



Click-Gap, paternalism, and tech giants' relationships with their users

J. L. A. Donohue¹

Received: 2 May 2023 / Accepted: 22 October 2023
© The Author(s) 2023

Abstract

The spread of misinformation and fake news raises important problems for our society and for our democracy. From the January 6 attack on the U.S. Capitol to vaccine hesitancy, from suppressing voter turnout to peddling conspiracy theories, we know that these problems are real and need to be taken seriously. While misinformation is not a new problem for democracy, it can spread more quickly and easily because of new media's design and popularity. Given these problems, it is encouraging that some technology companies are taking steps to reduce the spread of misinformation and fake news on the platforms they manage. Despite this seemingly positive development, some scholars have criticized some interventions designed to combat the spread of misinformation and fake news as paternalistic. For example, a 2019 Facebook intervention called Click-Gap aimed to reduce the amount of low-quality content (including fake news and misinformation) that users see on their NewsFeed. Click-Gap has been criticized as an instance of epistemic paternalism because it was adopted (1) with the goal of improving the epistemic status of its users and (2) irrespective of what the company believed the wishes of its users to be. If interventions like Click-Gap are problematic because paternalistic, those of us interested in the ethics of technology would face a dilemma—either endorse technology companies treating their users paternalistically or endorse their failing to act to combat the spread of misinformation and fake news on their platforms. Both options seem to me to be problematic. While paternalism may sometimes be permissible, I think we should be very hesitant to endorse a paternalistic relationship between technology companies and their users. The relationship does not seem to bear the right sort of structure to one in which paternalism might be appropriate, if it ever is. The second option seems, if anything worse: surely technology companies should not stand by and change nothing about their platforms despite the spread of misinformation and fake news in those spaces. In this paper, I argue that Click-Gap and interventions like it are not paternalistic, contrary to the conclusion of other scholars. Further, I will argue that the focus on paternalism itself is actually a red herring here. While not just any intervention or strategy that purports to reduce fake news and misinformation is permissible, we should want technology companies to take user well-being seriously and be able to take that well-being as a direct reason for action. Their doing so is not paternalistic nor even morally problematic, and it should not be criticized as such.

Keywords Paternalism · Facebook · Algorithms · Sphere of agency

As one of the leading platforms where people share information and express themselves, misinformation is an ongoing challenge for us. With millions of Americans using our services every day, there will always be things we miss. However, I believe we do more to address misinformation than any other company, and I am proud of the systems we have built.

—Mark Zuckerberg's written testimony before Congress; March 25, 2021¹

1 Introduction

The spread of misinformation and fake news raises important problems for our society and for our democracy. From the January 6 attack on the U.S. Capitol to vaccine hesitancy, from suppressing voter turnout to peddling conspiracy

✉ J. L. A. Donohue
jlad@uark.edu

¹ Department of Philosophy, University of Arkansas, 318 Old Main, Fayetteville, AR 72704, USA

¹ Hearing on “Disinformation Nation: Social Media's Role in Promoting Extremism and Misinformation”; <https://energycommerce.house.gov/committee-activity/hearings/hearing-on-disinformation-nation-social-medias-role-in-promoting>.

theories, these are serious problems. While misinformation is not a *new* problem for democracy, it can spread more quickly and easily because of new media's design and popularity.² Given these problems, it is encouraging that some technology companies are taking steps to reduce the spread of misinformation and fake news on the platforms they manage.³

Despite this seemingly positive development, some interventions designed to combat the spread of misinformation and fake news have been criticized as paternalistic. For example, a 2019 Facebook intervention called Click-Gap aimed to reduce the amount of low-quality content (including fake news and misinformation) that users see on their NewsFeed. In recent work, Alan Rubel, Clinton Castro, and Adam Pham have criticized Click-Gap as an instance of epistemic paternalism because it was adopted (1) with the goal of improving the epistemic status of its users and (2) irrespective of what the company believed the wishes of its users to be ([9, 38]). If interventions like Click-Gap are problematic because paternalistic, those of us interested in the ethics of technology would face a dilemma—either endorse technology companies treating their users paternalistically or endorse their failing to act to combat the spread of misinformation and fake news on their platforms. Both options seem to me to be problematic. Even if we think that paternalism may sometimes be permissible, I think we should be very hesitant to endorse a paternalistic relationship between technology companies and their users. The relationship does not seem to bear the kind of structure in which paternalism might be appropriate, if it ever is.⁴ The second option seems,

if anything worse: surely technology companies should not stand by and change nothing about their platforms despite the spread of misinformation and fake news in those spaces.

In this paper, I argue that Click-Gap and interventions like it are *not* paternalistic, contrary to the conclusion of these other scholars.⁵ Further, I will argue that the focus on paternalism itself is actually a red herring here. While not just any intervention or strategy that purports to reduce fake news and misinformation is permissible, we should *want* technology companies to take user well-being seriously and be able to take that well-being as a *direct* reason for action. Their doing so is not paternalistic nor even morally problematic, and we should not criticize it as such.

Part of my argument will involve articulating why we care about paternalism in the first place—why paternalistic behavior is in general problematic and requires our moral attention.⁶ I will argue that underlying our moral concern with quintessentially paternalistic behavior is *not* a concern that others have taken our well-being as relevant to their action but rather, that they have failed to respect our sphere of legitimate agency.⁷ They have acted in a space that properly belongs to us to make decisions about, and *that* is morally problematic. If my argument is successful, we should concern ourselves not with paternalism as such (at least not as it is widely understood) but rather interferences with our sphere of legitimate agency.

I will conclude with some thoughts about the implications of my argument for technology companies' relationships with their users. While I hope to show that they *can* (and I think they should!) take the well-being of their users seriously and incorporate it into their reasons for action, I also argue that they do *not* have complete legitimate control over parts of their platforms that might at first pass seem to be fully within their sphere of legitimate agency. This means that both users and society have a strong claim to

² See (among others), Allcott et al. [2], Anderson [3], Brown [6–8], Castro et al. [9], Coleman [13], 10, Millar [29], Nguyen [33], Oremus et al. [35], Pariser [37], Rubel et al. [38], Sunstein [42] and the World Health Organization [43].

³ For example, Facebook adopted an intervention called “Click-Gap” in 2019. (I explain in more detail in the body of the paper about what this intervention does and how it works.) Other examples include Reddit's policy of “quarantining” certain “communities” (also known as “subreddits”), YouTube's policy of removing some manipulated or misattributed content and giving strikes to YouTube accounts that engage in behaviors that violate their misinformation policy, and Twitter's policy of removing or labeling “Tweets” that violate its misinformation policy (2023).

⁴ Perhaps we might endorse paternalistic treatment of children by their parents or of citizens by their governments (though I am skeptical even in these cases), but these relationships seem importantly different from the relationship between tech companies and their users. Parents are charged with taking care of their children, and at least part of the role of government seems to be to take care of its citizens and care for their well-being. But private, for-profit companies are not in a position of caring for their users. That is not their role nor relationship. I think we should be *very* hesitant to endorse paternalistic treatment in such cases, even more hesitant than we might be in care-giving relationships. For this reason I am especially concerned with Rubel et al.'s treatment of this issue: “epistemically paternalistic policies should be a perennial part of the internet information envi-

Footnote 4 (continued)

ronment” (2020, 22; 2021, 121). I say more about my concerns with Rubel et al.'s position below. I explore in more detail the relationship that must exist between a paternalist and a target for permissible paternalism to be on the table in [Redacted].

⁵ I have in mind especially Rubel et al., who argue that Click-Gap and related interventions are paternalistic and justified (2020; 2021). Some scholars follow them in their classification, including Habgood-Coote [17] (Forthcoming, 11), Handfield [18] (Forthcoming, 2), and (Kitsik [24], 1).

⁶ As other scholars have noted, it is not clear that epistemic paternalism is even *pro tanto* problematic, at least if it is understood in ways that it is typically defined [1, 23, 24, 28]. So we need clarity regarding what is morally problematic about paternalism in the first place, such that it is worthy of our moral attention.

⁷ The idea that what matters to us about paternalism is the disrespect such action typically shows for our sphere of agency is inspired by Shiffrin's *account* of paternalism [41].

having some control over parts of the platforms and how those platforms operate.⁸ In the end, I hope to have brought us closer to understanding how technology companies *ought* to interact with their users: they cannot treat them paternalistically, but paternalism was never the right framework for evaluation in the first place.

2 What is Click-Gap?

Some of the details of Click-Gap are important to understanding my argument, so let me begin with an explanation of the intervention itself. Click-Gap is a change to Facebook's NewsFeed algorithm that Meta adopted in April 2019.⁹ The NewsFeed is one of the main ways users interact with Facebook: posts appear in their NewsFeed, and the order in which those posts appear is determined by a proprietary AI-based algorithm created and maintained by the company.¹⁰ The order is not determined by a simple metric, such as which of the user's friends posted first. Rather, it is determined by a variety of factors that the algorithm takes into account, such as how engaging the post is likely to be, how engaging the post already has been, and how close the user is to the poster, among others. (Note that all of these terms: engaging, close, etc. have to be quantified in some way, and so Meta makes decisions about what counts as e.g., "engaging" and "close".) Users have some control over their NewsFeeds, as they can choose who to be "friends" with on the platform and can choose to "unfollow" or "block" some of those friends if they would prefer content posted by those friends not show up in their NewsFeed. Users also impact what appears in their NewsFeed by interacting with content on Facebook: they can choose which content they prefer to interact with when they "like" or "comment" on posts, and this information in turn is fed into the NewsFeed arrangement algorithm. Exactly how all of these factors

work together to result in the NewsFeed users actually seen is primarily hidden and fully under Meta's control.¹¹

Meta makes fairly regular updates to the NewsFeed algorithm, some more significant than others. For example, a 2018 NewsFeed algorithm update was purportedly aimed at encouraging interaction, and so it elevated posts by friends and family as well as viral memes and divisive content, since such posts and content tend to receive the most interaction from users [21]. An earlier update had prioritized time spent on the platform rather than engagement during that time, and so it had elevated a different kind of post, including professionally produced videos [35].

Meta adopted Click-Gap with the professed goal of taking steps to slow the spread of misinformation on its platform. The idea behind the algorithmic intervention is actually rather simple: rank posts lower in users' NewsFeeds if they suffer from a larger "Click-Gap" than others. What is a "Click-Gap"? It is the gap between two numbers: (A) the number of clicks generated to a site from *within* Facebook and (B) the number of clicks generated to a site from *outside* Facebook. Take a simple example. Suppose I post a link to an article by the New York Times. This article is unlikely to have a large "Click-Gap". Why? Internet users are fairly likely to navigate to the New York Times and read its articles from outside of Facebook. They might navigate there directly, follow a link someone emails them, or follow a link from another website. While they might also navigate there from within Facebook, the gap between the number of individuals who navigate there from within Facebook (A) and those who navigate there from outside it (B) is unlikely to be very large. Contrast this with a case of an article written by the Internet Research Agency. These articles are *designed* to generate lots of clicks from within Facebook and do not have significant followings outside Facebook.¹² And so the difference between the number of clicks they generate from within Facebook (A) and the number they generate from without (B) is likely to be large. The Click-Gap intervention is designed to recognize the size of this gap and downgrade posts in users' feeds that have large gaps. The goal of downgrading such posts by adopting Click-Gap was to "identify and demote low-quality content, such as fake news, in users' NewsFeeds to prevent it from going viral within Facebook" [38], 122. Note again that we have a *proxy*: Click-Gap doesn't set out *directly* to attempt to determine if a particular article is fake news but rather demotes content with a larger gap. We have, it seems to me, good reason to think this kind

⁸ Cf. Coleman [13] and Lazar [26].

⁹ At the time that it made the change, the company was called Facebook, but it has more recently changed its name to Meta. Because the name change facilitates distinguishing more easily between the company and the social media platform, I will (a bit anachronistically) call the company "Meta" throughout this paper and reserve "Facebook" for the social media platform itself.

¹⁰ After this paper was written, Meta introduced a new update to the Facebook platform: there are now different feeds users can scroll through, including "Home" and "Feeds". My argument here focuses on the NewsFeed arrangement algorithm as it operated before the most recent update in summer 2022. It will be relevant to other updates as well, though, since it has implications for how tech companies should conceive of their relationship with their users and their users' well-being generally speaking.

¹¹ For a useful history of the NewsFeed arrangement algorithm, see Oremus et al. [35].

¹² The Internet Research Agency is a Russian company engaged in online influence operations on behalf of Russian business and political interests [11].

of proxy will at least approximately track fake news, but it is (1) unlikely to be perfect and of course may demote some “good” content and (2) worth noticing that the intervention doesn’t directly target the aimed-at content. There is some evidence that Click-Gap and related interventions have had a positive impact at slowing the spread of misinformation and fake news on Facebook [2].

There are some worries one might lodge here about Click-Gap. Will it inappropriately privilege mainstream news media over independent or small publications? It very well might. And we could worry about whether that was a sacrifice worth making in order to reduce the spread of fake news and misinformation. These and related questions are difficult and interesting, but they are not the focus of my argument here. I am focused on a particular kind of criticism that has been wielded against Click-Gap: namely, the criticism that Click-Gap is paternalistic. And this criticism will help to illuminate a set of questions about Meta and similarly-situated technology companies. What is the right way to think about the relationship between them and their users?

3 Click-Gap and Paternalism

Despite its success at reducing the spread of misinformation and fake news on Facebook, the Click-Gap intervention has been met with some (mild) criticism. Though scholars are largely in favor of the intervention, it has been criticized as being paternalistic. In their recent book *Algorithms and Autonomy: the Ethics of Automated Decision Systems*, Alan Rubel, Clinton Castro, and Adam Pham argue that Click-Gap is an instance of what they call justified epistemic paternalism. They conclude that “epistemically paternalistic policies should be a perennial part of the internet information environment” ([9], 22; [38], 121). I am concerned that adopting a view according to which social media companies are permitted and encouraged to treat their users paternalistically leaves decision-making and power in precisely the wrong place, misunderstanding the relationship between tech companies, their users, and the legitimate spheres of agency of each.¹³ As we work to understand and evaluate their argument, it will be helpful to have a conception of paternalism in mind. Let us adopt Rubel, Castro, and Pham’s conception of paternalism, since theirs is the argument I interrogate:

Paternalism: *P* acts paternalistically toward *S* by ϕ -ing if and only if:

Insensitivity: *P* does so irrespective of what *P* believes the wishes of *S* might be.

Expected Improvement: *P* does so just because *P* judges that ϕ -ing might or will advance *S*’s ends (*S*’s welfare, interests, values or good) ([38], 125, [39], 127).¹⁴

Suppose I require my son, James, to wear pants to school because it is cold outside and pants are more likely than shorts to keep him comfortable and healthy. And I require him to do so without regard for what I take his wishes on the matter to be. Because I act without regard to James’ wishes, my action satisfies *insensitivity*. Because I act *just because* I judge that James will be better off if he wears pants and not shorts, my action satisfies *expected improvement*. On this definition I act paternalistically toward James by requiring him to wear pants.

Click-Gap is understood not just as paternalistic, though, but rather *epistemically paternalistic*. So we need to adjust our conception a bit—

Epistemic paternalism: *P* acts epistemically paternalistically toward *S* by ϕ -ing if and only if:

Insensitivity: *P* does so irrespective of what *P* believes the wishes of *S* might be.

Expected Improvement: *P* does so just because *P* judges that ϕ -ing might or will make *S* epistemically better off ([38], 126).

Return to me and James. Suppose I require him to practice his “sight words” a few times a week because I think that such practice is likely to improve his reading skills.¹⁵ And I require him to do so without regard for what I take his wishes on the matter to be. On this definition I act epistemically paternalistically toward him.¹⁶

¹⁴ Though it is outside the scope of this paper to detail my objections to Ryan’s conception of paternalism (that is in turn adopted by Rubel et al.), I can share some examples [9, 38, 39]. This account labels as paternalistic quite a bit of behavior that does not seem paternalistic, precisely in virtue of stripping out Dworkin’s interference condition, a move that Rubel, Castro, and Pham favor [14]. Suppose I stop my car at a red light, not because I am concerned with traffic laws or my own safety but *just because* I do not wish to hit the pedestrian crossing the street, judging that she will be better off if I do not hit her. And suppose it does not matter to me at all whether or not she’d prefer I keep going: that is, I stop irrespective of what I take her wishes to be. It does not seem (to me) that I have acted paternalistically, and yet Ryan’s account will classify such a case as paternalistic.

¹⁵ “Sight words” are the 100–200 most common words in the English language. Students learning to read are taught to learn these words by “sight” rather than by having to sound them out phonetically. This has been shown to help with reading fluency [10, 30, 31]. More information is available here: <https://sightwords.com/sight-words/dolch/>.

¹⁶ Here I gloss over some details, including exactly what is supposed to (according to Rubel, Castro, and Pham) count as “epistemically better off” ([9], 35; [38], 126). As it happens, I think this *isn’t* the right way to understand paternalism. (See footnote 14 above.) But it is a definition that captures much of what many people think pater-

¹³ Cf. Coleman [13] and Lazar [26].

Why think that Click-Gap is epistemically paternalistic according to this definition? Click-Gap was adopted platform-wide, regardless of users' preferences. So it affected the NewsFeed algorithm even for users who would have preferred *not* to have it affect the way information was presented to them, meeting *insensitivity*. Meta claims to have adopted the intervention as part of an initiative aimed at making users better informed [32], and so Click-Gap also seems to meet *expected improvement*. It seems, then, that Click-Gap is epistemically paternalistic, and this is what Rubel, Castro, and Pham conclude [9, 38]. They do not argue that the intervention is impermissible on that basis, however. Rather, they argue that it is an instance of *justified* epistemic paternalism because it is a "policy that all effected [sic] parties could reasonably endorse" ([38], 128).¹⁷ They argue that the policy is not a threat to users' autonomy nor their freedom and that it is implausible to think that the intervention will harm users. So users can reasonably endorse the policy. They further argue that while purveyors might have interests impeded by Click-Gap, the benefits of the intervention far outweigh that interest. And so, in their view, Click-Gap is justified. Because they take Click-Gap to be an instance of epistemic paternalism and they take it to be justified, they conclude that "many epistemically paternalistic policies can (and should) be a perennial part of the internet information environment" ([9], 22; [38], 134). It is this claim that concerns me, as I do *not* think that tech companies are the right sorts of entities we should entrust with treating users in a paternalistic way, let alone endorse them doing so as a perennial part of the internet information environment. To Rubel, Castro, and Pham, what matters to whether an intervention is permissible is not whether or not it is paternalistic but whether or not it is autonomy-violating. While I think they are right to be concerned about autonomy and not about paternalism *as such*, I think they have lost track of why paternalism should concern us in the first place and how that concern can be illuminating in the Click-Gap case. In the next section, I will raise problems both for this classification of Click-Gap

as paternalistic and for using paternalism as a framework for evaluating interventions such as Click-Gap *at all*.

4 Bad Moral Advice

To help introduce and illustrate the main problem with Rubel, Castro, and Pham's analysis as I see it, imagine we adopt an admittedly somewhat contrived version of Meta's perspective. Suppose you are part of a (fictitious) ethics committee in a (fictitious) version of Meta. To keep track of the real and the fake, call the fictitious company MetaE. Suppose that at MetaE the ethics committee has real power within the company. MetaE cares about profits, but it also cares about getting things right, ethically speaking. That is, *if* something is the company's responsibility, they want to do something about it. But they remain unwilling to do just anything to make the world better: they are still a business aiming to make a profit.

Now suppose MetaE has noticed the problems I surveyed above in the introduction, and you (the ethics committee) inform them that they are partly responsible for those problems and thus have an obligation to act. MetaE has agreed to act if you can give them a recommendation for steps to take that would be good. Those steps need not solve all the problems, but they must be morally speaking on the right track. Suppose now the engineers let you know of a potential intervention—Click-Gap. Overall, it seems like it would reduce the spread of fake news on the platform. So you sit down to consider whether you should recommend that MetaE adopt it. But wait! Click-Gap is paternalistic. And why is it paternalistic? It is paternalistic *because our motive is improving our users' epistemic status* without regard for their preferences. It would not be paternalistic if their epistemic status happened not to matter to us, but since it does, Click-Gap is paternalistic. This line of reasoning seems to direct us as an ethics committee either to (1) not care about our users' improvement, (2) look for interventions that are *not* aimed at their improvement, or (3) always give users choices to reject the change. Because paternalism is a motive-based action, if we can just get ourselves not to care about our users' well-being and make the changes totally out of our own self-interest, then we avoid paternalism!

But notice how odd this is. If an agent wants to avoid acting paternalistically, the easiest way for her to do that is *not* to aim at improving the well-being of those around her. And that is absurd advice to give to the reasoning agent. If anything, we should want tech companies to care *more* about improving the well-being and epistemic statuses of their users, not less. And criticizing them when they take steps to make changes to their platforms with their users' well-being in mind seems to undermine precisely this objective.

Footnote 16 (continued)

nalism is about, so it will work for my purposes in this paper. The problems I will raise for (1) understanding paternalism in this way and (2) criticizing someone for acting paternalistically (when this is what we mean by that) generalize to any definition of paternalism that relies on some version of expected improvement that requires P to have the motive of improving S in some way or of making S in some way better off. That is, my argument does not depend on the (not uninteresting) details that I gloss over here.

¹⁷ Thank you to an anonymous reviewer for *AI and Ethics* for pushing me to expand on the nuance in this part of Rubel, Castro, and Pham's argument.

Here one might object that my criticism of the bad moral advice seems to rest on a kind of basic mistake. After all, avoiding paternalism is not the goal in itself. The goal is to act in a way that is all-things-considered justified. If paternalism is sometimes justified, then what will matter is not whether Click-Gap is paternalistic but whether it is justified. One might cap off the objection by pointing out that some of the very critics I discuss here, including Rubel, Castro, and Pham themselves (as we just saw in the previous section), argue that it is an instance of *justified* paternalism.

The objector is right about the critics: they do argue that Click-Gap is paternalistic but justified nonetheless. But I think the reasons to worry remain. If we think that Click-Gap is justified paternalism, we must think two things, each of which seems problematic. First, we must agree that at least sometimes Meta is justified in treating its users paternalistically. Second, we must wonder why we bothered to determine whether Click-Gap was paternalism in the first place. The first is problematic because Meta and companies like Meta are not in the right relationship with their users to be engaged in justified paternalism. Paternalism, if it is ever justified, requires that the agent acting paternalistically be in an appropriate position with respect to the subject of her paternalism: the most accepted case is that of a parent or caregiver and her child. But Meta and companies like it are *not* caregivers of their users. They do not bear the right sort of relation to their users to make paternalistic treatment justified.¹⁸ The second is problematic because it is unclear why we should care about whether something is paternalistic if the fact that it is paternalistic is completely irrelevant to its ethical evaluation.

Fortunately, these two problems point us in a better direction. If we can determine what the core moral concern is when one agent treats another paternalistically, we can allow that concern to be the focus of our evaluation. Below, I will argue that what matters to us when we are worried about paternalistic behavior is not whether a particular action aims at our improvement, not even whether it aims at our improvement while insensitive to our preferences. What matters to us is whether an action constitutes another agent doing something she doesn't have the legitimate agency to do—that is, doing something that is in *my* sphere of legitimate agency rather than her own. And so *that* should be our framework for evaluation, not a focus on paternalism. In the next section, I will say a bit more about what I mean by sphere of legitimate agency and how it can serve as a helpful framework when thinking about the relationship between technology companies and their users. Reorienting

our thinking around interventions on social media platform algorithms is of crucial importance if we are to get the moral evaluations right.

But before I turn to those suggestions, let us notice that it seems unlikely in the non-fictitious case that Meta's Click-Gap intervention actually *was* epistemically paternalistic in the sense captured by the motive-based definition above. For while it seems entirely possible that Meta adopted Click-Gap at least *in part* to make users better off, it seems very unlikely that Meta adopted Click-Gap *just because* they judged that doing so would make users epistemically better off. But this means that Click-Gap does not meet *expected improvement*, since that condition requires that to be paternalistic an intervention be adopted *just because* of the paternalistic motive. This plays right into my complaint: it will turn out that Click-Gap was *not* paternalistic if in addition to being motivated by making users better off (which seems like a good thing to be worried about!), Meta was motivated by making more money or other self-oriented reasons. And *that* makes it seem like we as moral philosophers are *recommending* worrying *less* about others' well-being. This is not what we ought to do.¹⁹

One final objection to my line of thinking here. One might worry that my fictitious MetaE is just reasoning badly. They are playing the game of moral minimalism: what is the least I can do and still meet my moral obligations? And that is not the right way to reason, morally speaking. Rather, one should aim to do good, to act rightly for the right reasons, not just do the bare minimum to avoid paternalism by avoiding acting for the well-being of others. I am not even sure I disagree with the objector here. I do think we should aim to act rightly for the right reasons. Nonetheless, my objection to the paternalism line of reasoning stands. If it does not matter whether we avoid paternalism, if that is not important to us as a goal in itself, then why should we worry at all about whether an action is paternalistic? Why not just move straight to worrying about whether the action is right and / or good? Or, to put it in a way that is grist for my mill, why not worry directly about the underlying moral concern, about what it is that makes paternalism wrong, when it is?

¹⁸ In what relationships might paternalism be justified? I take this question to be both interesting and not fully within the scope of this paper. I take it up more directly in [Redacted]. See also footnote 4.

¹⁹ Relatedly, we might doubt whether Click-Gap meets *insensitivity*, since we have evidence that part of why Meta adopted the intervention was because it was in line with what their research suggests users want to see. But that is not to adopt the intervention irrespective of users' wishes. (I am grateful to an anonymous reviewer for *AI & Ethics* for bringing this point to my attention).

5 What is Wrong with Paternalism?

So one option would be to set aside completely our worries about paternalism and go directly to the question whether or not an action, such as adopting the Click-Gap intervention, is wrong. But the benefit, it seems to me, of concerning ourselves with paternalism in the first place is that such a detour ought to give us resources for making that determination. That is, thinking about whether a particular action is paternalistic is illuminating when and because it helps us to determine whether that action is right or wrong.²⁰ This means that what we need to know is what *underlies* our concern about paternalism. What makes it something we care about in the first place?

To accuse someone of paternalism is to criticize them. This does not *entail* that all paternalistic behavior is wrong, but when we allege that a friend, doctor, teacher has acted paternalistically towards us, we are not giving a non-moralized description of their behavior. Notice that if I require my son James to eat his vegetables with his dinner²¹ or wear his “itchy” coat to school because it is 16°F outside and you allege that I am acting paternalistically toward him, the charge seems out of place. While it may be true, it reads inapt, hollow, a kind of category mistake. I would probably reply that yes, I am acting paternalistically toward him, but since I am his mother, why does it matter? Things seem different, though, if we change the content of the action. Suppose I require James to wear a coat of a particular color rather than another. This behavior strikes us as odd, even morally problematic. We might describe it as overly controlling. And the oddness remains *regardless* of whether I (as his mother) believe that wearing a particular color is in his best interest or just more pleasing to me aesthetically.²²

So even in the core case of a mother interacting with her child, some sort of behaviors undertaken irrespective of the child’s beliefs (*insensitivity*) and to improve the child’s well-being (*expected improvement*) strike us as morally

problematic and others do not. Noticing the difference between these cases is a first step toward understanding the moral concern underlying paternalism. Why does it strike us as problematic when the mother requires the child to wear a certain color but not when she requires a coat? Her child’s wearing a coat when it is cold outside seems like the kind of thing over which her agency legitimately extends. The color of his coat, on the other hand, does not.²³ It isn’t that the mother took the son’s well-being as relevant to her action, for often she should. Nor is it that she acted irrespective of his wishes, for often that is perfectly fine for her to do. What bothers us is that she acted irrespective of his wishes *in a domain of agency* that seems like it should be up to the child. She acted and made decisions within *his* sphere of legitimate agency. Of course, it is often messy and difficult to determine where my legitimate agency ends and yours begins. *Whether* or not to wear a coat is often up to the mother, but there will be temperature ranges where it seems up to the child. What *color* coat to wear is usually up to the child, but there could be cases in which it is up to the mother.²⁴ The point, though, is that *this* seems to be where the questions of moral interest lie: in whose sphere of legitimate agency is this potential action? Who should get to decide whether and how it goes?

When I am concerned that another person has acted paternalistically toward me, I am concerned that she has substituted her own judgment for mine in a domain of agency that is not legitimately under her control. Generalizing this insight, we can say that what makes paternalism something to worry about, morally speaking, is not that one agent has taken the well-being of another as a reason for action, for often they *should* do so. Nor is it that an agent acts irrespective of the wishes of another agent, for often this is perfectly fine to do. What makes paternalism an area of moral concern is that it often involves one agent substituting her own judgment for that of another in a domain of agency that is not her concern.²⁵ Notice that this makes wrongful paternalism *content sensitive*. That is, we will not be able to tell when an action constitutes paternalism in the worrying sense without looking to the content of the action itself. Looking at the relationships between the agents, their beliefs about

²⁰ Pallikkathayil makes a similar point about moral reasoning surrounding coercion [36].

²¹ Interestingly, even a seemingly mundane example like this one faces problems. There is fairly compelling evidence that forcing children to eat their vegetables contributes to an unhealthy relationship with food that can plague them into adulthood [15]. This raises an interesting question about the required degree of *certainty* when acting paternalistically. Of course, most scholars agree that it needn’t be the case that behavior *actually* improve the well-being of the subject (though Dworkin’s view if taken literally *does* have this implication) (2017). And some go further and do not even require full *belief* that it will improve the well-being of the subject, settling for positive epistemic leaning [39].

²² Suppose for the sake of the example that both coats are appropriate attire and that I believe that they are both appropriate. I simply want him to wear blue rather than red.

²³ If the red coat is upstairs, the blue coat already in my hand, and we are on our way out the door, then my insistence on the blue coat may be in my sphere of legitimate control, not because the color of the coat is but because it is my responsibility to ensure we arrive at, say, the doctor’s office on time.

²⁴ Perhaps they are taking photographs as a family, or he needs to wear a certain coat as part of a uniform, or it is a special occasion during which only certain coats are appropriate.

²⁵ This account of what is wrong with paternalism is directly and largely influenced by Shiffrin’s account, but I do not claim she would endorse this way of putting things (2000).

well-being, and their beliefs about each others' wishes may be illuminating, but it will be insufficient to settle the evaluation. We will need to look to the action itself and determine in whose sphere of legitimate agency it belongs.

A final note on this point. If I am right that what matters to us morally speaking about paternalism is respect for an agent's legitimate domain of control, why do so many accounts of paternalism focus on some version of *expected improvement*? Why has it seemed to so many theorists that well-being or improvement is important to the concept of paternalism?²⁶ Here I owe an error theory. I suspect that it is because one of the fundamental commitments of liberalism is that an agent's well-being is up to her. That is, it is within her sphere of legitimate agency to decide what constitutes a good life and how she will pursue it. According to that story, then, another agent's acting just so as to improve her well-being without caring at all what she thinks about it will disrespect her as an agent because it will infringe on matters that are legitimately hers to decide. But well-being is both too narrow and too wide as a proxy for legitimate agency. While it is true that my well-being is often up to me, our lives are fundamentally interconnected. I will not thrive if I am the only one who cares about my well-being. We can agree with liberal commitments about the importance of individual liberty without constraining ourselves away from caring about the well-being of others, so long as we refrain from interfering with matters that are properly theirs to determine.

6 Upshot for Click-Gap

Where does this leave us with Meta and Click-Gap? At first, it might seem like the arrangement of the NewsFeed is something that is completely within Meta's sphere of legitimate agency. After all, they own and operate the platform and have control over its arrangement. That is, they have *actual* control. Does it follow that they have *legitimate* control? It would certainly make the argument here easy and simple if they did. If the NewsFeed were completely within Meta's sphere of legitimate agency, there might be other complaints about the NewsFeed and the algorithm determining its

arrangement we could make, such as worries about sapping our attention or contributing to the spread of misinformation. But it wouldn't be the proper subject of complaints based on worries about paternalism. Why? Because it wouldn't raise the kind of concern that underlies our reason for caring about paternalism in the first place—another agent's interfering with matters that legitimately lie within our control.

Unfortunately for that simple story, I think that the arrangement of the NewsFeed is *not* something that is completely within Meta's sphere of legitimate agency. I think users have claims to it as being part of their domain of legitimate control as well. A full argument in support of this idea is outside of the scope of this paper, but allow me to offer some reasons for thinking that the arrangement of the NewsFeed is in a *mixed* sphere of legitimate agency between individual users, society at large, and Meta as a company. First, Facebook's sheer size is a reason to think society has a claim to control decisions made about the NewsFeed arrangement.²⁷ When something has the potential to have such a large and profound impact on what people think, believe, and value, society has a claim to participation in decisions about it.²⁸ Second, in Facebook, Meta created a dependency from which they benefit. Because users have no other option but Facebook if they want certain valuable goods, they have a claim to some control over the NewsFeed arrangement.²⁹ Third, the very fact that Meta regularly changes the algorithm determining the arrangement of the NewsFeed gives users a claim to some control over it. Users signed up for one thing, began to use it, and contributed to its success. These contributions give them a claim to have a voice when it changes. None of these reasons is meant as a full explanation of the idea that the arrangement of the NewsFeed is in a mixed sphere of legitimate agency. But taken together, they are offered as tentative support of that intuition.³⁰

Suppose I am right that the arrangement of the NewsFeed is within a *mixed* sphere of legitimate agency, belonging

²⁶ For example, Dworkin argues that the paternalist believes her action "will improve the welfare of [her target], or in some way promote the interests, values, or good of [her target]" (2017). Clarke argues that paternalistic behavior involves acting "in order to promote the [target's] good" ([12], 82). Kleinig argues that the paternalist "acts to diminish her [target's] freedom, to the end that [the target's] good may be secured" ([25], 19). Brock defines paternalism as "action by one person for another's good, but contrary to their present wishes or desires..." ([5], 238). More examples include (Husak [21], 40, Husak [22], 467), (de Marneffe [27], 65), (Ben-Porath [4], 249), and (Haybron and Alexandrova [19], 163).

²⁷ In 2021 around seven-in-ten U.S. adults said they use Facebook [16]. That means the arrangement of the NewsFeed affects a tremendous amount of people.

²⁸ How should we understand "society", especially given Facebook's global reach? In the context of this argument I am thinking of U.S. society in particular, but we might wonder if this is the right sphere to have in mind for a multinational corporation. I think fully addressing this question is outside the scope of the paper, but it does seem that the global reach of many big tech companies will impact in whose sphere of legitimate agency we are operating. I am grateful to an anonymous reviewer for *AI and Ethics* for this question.

²⁹ For example, because the platform is quite large, connecting with friends and family may only be possible on Facebook. Sometimes I must use Facebook if I want to connect with businesses that use the platform, or if I am a business that wants to connect with customers. Thank you to an anonymous referee for *AI and Ethics* for bringing some of these valuable goods to my attention.

³⁰ Cf. Lazar [26].

both to the users and to Meta. What is the upshot for Click-Gap? How do we navigate spaces where the decisions properly belong to more than one of us, for us to decide and settle together? I think questions about how we can permissibly navigate spaces of mixed agency are both difficult and pressing: I think that *this* is where the real moral work about the relationship between big tech and its users needs to be done. One thing Meta might owe users is more *choice* and greater *control* over the arrangement of their NewsFeeds. Perhaps users are owed the option to opt out of updates to the algorithm that governs the NewsFeed arrangement, but I am a bit skeptical of a strict right of this kind for several reasons. First, I do not think we have good reason to think that any particular arrangement of the NewsFeed is a default moral baseline to which users are entitled nor a default moral baseline changes to which require our moral attention as opposed to the original arrangement itself. Second, if a change to the algorithm like Click-Gap is being adopted in order to reduce the spread of misinformation and fake news, allowing users to opt out of it could undermine the very goal it sets out to achieve. In fact, Click-Gap seems to fare better here partly *in virtue* of being directed to a societal good—reducing the spread of misinformation and fake news. Of course, we might have further reason to worry, though: is Meta the right kind of entity to be deciding what societal goods are worth pursuing and significant enough to require changes to the algorithm? We might think that the mixed sphere of legitimate agency point actually implies not that individual users should have control over accepting or rejecting interventions on the algorithm but that elected officials or regulators should.³¹

In sum, if I am right that the arrangement of users' NewsFeeds on Facebook belongs to a mixed sphere of agency, then the choice of what that arrangement should look like is not fully within Meta's normative purview.³² Importantly, this view does *not* privilege as a normative baseline any particular arrangement, though. I am not arguing that users are *just* owed consultation or control when Meta makes a *change*. If the arrangement ought to be up to all of us, then this is true of the arrangement all the way down, as it were, not just when Meta adopts a change (whether in its own interest or that of its users or both).

³¹ Cf. Lazar [26].

³² What about completely removing/forbidding/censoring some NewsFeed items, such as e.g., illegal content? Such removal doesn't seem like it would constitute a paternalistic intervention, since it doesn't seem to be in users' sphere of legitimate agency to view, read, or engage with illegal content (though this might depend on a minimally just state). We might also think that in virtue of the content's being determined by the state to be illegal in the first place this already represents respect for the idea that what content to remove is in a mixed sphere of legitimate agency. Thank you to an anonymous reviewer for *AI and Ethics* for raising this point.

How does my analysis here, emphasizing our sphere of legitimate agency, differ from Rubel, Castro and Pham's ultimate analysis ([9, 38])? After all, they end up arguing that Click-Gap is not unjustified because it is not a threat to users' autonomy nor their freedom and that it is implausible to think that the intervention will harm users. In fact, they argue that "[I]imiting persons' ability to expose others to misinformation does nothing to undermine means of social and relational support or create impediments to one's ability to exercise de facto control over their life" ([38], 127). What is the difference between arguing that Click-Gap does not impede upon users' sphere of agency and arguing that it does not impede upon their autonomy? This is a good question.

First, I think it is important to notice that Rubel, Castro, and Pham have divorced their conception of paternalism and their conception of autonomy. So while I argue that my conception of "sphere of legitimate agency" and its infringement underlie our moral concern with paternalism, they think that paternalism is nothing to worry about because it is not connected (directly) with autonomy. Second, they seem to conceive of autonomy as a sort of value that can be maximized, rather than as the kind of value that directs us to respect an individuals' decision-making sphere as being her own. You might support my autonomy (as Rubel, Castro, and Pham understand it) by ensuring that I only ever hear true information. For then my attitudes would be authentically adopted in their view, since they would be attitudes I would "endorse upon critical reflection as consistent with [my] beliefs and values over time" ([38], 123). And yet limiting my access to diverse views and opinions, even if those views are false, might be an impermissible interference with my sphere of legitimate agency. While fully exploring the difference between their conception of autonomy and my conception of sphere of legitimate agency is outside the scope of this paper, I think it comes down to a difference of orientation: they focus on autonomy as a kind of value that can be increased or decreased, and I think of sphere of legitimate agency as a kind of sphere of control that others are not allowed to impinge, even if doing so might in some way make me more autonomous.

Let us consider one non Click-Gap example so that we can see the implications of my argument outside of this particular intervention. Consider YouTube, an online video sharing and social media platform. YouTube has adopted Community Guidelines and uses AI to aid in its enforcement of those guidelines. From their Community Guidelines FAQ:

YouTube's automated flagging systems start working as soon as a user attempts to publish a video or post a comment. The content is scanned by machines to assess whether it may violate YouTube's Community Guidelines. YouTube also utilizes automated systems to prevent re-uploads of known violative content,

including through the use of hashes (or “digital fingerprints”) [44].

Should we think that adopting Community Guidelines is something within YouTube’s sphere of legitimate agency? It seems we have good reason to think that no, adoption of such guidelines is not solely within the company’s sphere of legitimate agency. Rather, it also belongs to users to help determine what those guidelines should be and how they should be enforced. For many of the same reasons I articulated above when discussing Meta. I think we have good reason to believe that YouTube Community Guidelines are the sort of thing that is in a *mixed* sphere of legitimate agency between individual users, society at large, and YouTube’s parent company (Alphabet Inc.). As with Meta, I take it that a full argument in support of this idea is outside of the scope of this paper *and* the right place to direct our argumentative energy as we consider what ethical social media might look like. As with Facebook, YouTube is quite large, with an estimated 2.7 billion active users [40]. When something has the potential to have such a large and profound impact on what people think, believe, and value, society has a claim to participation in decisions about it.³³ As with Facebook, YouTube offers certain valuable goods, like connecting with other content creators, that are not available elsewhere. As with Meta, I do not intend any of these reasons as a full explanation of the idea that the arrangement of the NewsFeed is in a mixed sphere of legitimate agency. But taken together, they are offered as tentative support of that intuition.³⁴

If I am right that YouTube’s Community Standards are part of a mixed sphere of legitimate agency, then standards will have to be determined *together*. Exactly what shape that ought to take is not determined by my analysis here, but YouTube will not be able simply to adopt the standards that it deems best without actual empowered input from users. Not because it is problematic for YouTube to adopt such standards in order to benefit users or to make them better off nor because it is permissible for YouTube to have no standards and maintain the status quo. There is no presumption in my analysis that having no standards is somehow the ethically superior baseline to work from. But to properly respect the agency of its users, YouTube would need to find a way to incorporate their input in a meaningful way into determining its policies and their enforcement.

7 Upshot for Big Tech

What conclusions, if any, can we draw about the relationship between big tech and its users generally speaking? We might have thought that technology platforms are fully within the legitimate sphere of agency of the companies that control them. That is, we might have thought that companies can do as they want to their technologies: change the algorithm, update the application, pursue different goals or plans. Users always have a choice: they can choose to use the technology or not. If a user is unhappy with a change or adjustment, they can switch to a different technology. Allow the free market and competition, as it were, to offer feedback to companies about the changes they make. The criticism from paternalism already problematizes this model: it directs companies not to make changes with the goal of promoting their users’ well-being if those changes are also insensitive to users’ preferences. I argued that the criticism misses the mark and that it is problematic because it directs companies *not* to take user well-being as a direct reason for action.

If my argument succeeds, then, tech companies *can* (and should!) take their users’ well-being as direct reason for action. But the cards are not all theirs to play. While they can promote user well-being, some parts of their technologies are not fully within their legitimate sphere of agency, in part due to size, in part due to societal dependence, in part due to users’ crucial role in building the technologies into what they are. And when a technology is no longer fully within the legitimate sphere of agency of the company, it ought to act differently in how it conceives of who that technology belongs to and who gets a say in where that technology goes next.

Exactly how to determine when a particular technology enters the *mixed* sphere of legitimate agency strikes me as a very difficult question, one we will have to answer individually based on the content and context of that technology. Will ad recommenders be something over which companies have sole legitimate agency and need not accept input from users or society? I am skeptical, at least when it comes to ads offered on platforms as large as e.g., Facebook and YouTube. The size and attention these companies command give us good reason, it seems to me, to think that many decisions they are currently making in isolation belong, instead, to all of us.³⁵

³³ Cf. Lazar [26].

³⁴ Cf. Lazar [26].

³⁵ Cf. Lazar [26].

8 Conclusion

In the end, Facebook users do not have a complaint against Click-Gap (neither as paternalistic nor all-things-considered), but they do have a complaint *in general* that they are not being given opportunities to have (individual) control over the NewsFeed arrangement nor (collective) say/impact on the algorithm.

Importantly, notice that it does not matter to our evaluation whether Newsfeed arrangement algorithm intervention is undertaken *in order to make users better off*. And as we have seen, this is good for a few reasons. First, we will often be unable to know or determine what the reasons are that companies are undertaking changes or interventions in this space. They may *say* that it is for the benefit of their users, but we often may suspect an (at best) mixed motive. Second, we do not have good reason to pay special moral attention to interventions undertaken for any particular reason rather than another. We should treat all interventions as worthy of evaluation and critique, regardless of their proffered explanations.

What is the upshot of a framework of evaluation for our fictitious MetaE and the ethics committee on which we are pretending we sit? The news is both good and bad: we still face complicated questions about our impact on the world, what we are responsible for, and whether interventions we want to undertake are in our sphere of legitimate control or another's. But we do not have to try to *avoid* acting in the best interest of our users: we can take making them better, making the user experience better to be directly important to us, not just instrumentally important or among our at-best mixed motives. And that, I think, is a very good thing.

Acknowledgements I am grateful for help and critical comments from Jeffrey Behrends, Susan Brison, Willian Cochran, Trystan Goetze, David Gray Grant, Ryan Jenkins, Amber Kavka-Warren, and C. Thi Nguyen. Previous versions of this paper were presented at the Cyberethics Forum Works in Progress Group, the Harvard / MIT Conference on Ethics of Computing & AI, the Dartmouth Workshop on Ethics and Information Technology, and the Philosophy, Politics, and Economics Society Seventh Annual Meeting, and I am grateful to the audiences at all of those events for stimulating discussion and pressing questions that were tremendously helpful to my thinking about these issues. I have benefited enormously over the years from conversations about paternalism, moral complicity, and interpersonal moral responsibility with Kevin Donohue, Daniela Dover, Barbara Herman, Pamela Hieronymi, Gabrielle Johnson, A. J. Julius, Ally Peabody-Smith, and Seana Shiffrin. Finally, I want to acknowledge the detailed and helpful comments received from two anonymous referees and the editors of *AI and Ethics* that greatly improved the paper.

Data availability No data was used for the research described in the article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source,

provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahlström, K.: Epistemic paternalism: a defence. Palgrave Macmillan, Houndmills (2013)
- Allcott, H., Gentzkow, M., Chuan, Yu.: Trends in the diffusion of misinformation on social Media. *Res. Politics* **6**(2), 2053168019848554 (2019). <https://doi.org/10.1177/2053168019848554>
- Anderson, K.: Truth, lies, and likes: why human nature makes online misinformation a serious threat (and what we can do about it) student articles. *Law Psychol. Rev.* **44**, 209–244 (2019)
- Ben-Porath, S.: Paternalism, (school) choice, and opportunity. In: Coons, C., Weber, M. (eds.) *Paternalism: theory and practice*, pp. 247–65. Cambridge University Press, Cambridge (2013)
- Brock, D.: Paternalism and promoting the good. In: Sartorius, R.E. (ed.) *paternalism*, pp. 237–260. University of Minnesota Press, Minneapolis (1983)
- Brown, É.: Propaganda, misinformation, and the epistemic value of democracy. *Crit. Rev.* **30**(3–4), 194–218 (2018). <https://doi.org/10.1080/08913811.2018.1575007>
- Brown, É.: Fake news and conceptual ethics. *J. Ethics Soc. Philos.* **16**, 144 (2019)
- Brown, É.: Free speech and the legal prohibition of fake news. *Soc. Theory Pract.* **49**(1), 29–55 (2023). <https://doi.org/10.5840/soctheorpract202333179>
- Castro, C., Adam, P., Alan, R.: Epistemic paternalism online. In: Bernal, A., Axtell, G. (eds.) *Epistemic paternalism: conceptions, justifications, and implications*, pp. 22–44. Rowman and Littlefield International, Lanham (2020)
- Chall, J.S.: *Stages of reading development*. McGraw-Hill (1983)
- Chen, A.: What Mueller's indictment reveals about russia's internet research agency. *The New Yorker*, 16 Feb 2018. sec. tags. (2018). <https://www.newyorker.com/news/news-desk/what-muellers-indictment-reveals-about-russias-internet-research-agency>
- Clarke, S.: A definition of paternalism. *Crit. Rev. Int. Soc. Pol. Phil.* **5**(1), 81–91 (2002). <https://doi.org/10.1080/13698230212331294554>
- Coleman, M.C.: Attempting to stop the spread: epistemic responsibility and platformed responses to the COVID-19 'Infodemic.' In: Zeiser, P., Beasley, B. (eds.) *Social media ethics and COVID-19: well-being, truth, misinformation, and authenticity*, pp. 9–30. Lexington Books, Lanham (2023)
- Dworkin, G.: Paternalism. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2017. Metaphysics Research Lab, Stanford University. (2017). <https://plato.stanford.edu/archives/spr2017/entries/paternalism/>
- Gouveia, M.J., Canavaro, M.C., Moreira, H.: How can mindful parenting be related to emotional eating and overeating in childhood and adolescence? The mediating role of parenting stress and parental child-feeding practices. *Appetite* **138**(July), 102–114 (2019). <https://doi.org/10.1016/j.appet.2019.03.021>
- Gramlich, J.: 10 Facts about Americans and Facebook. Pew Research center (blog). 1 Ju 2021. (2021). <https://www.pewresearch.org/fact-tank/2021/06/01/facts-about-americans-and-facebook/>

17. Habgood-Coote, J.: Towards a critical social epistemology of social media. In: J Lackey, A McGlynn (eds) *Oxford Handbook of Social Epistemology* (Forthcoming)
18. Handfield, T.: Regulating social media as a public good: limiting epistemic segregation. *Soc. Epistemol.* (2023). <https://doi.org/10.1080/02691728.2022.2156825>
19. Haybron, D.M., Anna, A.: Paternalism in economics. In: Coons, C., Weber, M. (eds.) *Paternalism: theory and practice*, pp. 157–177. Cambridge University Press, Cambridge (2013). <https://doi.org/10.1017/CBO9781139179003.009>
20. Hagey, K., Horwitz, J.: Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead. *Wall Street JI*, 5 Sep 2021, Sec. Tech. (2021). <https://www.wsj.com/articles/facebook-algorithm-change-zuckerberg-11631654215>.
21. Husak, D.: Penal paternalism. In: Coons, C., Weber, M. (eds.) *Paternalism: theory and practice*, pp. 39–55. Cambridge University Press, Cambridge (2013). <https://doi.org/10.1017/CBO9781139179003.003>
22. Husak, D.: Paternalism. In: Marmor, A. (ed.) *The Routledge Companion to Philosophy of Law*, pp. 467–480. Routledge Philosophy Companions, Routledge (2015)
23. Jackson, E.: What's epistemic about epistemic paternalism? In: Matheson, J., Loughed, K. (eds.) *Epistemic autonomy*, pp. 132–150. Routledge (2022)
24. Kitsik, E.: Epistemic paternalism via conceptual engineering. *J. Am. Philos. Assoc.* (2022). <https://doi.org/10.1017/apa.2022.22>
25. Kleinig, J.: *Paternalism*. Rowman and Allanheld, Totowa (1984)
26. Lazar, S.: AI and human values. The tanner lectures on human values (2023) <https://hai.stanford.edu/events/tanner-lecture-ai-and-human-values-seth-lazar>
27. de Marneffe, P.: Self-sovereignty and paternalism. In: Coons, C., Weber, M. (eds.) *Paternalism: theory and practice*, pp. 56–73. Cambridge University Press, Cambridge (2013). <https://doi.org/10.1017/CBO9781139179003.004>
28. Medvecky, F.: Epistemic paternalism, science, and communication. In: Axtell, G., Bernal, A. (eds.) *Epistemic paternalism: conceptions, justifications and implications*, pp. 79–89. Rowman and Littlefield International, Lanham (2020)
29. Millar, B.: Misinformation and the Limits of Individual Responsibility. *Soc. Epistemol. Rev. Reply Collective* **10**(12), 8–21 (2021)
30. Morris, D., Trathen, W., Gill, T., Schlagal, R., Ward, D., Frye, E.M.: Assessing reading rate in the primary grades (1–3). *Read. Psychol.* **38**(7), 653–672 (2017). <https://doi.org/10.1080/02702711.2017.1323057>
31. Morris, D., Trathen, W., Perney, J., Gill, T., Schlagal, R., Ward, D., Frye, E.M.: Three DIBELS tasks vs. three informal reading/spelling tasks: a comparison of predictive validity. *Read. Psychol.* **38**(3), 289–320 (2017). <https://doi.org/10.1080/02702711.2016.1263700>
32. Mosseri, A.: Working to Stop Misinformation and False News. Meta Newsroom (blog). 6 Apr 2017. (2017). <https://about.fb.com/news/2017/04/working-to-stop-misinformation-and-false-news/>
33. Nguyen, C. T.: Escape the Echo Chamber. Aeon. (2018). <https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult>. Accessed 29 Oct 2020
34. Nguyen, C. T.: Echo chambers and epistemic bubbles. *Episteme* **17**(2), 141–161 (2020). <https://doi.org/10.1017/epi.2018.32>
35. Oremus, W., Alcantara, C., Merrill, J. B., Galocha, A.: How facebook shapes your feed: the evolution of what posts get top billing on users' news feeds, and what gets obscured. *The Washington Post*, 26 Oct 2021. (2021). <https://www.washingtonpost.com/technology/interactive/2021/how-facebook-algorithm-works/>
36. Pallikkathayil, J.: The possibility of choice: three accounts of the problem with coercion. *Philosophers' Imprint* **11** (2011)
37. Pariser, E.: *The filter bubble: what the internet is hiding from you*. Penguin Press, New York (2011)
38. Rubel, A., Castro, C., Pham, A.: *Algorithms and autonomy: the ethics of automated decision systems*. Cambridge University Press, Cambridge (2021). <https://doi.org/10.1017/9781108895057>
39. Ryan, S.: Paternalism: an analysis. *Utilitas* **28**(2), 123–135 (2016). <https://doi.org/10.1017/S0953820815000254>
40. Shewale, R.: YouTube Statistics For 2023 (Demographics & Usage). *DemandSage* (blog). 5 Sep 2023. (2023). <https://www.demandsage.com/youtube-stats/>
41. Shiffrin, S.V.: Paternalism, unconscionability doctrine, and accommodation. *Philos. Public Aff.* **29**(3), 205–250 (2000)
42. Sunstein, C.R.: *#Republic: divided democracy in the age of social media*. Princeton University Press (2018). <https://doi.org/10.2307/j.ctv8xnhtd>
43. World Health Organization: Coronavirus Disease 2019 (COVID-19) Situation Report. 86. World Health Organization. (2020). <https://www.who.int/publications/m/item/situation-report--86>. Accessed 3 Oct 2023
44. YouTube: YouTube Community Guidelines Enforcement FAQs. Transparency Report Help Center. (2022). <https://support.google.com/transparencyreport/answer/9209072?hl=en>. Accessed 3 Oct 2023

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.