manipulation.[33] That being said, it is, as a matter of fact, hardly even an inconvenience to defy the choice architect in virtually every case being discussed, and surely that counts against the nudges being manipulative.[34]

3. *Social-media companies nudge us towards the bad*

If the analysis ended at this point, it would seem as though the use of nudges on social-media sites is unobjectionable. The truth is more complicated. The canonical objections to nudging – which usually revolve around manipulativeness or the promotion of irrationality – do not succeed in these cases, but the problem is a far-deeper one: social-media companies nudge us towards what is bad for us.

One way that they do this is by nudging us to return to their platform while we are spending time away from it. The most obvious such nudge is the notification regarding some content on their platform when we are not currently using it. For instance, the Facebook app on our smartphone notifies us of some going-on on the platform when we are checking our email. A second way is by promoting continued engagement with their site, such as nudging people by promoting content that the software designers of the site have predicted they might like or by automatically loading more content

---

[32] The ethical problem here is that Wilkinson 2013: 352 knows that the most common form of nudging is to change the default option such that people need to opt *out* instead of opting in. Therefore, this common form of nudging seems to, by definition, have an escape clause built into it: they simply opt out, but maybe "while people have the formal freedom to opt out, in fact the nudging methods are so powerful that people go in the direction they are nudged and do not opt out when the nudges are poorly designed or do not suit them." Wilkinson's solution is to posit that the choice architect ought to *sincerely* intend the nudged to take the escape clause. This is a tricky problem, because business-owners generally do not intend that people *not use* their products, for instance; it seems implausibly strict to require that someone *intend* that a user not use their platform. It might be more plausible to maintain, instead, that the choice architect should sincerely intend that someone defy the nudge *if that person wants to defy the nudge*. Specifically, the choice architect should sincerely intend that someone should be free to do what they want, and if they do not want to use the platform, they should not use it. A restaurant-owner might sincerely intend for hungry patrons to eat at their restaurant, while also intending that those who are not interested in eating there go somewhere else.

[33] There is a strong consensus that manipulation requires that the manipulator *intend* to manipulate the victim. See Faden et al. 1986 and Wilkinson 2013 for examples and discussions.

[34] I say 'virtually all cases' because I am sensitive to the often-mentioned counter-example of the difficult process of deleting one's Facebook account.

after we are done consuming some content already. The difference between these two nudges lies in whether we are not on the platform at the moment and are being nudged to return to it or whether we are on the platform and being nudged to not leave it. Morally, there is no difference, since spending more time on social-media sites is bad for us, and that is the outcome of both kinds of nudges.

That more time spent on these sites is bad for our well-being is an empirical fact, one that we have ascertained through psychological research.[35] This is especially true of the demographic most likely to use social-media sites, namely, teenagers and young adults.[36] One striking feature of this research is that it shows that the more time we spend on social-media sites, the more pronounced the effects on our subjective and objective well-being become.[37] The implication is that there is much less harm involved in spending little time on social-media sites. *Social-media sites are not bad for us when we spend only a little bit of time on them.* There are many possible reasons for this. Perhaps when we spend little time on them daily, we tend to use them purposively during that time, maybe by speaking to far-flung friends. Maybe by spending little time on them daily, we are not so exposed to images from other people's lives that invite comparisons with our own that are not flattering to us. No matter what the reason is, although more research on new technologies is always welcome, it seems warranted to say that

[35] For instance, consider Riehm et al. 2019, which found that time spent on social-media sites was associated with an increase in mental-health problems, even after controlling for the participants' mental-health history. We can reproduce this effect in controlled trials to discern causation, too: Hunt et al. 2018 found that adolescent participants who were allotted only ten minutes daily to spend on each of Facebook, Snapchat, and Instagram's platforms "showed significant reductions in loneliness and depression over three weeks" compared to the group of adolescents who had unlimited access to these platforms. This is not to say that social-media use *itself* is bad or is necessarily bad. See Farim et al. 2020's systematic review of sixteen research studies: overall, the results were that social-media use can be both positive and negative. Consider Escobar-Viera et al. 2018's literature review of studies concerning the experience of LGBT minorities, included among the sixteen reviewed by Farim et al. 2020, which found that social-media sites can be productive places to get support and share experiences, but that it can also become a source of stress and cause of depression.

[36] See, for instance, Twenge and Campbell 2019, which found that adolescents who used social media lightly (i.e., less than one hour a day) reported substantially higher psychological well-being than those who used it heavily (i.e., more than five hours a day); specifically, "heavy users (vs. light) of digital media were 48% to 171% more likely to be unhappy, to be in low in well-being, or to have suicide risk factors such as depression, suicidal ideation, or past suicide attempts." They also found that the greatest drop in well-being occurred as one moves from moderate use to heavy use. However, studies that include older people do find that similar problems occur in these demographics too, although with less frequency; see Schimmele et al. 2021, which looks at Canadians between the ages of 15 and 64 and partially attributes the greater frequency of negative effects among the younger users to their more frequent use.

[37] The difference between subjective and objective well-being in this context reflects methodological differences: sometimes, the researchers will ask users to report their own well-being (i.e., by asking how the respondents feel), which measures subjective well-being; sometimes, the researchers will objectively measure the effects on users, such as by tracking the amount of sleep lost. Generally, research studies do both.

our well-being is negatively affected by spending more time on the platforms; the more time, the greater the negative effect.

The moral problem with this is that spending more time on the platforms is exactly what the nudges are designed to get us to do. We are hardly ever nudged to sign up for social-media sites in the first place.[38] We are nudged to spend more time on them when we are on them and to return to them when we are not. We are being nudged towards doing something bad for us, and that is the problem. This is a major ethical problem for the choice architects of these sites because nudging is intended as a libertarian-paternalist tool to promote what is good for us. Sometimes libertarian paternalists explain nudging in morally neutral language, such as when they say that a nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentive."[39] However, morally neutral descriptions of nudging are exceptional. Consider: "the golden rule of libertarian paternalism [is to…] offer nudges that are most likely to help and least likely to inflict harm."[40] Here is another handy and articulate formulation:

Libertarian Paternalism is the set of interventions aimed at overcoming the unavoidable cognitive biases and decisional inadequacies of an individual by exploiting them in such a way as to influence her decisions (in an easily reversible manner) towards the choices she herself would make under idealised conditions.[41]

Nudging is meant to be *towards the choices* that the nudged would choose. Certainly, the choice architect might be wrong about what the nudged would choose, and I take it that in these cases it is helpful that the nudges come with a low cost of defiance.

Even if we prefer to accept a morally neutral description of a nudge, it surely matters that nudges are always justified on *consequentialist* grounds.[42] The libertarian paternalist is concerned with *bringing about a kind of outcome*. No ethicist has tried to justify nudges on some other grounds. The nudges that we see on social-media sites are bad for us and cannot be justified by a consequentialist framework. If the very reasoning that justifies nudges in cases where they are justified does not justify them in social-media contexts, then these latter nudges seem condemned.[43]

---

[38] At least not by the social-media companies themselves: they have no obvious ability to reach out to people who have not signed up for their platforms.

[39] This is from Sunstein and Thaler 2008: 6.

[40] This is from Sunstein and Thaler 2008: 72.

[41] This is from Rebanato 2012: 6.

[42] For instance, when Sunstein 2012: 23 is pitching libertarian paternalism, he explains that "there are many opportunities for improving welfare without intruding on freedom of choice." The idea is that libertarian paternalism is consequentialist: specifically, it is a kind of welfare utilitarianism. We nudge in order to improve welfare.

[43] Pointing out the utilitarian background is important for cutting off the objection that we might

The same goes for other business practices in the offline world, even though these are not the subject of this paper. For instance, there seems to be something objectionable about a bartender nudging someone towards their twentieth glass of alcohol. In fact, we might even expect that the bartender should cut off that person entirely.

There is one more similar nudge that we ought to consider: the promotion of content. Social-media companies often promote content to users on the basis that the content is popular or "trending" at the moment or because the user's history on the site (or off the site) makes it likely in the eyes of the software designers that the user will engage with the content. At bottom, this is another kind of nudge to spend more time on the platform. The idea is that users prefer to spend more time on a site when they are engaging with content that others have liked or that the site has predicted the user in question will like. What is different about these nudges is that they can promote content that is bad for us independently of how much time we spend on the site. The *Wall Street Journal*'s 2021 exposé of Facebook and Instagram's effects on the mental health of teen-aged girls showcases some powerful examples of this design, such as promoting eating-disorder content to girls who had initially searched for content related to health and exercise.[44] Promoting content is a form of a nudging, and nudging girls towards eating disorders is morally wrong.[45] In late 2021, a whistleblower, Frances Haugen, who had worked for Facebook as a product manager and had been hired to protect against election interference, disclosed that the company promotes this form of content because of how successful it is according to their engagement-based metrics.[46] In other words, the nudges are used because they are effective at getting us to spend more time on the platforms.

want to apply libertarian paternalism to *state action*, but we should not expect private companies to have the same duties to us or to society. However, the thoroughgoing utilitarian will hold that everyone should be held to the same standards of promoting well-being. This matters because, as we shall see in the rest of this paper, the moral problem is that some private social-media companies are failing in their moral obligations, as those obligations are understood by the same reasoning that legitimizes the use of nudging in the first place. Consequentialism legitimizes nudging *and* objects to the behavior of social-media platforms.

[44] See Wells et al., "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show," *Wall Street Journal* September 14, 2021. https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=hp_lead_pos7&mod=article_inline Accessed October 18, 2021.

[45] We might also think of the way that social-media companies nudge users towards conspiracy theories such as QAnon. Leaked Facebook documents have revealed internal research studies showing that users who show an interest in conservative content (e.g., Donald Trump and Fox News) were being recommended QAnon content within two days of joining Facebook. However, it is not obvious that merely having false beliefs about the lives and motivations of American politicians is, in fact, harmful. The promotion of eating-disorder content is more obviously harmful. For the reporting on these leaked documents, see Kari Paul and Dani Anguiano, "Facebook crisis grows as new whistleblower and leaked documents emerge," *The Guardian* October 23, 2021 https://www.theguardian.com/technology/2021/oct/22/facebook-whistleblower-hate-speech-illegal-report Accessed October 23, 2021.

[46] See Jeff Horwitz, "The Facebook Whistleblower, Frances Haugen, Says She Wants to Fix the

By the same token, some of these social-media sites know that we are being nudged towards what is bad for us. For years, Facebook has conducted research internally on the effects of the services of Instagram and Facebook on users' health. The *Wall Street Journal* published the documents released by the Facebook whistleblower, and we can use them to learn what Facebook itself has known about the damage time on its platforms does to its users. The research showed that as far as Facebook is aware, among teenagers, 63.37% report sleep issues, 60.10% report work stress, 57.56% report social-comparison issues, and 55.59% report body-image issues. Among adults, 57.77% report sleep issues, 57.54% report financial stress, and 52.54% report body-image issues. There are other problems, too: among teenagers, reports of sadness, anxiety, laziness, and a fear or feeling of missing out (FOMO) were all above 50%; among adults, reports of sadness, anxiety, FOMO, eating issues, loneliness, work stress, and social-comparison issues, were all above 40%.[47] One might wonder whether this internal research is as evidentially solid as peer-reviewed research, but the point is that Facebook cannot defend its use of nudging by saying that it believes that its platforms are good for its users.[48]

When we speak specifically of teenaged girls on Instagram, Facebook's internal research disclosed that 66% experienced negative social comparison (as opposed to 40% for boys); 52% said that these feelings were caused by images related to beauty; and 32% said that when they felt bad about their bodies, *Instagram made them feel worse*.[49] This last part is the crucial thing: a large minority of girls report that Instagram makes the problems worse. Yet, Facebook nudges

Company, Not Harm It," *Wall Street Journal* October 3, 2021 https://www.wsj.com/articles/facebook-whistleblower-frances-haugen-says-she-wants-to-fix-the-company-not-harm-it-11633304122 Accessed October 18, 2021.

[47] The *Wall Street Journal*'s staff published this leaked document on September 29, 2021, which had appeared on a Facebook internal site in 2019 as "Teens & young adults on IG [that is, Instagram] and FB [that is, Facebook]" https://s.wsj.net/public/resources/documents/mental-health-findings.pdf. Accessed October 18, 2021.

[48] Respondents were asked "in the last 30 days, have you experienced…?". It seems that Facebook then listed these options. There were 22410 respondents across multiple countries including the United States of America and Japan. I do not know how the respondents were chosen. It could well be that there is some self-selection bias among respondents who completed the survey; that is, it could be that the people with pronounced negative effects were more likely to respond to the survey, whereas people with generally positive or indifferent experiences were more likely to decline to respond. Since the research was released by a whistleblower, we lack the full context required to interpret the data. I do not use this research to conclude that use of Facebook's platforms has all of these effects; I use it to conclude that Facebook cannot defend itself by saying that it believes that time on its platforms is good for users. When we take the peer-reviewed research that *is* available and combine it with the data that Facebook has collected internally, we see that there is no way for the company to plead ignorance. *As far as they know, they are nudging people towards what is bad for them*.

[49] The *Wall Street Journal*'s staff published this leaked document on September 29, 2021, which had appeared on a Facebook internal site in 2020 as "Teen Girls [sic] Body Image and Social Comparison on Instagram – An Exploratory Study in the US" https://s.wsj.net/public/resources/documents/teen-girls-body-image-and-social-comparison-on-instagram.pdf Accessed October 18, 2021.

them to spend more time on the platform. A nudge towards what is bad for us should be condemned on the same consequentialist grounds that justifies other nudges. The fact that Facebook has evidence that that towards which we are nudged is bad for us is even more damning.

Some of these nudges do not pass the publicity-principle test. I said earlier that most nudges on social-media platforms do comply with this principle: after all, most of the time, the companies in question are being transparent in the first place, and few people would see something wrong in principle with a company nudging you to use their free product. However, I think that the promotion of eating-disorder content and of other content that harms young girls' sense of themselves *are* nudges that Facebook would be reluctant to publicly defend. We can conclude this from the fact that Facebook *is* reluctant to publicly defend these nudges. Even though spending increasing amounts of time on Facebook is bad for us, I do not think that Facebook would be reluctant to defend its use of nudges in general, since there is nothing particularly embarrassing or shameful about, e.g., notifying users about a post that was recently made by a friend. However, the public outrage — which was sparked by these leaked internal documents, which prompted the whistleblower to testify in the United States Senate, and which was in large part a response to the revealed damage to teenaged girls' mental health — and Facebook's own response, which centered around denying the allegations and then also pledging to develop new "safety features" for teenagers, reveal that there was something especially egregious about this aspect of their platforms.[50]

4. *The same technologies can be used for good*

It is important that we recognize the moral permissibility of nudges in principle on social-media platforms. The problem with currently existing nudges online is that they nudge us towards what is bad. However, as long as we affirm that nudging in principle is morally permissible, we can use the same choice-architecture designs to nudge people towards what is good, namely, spending less time on these sites. For instance, when someone watches more than, say, five or six YouTube videos in a row, perhaps there ought to be a message that comes up that informs them of how much time they have been spending, and they have to look at this message for, say, ten or fifteen seconds before dismissing it. Other examples include making notifications opt-in, such that they are by default disabled; perhaps notifications should be automatically disabled after a user has spent more than twenty or thirty minutes on the platform that

---

[50] See Samantha Subin, "Facebook says it will add new safety features, notably for teens on Instagram, after bombshell whistleblower leak," *CNBC* October 11, 2021 https://www.cnbc.com/2021/10/11/facebook-will-add-new-safety-features-for-teens-following-whistleblower-leak.html Accessed October 18, 2021.

day. At a minimum, social-media companies ought to not promote content that leads to mental-health problems, such as eating disorders; it might be best for them not to promote content at all and simply let users engage with whatever they happen upon. This might help ameliorate other often-discussed philosophical problems with the design of the Internet generally, such as the loss of serendipity.[51] However, I am not maintaining that there is a moral obligation for social-media companies to nudge us towards the good, although there is a moral obligation for them not to nudge us towards the bad. In practice, since nudges are often unavoidable, failing to nudge us towards the bad will look like nudging us towards the good. My point here is merely that since nudges are in principle morally permissible, social-media companies can permissibly exploit the same choice architecture – such as the power of default rules – for the sake of the good.

It is not clear whether it is too much to ask social-media platforms to change themselves. Whether the government ought to get involved is a question for coercive, not libertarian, paternalists. Facebook's vice president of global affairs has said that "we have no commercial incentive to do anything other than try and make sure that the experience is positive."[52] This is surely true in one sense but false in the sense that he meant it. He said this to dismiss the allegations that Facebook had contributed to mental-health issues, and in this sense, it was false: the psychological research is clear that people overwhelmingly tend to be attracted to negative content.[53] The claim that Facebook has no commercial incentive to promote negative content, therefore, is baseless. Any social-media site can easily leverage this feature of human psychology to increase user engagement. In another sense, the claim is true: the great public backlash, sparked as the effects of social-media sites are felt, noticed, and decried, gives the platforms a considerable commercial incentive to *not* harm users' mental health, even if doing so would be (and often is) in their immediate financial interests.

[51] See, for instance, Sunstein 2017 and Reviglio 2019 for discussions of serendipity. Especially for Sunstein, something valuable is lost when what we see online are only those things that we have selected to see (or those things that various companies have predicted that we would like to see). The serendipitous stumbling upon something that we did not think that we would like (either because we tend to dislike that sort of thing or because we were previously unexposed) is an important consideration for these thinkers, among others.

[52] See Samantha Subin, "Facebook says it will add new safety features, notably for teens on Instagram, after bombshell whistleblower leak," *CNBC* October 11, 2021 https://www.cnbc.com/2021/10/11/facebook-will-add-new-safety-features-for-teens-following-whistleblower-leak.html Accessed October 18, 2021.

[53] It is hard to prove negative content on social-media sites gets more attention, but we can study this indirectly. For instance, Soroka et al. 2019 explains that negative biases "in human cognition are well documented" and further discovered that their effects on news coverage (and news production) are profound and exist that because of the fact that negative content causes more engagement from consumers (which is surely the same effect that is leading Facebook to prioritize negative content for the promotion of user-engagement, as per the whistleblower's testimony). (Soroka et al. 2019 did find that there individuals tended to vary in their responsiveness to negative content and that this might lead to more positive content in the future.)

Further evidence that it is not unreasonable to be optimistic about social-media companies reforming themselves is that some already do nudge users in the right direction. TikTok, for instance, implemented a nudge of just this sort: notifying users when they have spent a certain amount of time on the platform. TikTok, not uniquely, also gives users the ability to set (or entirely disable) daily screen-time limits, such that an interested user will need to re-enter their password if the user wants to exceed their limit. This illustrates nudging well: helpful, yet trivially easy to defy. Moreover, Facebook's response to the immense public backlash in light of the whistleblower's testimony to the United States Senate included saying that they will be developing a feature on Instagram to nudge people away from the site after spending a certain amount of time on it. Facebook and I are on the same page with how to proceed. Backlash is apparently sufficient to persuade these private companies to change their choice architecture.

It is not obvious that these nudges will be powerful enough to mitigate or prevent the damage done by social-media platforms to users' mental health.[54] If nudging people away from these platforms does not appreciably improve the situation, we might consider more aggressive action, but nudges are convenient tools since they preserve the freedom of choice of would-be users. Having social-media companies reform themselves in a way that preserves freedom of choice and that protects mental health would be a happy alternative to both the status quo and the possible world in which the government uses coercion, for both moral reasons and for pragmatic considerations about the efficacy of government regulation of the Internet. At a minimum, by transforming nudges *towards* social-media sites into nudges *away* from those sites, we improve the moral status of the choice architecture of these sites and of the Internet generally.

Bibliography

Cohen, Shlomo. "Nudging and Informed Consent," *The American Journal of Bioethics* 13 6 (2013): 3-11.

Conly, Sarah. "Coercive Paternalism in Health Care: Against Freedom of Choice," *Public Health Ethics* (2013): 241-245.

Dunt, Ian. "Nudge nudge, say no more. Brits' minds will be controlled without us knowing it," *The Guardian* February 5, 2014 https://www.theguardian.com/commentisfree/2014/feb/05/nudge-say-no-more-behavioural-insights-team, Accessed October 20, 2021.

---

[54] Social-media platforms are designed to be addictive. (For a discussion of *behavioral* addictions, such as addictions to gambling, sex, and, indeed, smartphone-related technology, as opposed to *substance* addictions, such as to nicotine, see Alter 2017.) It might be that nudging people away from something to which they are addicted is simply insufficient to see a meaningful decrease in the time spent on the platforms. Either way, it is morally important for social-media companies to cease nudging people towards what is bad for them; further, the cost of implementing libertarian-paternalist strategies is negligible when compared with the cost of implementing far blunter coercive-paternalist strategies, and this alone gives us a reason to consider the former first.

Escobar-Viera C, Whitfield D, Wessel C, et al. "For better or for worse? A systematic review of the evidence on social media use and depression among lesbian, gay, and bisexual minorities." *JMIR Mental Health* 5 3 (2018) DOI: 10.2196/10496

Farim, Fazida, and Azeezat A. Oyewande, Lamis F. Abdalla, Reem Chaudhry Ehsanullah, and Safeera Khan, "Social Media Use and Its Connection to Mental Health: A Systematic Review," *Cureus* 12 6: (2020) DOI: 10.7759/cureus.8627

Faden, Ruth, Thomas Beauchamp, and Nancy King (1986). *A History and Theory of Informed Consent*, Oxford: Oxford University Press.

Fox, Craig R. and Tannenbaum, David. "The Curious Politics of the 'Nudge'", *The New York Times* September 26, 2015 https://www.nytimes.com/2015/09/27/opinion/sunday/the-curious-politics-of-the-nudge.html Accessed October 18, 2021.

Gigerenzer, Gerd. (2015). "On the Supposed Evidence for Libertarian Paternalism" *Review of Philosophy and Psychology* 6: 361–383.

Kurtz, Sheldon F. and Saks, Michael J. "The Transplant Paradox: Overwhelming Public Support for Organ Donation vs. Under-Supply of Organs: The Iowa Organ Procurement Study." *Journal of Corporation Law* 21 (1996): 767–806.

Halpern, S. D., et al. "Harnessing the power of default options to improve health care." *New England Journal of Medicine* 357: (2007) 1340–1344.

Hill, Thomas. (1991) *Autonomy and Self-Respect*. Cambridge: Cambridge University Press.

Holte, Alex J. and Ferraro, Richard F. "True Colors: Grayscale Setting Reduces Screen Time in College Students." *The Social Science Journal* 2020. DOI: 10.1080/03623319.2020.1737461

Holte, Alex J., Giesen, Desiree T. and Ferraro, Richard F. "Color me calm: Grayscale phone setting reduces anxiety and problematic smartphone use." *Current Psychology* 2021. https://doi.org/10.1007/s12144-021-02020-y

Hunt, Melissa G., Rachel Marx, Courtney Lipson and Jordyn Young, "No More FOMO: Limiting Social Media Decreases Loneliness and Depression" *Journal of Social and Clinical Psychology*, 37 10: (2018) 751-768.

Horwitz, Jeff, "The Facebook Whistleblower, Frances Haugen, Says She Wants to Fix the Company, Not Harm It," *Wall Street Journal* October 3, 2021 https://www.wsj.com/articles/facebook-whistleblower-frances-haugen-says-she-wants-to-fix-the-company-not-harm-it-11633304122 Accessed October 18, 2021.

Lewis, Paul, "'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia," *The Guardian* October 6, 2017 https://www.theguardian.com/technology/2017/oct/05/smartphone-addiction-silicon-valley-dystopia Accessed October 19, 2021

Linkenbach, Jeffrey W. "The Montana Model: Development and Overview of a Seven-Step Process for Implementing Macro-Level Social Norms Campaigns," in Perkins, H. Wesley, ed. *The Social Norms Approach to Preventing School and College Age Substance Abuse*. New York: Jossey-Bass, 2003, 182-208.

Linkenbach, Jeffrey W. and H. Wesley Perkins. "MOST of Us are Tobacco Free: An Eight-Month Social Norms Campaign Reducing Youth Initiation of Smoking in Montana," in Perkins, H. Wesley, ed. *The Social Norms Approach to Preventing School and College Age Substance Abuse*. New York: Jossey-Bass, 2003, 224-234.

Mitchell, Gregory. (2005). "Libertarian Paternalism is an Oxymoron," *Northwestern University Law Review* 99 3: 1245-1277.

Morewedge, Carey K. and Giblin, Colleen E. "Explanations of the endowment effect: an integrative review," *Trends in Cognitive Sciences* 19 6: (2015) 339-348.

Newport, Cal. (2019) *Digital Minimalism: Choosing a Focused Life in a Noisy World*. New York: Portfolio.

Paul, Kari, and Dani Anguiano, "Facebook crisis grows as new whistleblower and leaked documents emerge," *The Guardian* October 23, 2021 https://www.theguardian.com/technology/2021/oct/22/facebook-whistleblower-hate-speech-illegal-report Accessed October 23, 2021.

Perkins, H. Wesley, ed. (2003) *The Social Norms Approach to Preventing School and College Age Substance Abuse*. New York: Jossey-Bass.

Prelec, Drazen and Duncan Simester. "Always Leave Home Without It: A Further Investigation of the Credit-Card Effect on Willingness to Pay," *Marketing Letters* 12 1: (2001) 5-12.

Raz, Joseph. (1986) *The Morality of Freedom*. Oxford: Clarendon Press.

Rawls, John. (1971) *A Theory of Justice*. Oxford: Clarendon Press.

Rebanato, Riccardo. (2012) *Taking Liberties: A Critical Examination of Libertarian Paternalism*, London: Palgrave Macmillan.

Riehm, Kira E. and Kenneth A. Feder, PhD, Kayla N. Tormohlen, et al. "Associations Between Time Spent Using Social Media and Internalizing and Externalizing Problems Among US Youth," *JAMA Psychiatry*. 76 12: (2019) 1266-1273. doi:10.1001/jamapsychiatry.2019.2325

Reviglio, Urbano. "Serendipity as an emerging design principle of the infosphere: challenges and opportunities," *Ethics and Information Technology* 21: (2019) 151-166.

Simon, H. A. "Rational choice and the structure of the environment." *Psychological Review* 63: (1956) 129–138.

Samantha Subin, "Facebook says it will add new safety features, notably for teens on Instagram, after bombshell whistleblower leak," *CNBC* October 11, 2021 https://www.cnbc.com/2021/10/11/facebook-will-add-new-safety-features-for-teens-following-whistleblower-leak.html Accessed October 18, 2021.

Schimmele, Christoph, and Jonathan Fonberg and Grant Schellenberg. "Canadians' Assessments of Social Media in Their Lives" *Economic and Social Reports from Statistics Canada* 2021 https://doi.org/10.25318/36280001202100300004-eng

Soroka, Stuart, Patrick Fournier, and Lilach Nir,"Cross-national evidence of a negativity bias in psychophysiological reactions to news" *PNAS* 116 38: (2019) 18888-18892.

Sunstein, C. and Thaler, R. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 70 (4): 1159-1202.

Sunstein, Cass R. *Why Nudge?* New Haven, CT: Yale University Press, 2012.

----. *Choosing Not to Choose*. Oxford: Oxford University Press, 2015.

----. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2017.

Tannenbaum, David, Craig R. Fox, and Todd Rogers. "On the misplaced politics of behavioural policy interventions" *Nature Human Behavior* 1: (2017) https://doi.org/10.1038/s41562-017-0130

Thaler, R. and Sunstein, C. (2008) N*udge: Improving Decisions about Health, Wealth, and Happiness*. New Haven CT: Yale University Press.

Twenge, Jean M., and Keith Campbell, "Media Use Is Linked to Lower Psychological Well-Being: Evidence from Three Datasets" *Psychiatric Quarterly* 90: (2019) 311-331.

van Deursen, Alexander J.A.M, and Jan A.G.M. van Dijk. "Using the Internet: Skill related problems in users' online behavior" *Interacting with Computers* 21: (2009) 393-402

Veetil, V.P. (2011). "Libertarian paternalism is an oxymoron: an essay in defence of liberty," *European Journal of Law and Economics* 31: 321-334.

Wall Street Journal staff, "Teens & Young Adults on IG and FB," *Wall Street Journal* September 29, 2021 https://s.wsj.net/public/resources/documents/mental-health-findings.pdf Date Accessed October 18, 2021.

Wall Street Journal staff, "Teen Girls [sic] Body Image and Social Comparison on Instagram – An Exploratory Study in the US," *Wall Street Journal* September 29, 2021 https://s.wsj.net/public/resources/documents/mental-health-findings.pdf Date Accessed October 18, 2021.

Ward, Adrian F., Kristen Duke, Ayelet Gneezy, and Maarten W. Bos, "Brain Drain: The Mere Presence of One's Own Smartphone Reduces Available Cognitive Capacity" *Journal of the Association for Consumer Research* 2 2: (2017) https://doi.org/10.1086/691462

Wells, Georgia, Jeff Horwitz, and Deepa Seetharaman, "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show," *Wall Street Journal* September 14, 2021. https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=hp_lead_pos7&mod=article_inline Accessed October 18, 2021.

Wechsler, Henry, Jae Eun Lee, Meicun Kuo, and Hang Lee. "College Binge Drinking in the 1990s: A Continuing Problem. Results of the Harvard School of Public Health 1999 Alcohol Study." *Journal of the American College of Health* 48 (2000): 199-210.