TikTok and Autonomy: Exploring the Effects of Algorithmic Paternalism on Well-being

Bachelor Thesis for the Bachelor programme Philosophy, Politics and Society

at the Faculty of Philosophy, Theology and Religious Studies

at Radboud University, Nijmegen

First Reader: Linde van Schuppen

Number of Words: 5999

Submission Date: 20.06.2023

Aaron von dem Bussche

Hereby, I, Aaron von dem Bussche, declare and ensure that this bachelor thesis titled *TikTok and Autonomy: Exploring the Effects of Algorithmic Paternalism on Well-being*, has been entirely written by me, that I have not used any other sources or tools than mentioned here, and that the passages in this work of which the verbatim content or meaning – including from electronic media – has been taken, have been referred to.

Introduction

It is easy to get lost in light entertainment. Quickly checking the phone once can lead to several hours of lost time in what feels like the blink of an eye, leading to a day spent without feeling any real accomplishment. Time spent on social media feels good while it is happening, with no strong emotions either way. However, it might lead to regret when it is discovered that while one could have done errands, exercised, or learned during that time or at least enjoyed forms of entertainment considered of higher quality, such as books, movies, or music. After all, time spent browsing timelines and frontpages is seldom remembered in detail, nor does it cause us to reflect consciously.

Nevertheless, it is strangely appealing, a space where we can do nothing without being bored, that presents us with things we enjoy (or at least are somewhat interested in) and never ends. If one subconsciously opens post for post, video for video, it might be comforting. However, upon short reflection, it seems eerie if not scary: These social media platforms know a lot about one's interests, maybe more than one would be able to tell a stranger and certainly more than one would like them to know.

By mindlessly consuming what the algorithm presents them with, the user gives away their choice to select content themselves out of convenience. TikTok has famously perfected this style of content consumption, advertising that “All you have to do is watch, engage with what you like, skip what you don’t, and you’ll find an endless stream of short videos that feel personalized just for you” (‘TikTok App Store Page’ 2023). Users praise its algorithm, dubbed the 'secret sauce' for its success. However, of course, behind the fun there lies a profit motive.

The transmission of increasing amounts of personal data and accompanying delegation of choice in what to consume to information technology companies dangerously undermines the individual’s autonomy. This can, at least in certain conceptions of the concept, greatly diminish the individual’s well-being.

The academic field concerned with the societal implications of information technology has been steadily growing. However, autonomy still only fills a niche, as most researchers are foremost troubled by privacy. So far, autonomy research has focused on the rather abstract and gone in the direction of giving recommendations for political legislature. Meanwhile, this thesis fills a gap in that it focuses on concrete implications to the individual’s autonomy and well-being through the example of TikTok. It introduces the concept of *algorithmic paternalism*, which has not previously existed in this field, and sets out to answer the question: *How does social media’s algorithmic paternalism harm user’s well-being*?

To answer this question, I will first introduce philosophical theories of well-being, autonomy, paternalism, and the connected concept of the *nudge*. Then, I will conceptualise my term *algorithmic paternalism* and present the digital evolution of the *nudge,* the *hypernudge*, utilised in the following

argumentative section. In the argumentative section, I lay out, working closely with the example of TikTok, how social media platforms impact individuals and can diminish their autonomy. Lastly, I will specify the implications of such diminished autonomy in the case of TikTok to individual well-being and summarise the paper's content.

Well-being

Theories of Well-being

Defining the 'correct' concept of well-being is a broad, fundamental philosophical issue. The question of what it takes to live well was already a disagreement between Plato and Aristotle. While theories of well-being nowadays have certainly evolved and are more comprehensible and extensive than back in ancient Greece, most ideas at the core of popular theories since then have remained stable. Most commonly, theories of well-being are divided into three categories: Hedonistic-, desire-, and objective list theories (Crisp 2021). In the following paragraphs, I will quickly summarise their main points.

Hedonistic theories

Hedonistic theories of well-being state that what is fundamental is the experience of pleasure and the absence of pain (Bentham 1970, 6). An individual's pleasure and pain are measured and weighed against each other, and if the pleasure outweighs the pain by a good margin, that person can be said to be happy, which is enough to constitute well-being. For the hedonistic model, it does not matter *why* a person is experiencing good feelings instead of negative ones. Under this theory, a person can be said to be living well even if their sources of happiness seem nonsensical to others. In this simplicity, which allows everyone to orient their life to their individual preferences, lies the appeal of this theory.

Hedonistic well-being has been attacked for prioritising short-term pleasure over long-term happiness. Another critique, prominently posed by Mill, a contemporary of Bentham, is that human flourishing does not just constitute a singular pleasure, but that certain pleasures, such as intellectual pursuit, possess a certain quality that makes them more desirable (Crisp about Mill 1997, 30). An answer to the first critique is that well-being is meant to be measured throughout a lifetime, not only for short amounts of time. To maximise well-being, it is wise to think of ways to *sustainably* maximise pleasure (and minimise pain) in the future, even if it means taking the more painful road short term. The answer to the second critique is less satisfying for some: Much to the displeasure of some philosophers like Mill, there is no further appreciation for fine arts, education, or other signs of quality and virtue in Bentham's hedonism. If people live happy but uneducated lives, hedonism has no problems calling them good lives.

Desire theories

Desire theories, also called desire satisfaction theories, state that an individual's well-being is determined by how successfully their desires, goals, or preferences are satisfied. Such desires could be having a successful career, good social life, or whatever else an individual wants. While most desire theories accept the concrete desires of a person, other theories define desires as what one *would want* if one were perfectly informed. Either way, desire theories, like hedonistic theories, are subjective. Therefore, whether a person's desires are reasonable or beneficial for them is irrelevant if these desires are fulfilled according to the theory. That said, desire theories are less permissible than hedonistic theories, as they demand that certain conditions, though chosen authentically by the individual themselves, are satisfied for them to experience well-being (Crisp 2021).

Objective list theories

Objective list theories disagree on the last point. As their name states, they are objective theories and leave less or no room for an individual's preferences. Objective list theories usually come with their own *(objective) list* of goods and goals everyone should try to attain. If the goals are being fulfilled, an individual gains well-being, else they do not. Many theories also distinguish between necessary and recommended goals, in which a life cannot be considered well lived if necessary goals are not accomplished, while recommended goals only improve well-being but are not necessary for a good life. Necessary goals could be the fulfilment of basic human needs, and recommended goals could be a fitting and successful career, artistic expression, happy family life and the like. Knowledge and friendship (Crisp 2021) could be in either. While the specific goals differ wildly between objective list theories and can include somewhat subjective goals open for interpretation based on personal preference and context, the objective nature of objective list theories is its main critique point. After all, humans have distinct personalities, cultures and preferences and might not all become happy by following the same goals.

Concerning autonomy, these theories diverge significantly. While hedonistic theories usually do not value autonomy as an intrinsic good, desire and objective list theories often consider it fundamental for human well-being. Desire theories do so as the free choice and pursuit of desires are at the heart of the theories. Without autonomy, there would be no way to live by one's authentic desires. In objective list theories, autonomy is usually an intrinsic, necessary good without which the good life cannot be attained. Therefore, while hedonistic theories do not value autonomy apart from possible effects on pleasure, desire theories see it as instrumental for well-being. In contrast, objective list theories usually consider it necessary and intrinsically valuable.

Autonomy and Paternalism

Autonomy

The concept of personal autonomy in philosophy is ambiguous. Many different fields and currents have conceptualised autonomy in very different ways, often allowing very different conclusions as to what constitutes autonomy concretely. To analyse a specific situation, at least a dozen philosophical definitions of autonomy can reasonably be considered. At least the basic idea behind the meaning of autonomy is relatively easy to explain: the idea that people "should make their own lives" (Raz 1986, 369). In general, people can and should make small and big life decisions for themselves and their own reasons.

Nevertheless, what exactly constitutes a free choice is unclear. Some claim that a lack of options for disadvantaged people constitutes non-autonomy. They own little monetary or other means; therefore, their availability of meaningful choices is much restricted. On the other hand, others claim that even if one is held at gunpoint and made to hand over one's wallet or is physically restrained, one is still autonomous if one's brain is not directly being interfered with. In that interpretation, it is not the meaningful capacity to exercise choice that matters but the *capacity to choose* alone (Husak 1981, 36).

Because finding a definite concept of autonomy is not possible, this paper will look at a contemporary conception of it that has been successfully applied in the field of philosophy of technology. Susser, Roessler and Nissenbaum, leading researchers in their field, devised a fitting model in their 2019 paper dealing with online manipulation, its consequences, and policy recommendations to combat it. They point to the frequently present distinction between competency and authenticity conditions in autonomy research. Competency means that to possess autonomy, one needs to have the “cognitive, psychological, social, and emotional competencies” (Susser, Roessler, and Nissenbaum 2019, 8) to evaluate one's choices, form intentions about them and bring these intentions to action. The authenticity condition denotes that one must critically reflect on one's values, desires, and goals and endorse them as authentically one's own.

The authors concede that this definition could be seen as overly rationalistic, as people rarely follow these exact steps to make their decisions, often following along with their habits without second thoughts. Furthermore, humans are too emotional to consider each action they take rationally. Susser et al. (2019) confirm the reasonableness of these objections, as in their paper, they clarify that their model is not to be taken as strict as it might seem at first glance: not only do they not just include beliefs and desires, but also emotions, convictions and experiences as valid aspects of evaluation, but they also do not demand the evaluation process mentioned above to take place before every action a human takes. Instead, it is enough if a person can evaluate their motivations and carry them into everyday life, even if the actions happen impulsively.

In their paper, the authors also state their agreement with feminist, relational approaches to autonomy and, therefore, against atomistic ones. Relational autonomy is the explicit formulation of the conviction that every person is embedded deeply in society, including social, cultural, historical, and political influence in their identity and possibilities. Social determinants, such as race, class, gender and ethnicity, profoundly impact the intersubjective and social dimensions of selfhood and identity (Mackenzie and Stoljar 2000, 4). It is accepted that personal autonomy is never absolute, as one is always in some ways limited by society, relations, obligations, and natural reality. Rather than to avoid all societal obligations, to live autonomously means to make one's own choices within society freely, form relations and overcome obstacles.

Paternalism and Nudges

Gerald Dworkin, the paradigmatic philosopher of paternalism, defined paternalism as “interference with a person's liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interests, or values of the person being coerced” (Dworkin 1971, 108). Dworkin makes distinctions between justified and non-justified paternalism. In his view, paternalism inhibits an individual's autonomy. However, it can be justified if it "preserves and enhances for the individual his ability to rationally consider and carry out his own decisions" (Dworkin 1971, 108). Following this, examples like paternalism preventing suicide or injury would be justified in all but edge cases (Husak 1981, 34), including laws mandating seatbelt use. It is also usually permissible to act paternalistic towards one's children. Therefore, while autonomy is, in his view, an essential good to be protected, people's autonomy can be violated in some instances, namely if it is reasonable to expect their autonomy to benefit from it ultimately.

There is an influential alternative model describing some cases of paternalism: Libertarian paternalism. Thaler and Sunstein (2003) define a policy as paternalistic if it is "selected with the goal of influencing the choices of affected parties in a way that will make those parties better off". What separates their model from Dworkins is that they do not believe in the necessity of coercive interference in a person's liberty of action for something to be paternalistic.

Paternalism can mean minor influences, *nudges*, such as placing food at a buffet in the order of healthy to unhealthy to make more people choose the healthy food. A nudge is a form of choice architecture that, as 'soft' design-based control, does not forbid the user any options while still altering behaviour in predictable ways. Nudges work by intentionally offering a "path of least resistance" to one or more options over others in a choice environment, making them the 'default options' (Thaler and Sunstein 2008, 83). Research shows that any arrangement of choices makes some choices more or less likely to be picked (Dworkin 2020), making nudges an impactful tool to control human behaviour.

Thaler and Sunstein also make the point that for many policy decisions, there is no viable alternative to this type of paternalism; is it not much more sensical to make decisions in the people's best interest instead of randomly or even against them? After all, almost everywhere a choice can be made, there is such a path of least resistance (Thaler and Sunstein 2008, 83). Since the advent of this concept, nudges have become increasingly popular in policy advice. The selling point of a nudge from the perspective of Thaler and Sunstein is that libertarian paternalism influences people for the better without infringing on their freedom of choice.

The following section introduces my concept of algorithmic paternalism. It applies it to the case of TikTok, examining the effects the use of TikTok can have on autonomy and well-being.

Examining TikTok's Algorithmic Paternalism

Algorithmic Paternalism and related concepts

Social media such as TikTok engages in what I term *algorithmic paternalism* when selecting content on behalf of a user's perceived interests. Algorithmic paternalism is the *use of algorithms to guide and influence user behaviour in a way that is claimed, perceived, or aimed to be in the user's best interest*. Algorithmic paternalism is a broad term encompassing a field of more specific tools and technologies, such as automated spam moderation, nudity filters, reminders, support chatbots, and nudges. Algorithmic paternalism is not a weighed term in itself, and many of its uses could be described as user-friendly design or as necessary for the functioning of certain web services. Algorithmic paternalism draws from a variety of similar, contemporary notions, but it has not previously been precisely formulated.

The exact term *"algorithmic paternalism"* has so far only been used once, in the field of medical research, substituting "medical paternalism" now that increasingly doctors' decisions are replaced by machine learning tools (McCradden and Kirsch 2023). Medical paternalism describes the antiquated, negatively connotated practice of a medical professional making choices on behalf of patients without their consent or involvement in the decision-making process. Brown (2020), on the other hand, speaks of *digital paternalism*, a concept that seems promising but is not sufficiently defined to be utilised in this paper. Additionally, I want to explicitly draw attention to the algorithms at the centre of the issue and therefore choose to go by *algorithmic paternalism* instead.

Yeung (2017) expands on the use of nudges, introduced as a concept earlier as an example of libertarian paternalism in the digital age. It is easy to see that nudges can be powerful on digital devices: After all, the arrangement of choices is fundamental to the structure of websites and apps. However, more importantly, since the popularisation of nudges by Thaler and Sunstein (2008), nudges have evolved from being static to being dynamic: Yeung describes the *hypernudge* as a potentially

much more invasive and manipulative type of digital intervention that utilises constantly updated, personalised, and targeted choice architecture.

Hypernudges can be used in various contexts, including advertisements, shopping sites, gambling sites, healthcare applications and social media. Whenever content, ads, products, or links are dynamically placed in a social media timeline or on a website, a hypernudge takes place. This can include innocent uses, like putting cute cat pics in a cat lover's timeline and nefarious uses, like presenting someone with money troubles ads for high-interest loans. While the nudge is described as 'soft' control, hypernudges can be extraordinarily strong, potentially being able to be hyper-personalised and utilise intimate details and weak points about their target. They can also reach millions of users simultaneously. Provided their use is claimed, perceived, or aimed to be in the user's best interest, hypernudges are one of the most abundant types of algorithmic paternalism.

Some have criticised nudges on account of using techniques that "deliberately seek to exploit cognitive weaknesses to provoke desired behaviour" (Yeung 2017, 7) and therefore, at least in cases in which not each individual nudge is transparently explained to the nudged, as unwarranted manipulation (Bovens 2008, 15). Thaler and Sunstein (2008), on the other hand, argue that this is not the case, as when following their guidelines and using nudges for government intervention, nudges involve neither coercion nor deception and are transparent in their intention to influence behaviour in a given direction. While they do not explicitly use the word manipulation, they are adamant that "freedom of choice" is always preserved, if not expanded, by their measures (Thaler and Sunstein 2008, 237). It is straightforward to say that at least in cases where nudges do not follow Thalers and Sunstein's guidelines and are not transparently explained to the nudged, a deceptive, manipulative element can exist.

Social media, Algorithms and TikTok

Social media networks are nowadays far more than just places to connect with friends and family. Behind the companies in the field worth hundreds of billions (Isaac 2021) stands a profit motive, primarily fuelled by advertising. The ARPU, average revenue per user for Facebook worldwide, was reported as $7.89 in the third quarter of 2020, extrapolating to $31.56 a year. In the US & Canada region, it even reached a whopping $158.52 a year (Gao 2020). As companies are always incentivised to maximise their profit, which is generated by showing users advertisements, social media companies are competing for the screen time of their users. In an *Attention Economy* (Davenport and Beck 2001), they have long understood that the best way to keep users active and attentive on the platform is by bringing users engaging content via algorithms. 'Engaging' does not need to mean pleasurable, as Facebook in the past has been found to have been disproportionally promoting content that enrages people instead (Merrill and Oremus 2021).

In general terms, algorithms are not more than a set of instructions that a computer can follow, whether simple or complex. In social media, these algorithms determine which content is shown to specific users and in which order. These algorithms are fed with data about the user, such as their behaviour on the platform, to analyse users' interests and show them content popular with users deemed similar. The analysis of users goes much more thorough than one might expect: It is not only surface-level platform behaviour, such as likes, that is analysed, but entire profiles are created about users, identifying their demographic, interests, and even political leaning (Merrill 2016). Exemplary, a lifestyle blog titled "The TikTok Algorithm Knew My Sexuality Better Than I Did" (MacGowan 2020). Of course, this information is not only used for targeted content but also targeted advertisement.

The pinnacle of the development to bring content to the user is TikTok. For the time, it has perfected algorithms to show users engaging content. Users searching for entertainment flock to the platforms with the 'best' algorithms giving them the most enjoyable mix of content. While other social media companies offer users more control to choose the content they consume, such as YouTube providing a selection of around eight algorithm-selected videos on desktop and Instagram initially displaying content from explicitly followed creators in the timeline, TikTok eliminates these options. On the "For You" page of the TikTok app, the user gets spoon-fed content one by one across the entire screen of their device. Many users praise TikTok's algorithms for consistently showing them content they enjoy better than other social media sites and media services. Meanwhile, according to experts, while TikTok uses more trackers than other social media platforms and has a very refined algorithm, it does not do anything extraordinary on the technical site that other social media networks do not (Smith 2021). However, its design puts selected content front and centre, making it the prime example to show the power of *algorithmic paternalism*.

Effects of TikTok on the User’s Autonomy

As required to fulfil the definition of algorithmic paternalism, TikTok claims that its algorithm's use is in its users' best interest. It promises to bring the user content that will "make their day" and promises "real, interesting, and fun" videos that "entertain and inspire" and "feel personalized just for you" ('TikTok App Store Page' 2023). TikTok assures that it will give the user short-term fun, and perhaps even lasting inspiration for outside the app. After all, that is what users visit the platform for. How are users impacted in practice?

To start with the positive effects, the algorithm does somewhat deliver on its promises. With its help, users can find content that they enjoy, leading to a temporary gain in hedonistic well-being. Through TikTok's algorithm's analysis of users, they can be matched with comforting content and creators that they can relate with and find supportive communities with potentially helpful advice applicable in real life. While these contents can be helpful in general, they can be especially valuable for juveniles that have trouble finding support and relating with their peers and families. For example, a queer kid growing up in a homophobic setting might be presented with affirmations and relatable content of other kids in the same position, which can give them the strength to believe in themselves and better deal with their surroundings. In this way, TikTok and its algorithms can form a safe haven and support network for people, letting them escape everyday life's hardships and restrictions.

On the other hand, these communities and bubbles are not always a positive influence. An especially gruesome example is the case described in an article titled "TikTok's Algorithm Keeps Pushing Suicide to Vulnerable Kids" (Carville 2023). While the topics the now deceased boy explicitly searched for were relatively positive or neutral (sports, batman, motivational speeches), his *for you* page picked by the algorithm showed "a steady stream of videos about depression, breakups, death and suicide". The algorithm likely reinforced itself on continuous subtle indicators in the boy's swiping behaviour on such topics. Another reminder that "engaging" can also mean "controversial or toxic" lies in the success of Andrew Tate on the platform, spreading misogynistic content that is highly influential with schoolboys, especially those ages 13-15, who are deemed especially vulnerable (Weale 2023). These examples show that the algorithm can negatively influence users and whole generations by introducing them to topics and reinforcing beliefs in which they are not explicitly interested.

Next, time spent on TikTok is not usually productive time. While TikTok can be used to access self-help resources, workout and cooking advice, and other content perhaps deemed useful, I argue that scrolling usually does not contribute as much to long-term growth and well-being as other activities could. For hedonism, this argument could be formulated as *consuming TikTok now for temporary pleasure is not worth it if it leads to diminished pleasure long-term;* for desire- and objective list theories, it could be that *consuming TikTok now stands in the way of fulfilling one's desires or goals, and is therefore not worth it*. Of course, this point can also be made regarding television and other media, and spending some time doing pleasurable activities is not usually seen as morally problematic if it does not pass the threshold to addiction, which can undermine autonomy's competency condition (Susser, Roessler, and Nissenbaum 2019, 8).

Limiting Choice by Design

I argue that autonomy is also harmed in those that are not addicted. The influence of hypernudges on the user is especially strong on TikTok, where *choice architecture* is distilled down from giving the user pre-selected and ordered choices to just showing one video at a time. While nudges are usually used in *digital decision guidance processes*, such as by offering suggestions that the user can select, TikTok threads the needle to an *automated decision-making process*. In TikTok's primary mode of consumption, the "for you" - page, the user's agency is put second: Instead of confirming the choice of an algorithm to engage with the content, they can instead skip the choice the algorithm already auto-plays in their face. When the user consumes content, the algorithm dictates what they see, with the user only steering after the fact. This is undoubtedly an extreme case of choice architecture, as the choice to start watching the video is already made for the user in advance. While engaging with a piece of content, the only options are to keep consuming, skip to the following video, or stopping to consume it. The autonomy of choice for selecting individual pieces of content in this mode is severely diminished.

The limited autonomy of choice in consumption is reminiscent of television. When watching TV, users can consume what is playing or keep zapping through channels until they find something personally appealing. Of course, they can also inform themselves of the running programs and actively choose what to watch. This is also possible in TikTok: Users can search for individual content creators and consume their videos or only watch the content of users they follow. Nevertheless, in both examples, especially for TikTok, many people lazily use the media to 'kill time' by zapping the abovementioned "for you" – page. While the options of uploading content oneself and consuming content created by other individuals without much budget or a licence are unique to the new form of media, as is the ability to comment and rate content, the principal difference in the mode of consumption is how personalised the content is. The comparison with television shows that TikTok's restricted autonomy to choose content is nothing new and not itself at the core of the problem. However, it can augment the effect of diminishing autonomy.

TikTok does not only give even less attention to users' autonomy of choice than other social media platforms in making the algorithm central to the platform’s design but also continuously analyses users over a magnitude of indicators. So to say, each moment staring into TikTok, TikTok stares into you. By subconsciously giving away secrets about one's identity and psychology, one becomes predictable and gives up power over oneself. In poker, this would set one up to be read and lose the game. Using TikTok, this information is fed back directly into one's source of information and entertainment in an endlessly feeding stream of short videos that can have subliminal effects.

Hypernudges and Manipulation

Past studies have already shown the potential of (hyper)nudges on Facebook to influence people to vote (Oremus 2012) and to manipulate their emotions through interference with the algorithm (Hill 2014). As in the second case, manipulation on TikTok could happen entirely subliminally but automated and to an even finer degree than in the studies conducted around ten years ago. If TikTok knows one's psychological vulnerabilities, it can subtly exploit them via hypernudges to steer a user's emotions, beliefs, and to a degree, actions.

Whether or not nudges can be considered manipulation, hypernudges certainly are. Hypernudges bring out the most manipulative aspects of nudges in a highly personalised, sophisticated manner and seek to bypass the user's rational decision-making. C*overt influences on a person's decision-making by targeting and exploiting their decision-making vulnerabilities* constitute precisely the definition of manipulation (Susser, Roessler, and Nissenbaum 2019, 6). Drawing back on the definition of autonomy, this type of manipulation causes people to act based on reasons they did not embrace themselves, towards ends they did not choose. Through manipulative mechanisms, social media companies such as TikTok greatly diminish their user's autonomy.

Such manipulation can be most apparent in the use of advertising to sell users products they 'did not know they wanted'. Users can be repeatedly exposed to ads and sponsored content that draws on their personal traits or insecurities. However, influences are not necessarily constricted to advertising, as they could also be used to keep us hooked to the platform by guiding us to *engaging* content, guide us to more advertiser-friendly content, manipulate political opinions and elections or potentially even target groups or individual people for whatever reason. With its opaque but potent power, if TikTok wanted, it could probably incite anti-government protests in many countries.

Political figures increasingly acknowledge these possibilities, for example, in the USA and India, where TikTok might be banned soon or is already banned. Reasons are security concerns, such as espionage and fears of influencing their citizens through harmful content (Morrison 2023) in the US, and "sovereignty and integrity of India", as well as "security of state and public order" in India (*BBC* 2020). Of course, these countries are more concerned about TikTok than other social media, as it is owned by a Chinese company, with the Chinese government being a geopolitical rival.

Geopolitical interests do not matter as much for the user's implications. Ultimately, it is a play with fire to consume platforms utilising algorithmic paternalism of this form, as we grant them control to manipulate us at will. At the same time, the platform's interests remain hidden. After all, there is a massive power and information asymmetry (Zuboff 2015, 85) here – while big data companies continue to accumulate and analyse massive amounts of information about users and society, users do not get to access information about the companies, algorithms, and accumulated data.

How TikTok Impacts Well-being

Different theories of well-being correspond differently to diminished autonomy in the case of TikTok.

Firstly, hedonistic models draw no necessary negative implications from it. All positive or negative impacts of diminished autonomy on well-being in this conception of well-being come from secondary effects concerning the pleasure-and-pain wager. TikTok can provide pleasure in the short term, although it might hinder long-term well-being as the time spent swiping through it could be used to increase well-being long-term. It is also imaginable that people feel less comfortable living with life decisions that they did not choose autonomously. On the other hand, some people might get manipulated into living a happier life, such as when TikTok subconsciously influences their sense of fashion to appeal better to others and, in turn, indirectly gain friends and relationships. In the end, placing your well-being to a degree at the mercy of a big data company is undoubtedly a gamble, but one to which hedonistic well-being does not necessarily object.

Secondly, desire theories are certainly more critical of compromised autonomy. After all, to live a good life according to desire theories, one should define and fulfil one's own desires. Through manipulation, TikTok can implant desires into users that are not authentically their own, driving them apart from their authentic desires and leaving them unfulfilled. Of course, TikTok can also aid in fulfilling one's authentic desires, as when using it as a source of information on hobbies and interests. However, users should be aware that TikTok is a tool that uses them right back with a significant power asymmetry between the social network and user, and of the potential gap between their authentic desires and those that might be implanted by TikTok that might negatively impact their well-being.

Lastly, in objective list theories, the influence of TikTok might help with some goals, such as social acceptance. Nevertheless, its stranglehold on autonomy, which is often seen as an intrinsic and necessary value for well-being, disqualifies any potential positive impacts according to this type of theory. Objective list theories hold that people should be autonomous as a condition for living a good life and are not convinced by possible positive side effects. Following objective list theories, people should either be perfectly informed about the technology and intentions of TikTok not to get manipulated or, as this is not possible owing to the power asymmetry between big data companies and users, not use the platform at all.

Conclusion

This paper discusses how using social media platforms can impact an individual's well-being through the example of TikTok.

It introduces the concept of algorithmic paternalism, a term for the use of algorithms to guide and influence user behaviour in a way that is claimed, perceived, or aimed to be in their best interest. Algorithmic paternalism is not, by definition, a negatively connotated term but is often perverted, such as with the use of manipulative hypernudges by social media companies.

To show how social media companies can impact well-being, the paper lays out in detail how the use of TikTok can impact users and diminish an individual's autonomy, both by a design that forbids autonomy of choice to a higher degree than comparable platforms and through algorithmic paternalism: Utilising the personalised choice architecture termed hypernudge in combination with detailed gathered and analysed user data, it is possible to exploit the decision-making vulnerabilities of individuals.

Diminished autonomy has consequences for well-being, but the specific effects depend on the theory of well-being utilised: While hedonistic well-being only makes out potential indirect connections from diminished autonomy to reduced well-being, desire theories are concerned that users could get manipulated into following desires that are not authentically their own and are left unfulfilled, and most objective list theories see autonomy as a necessary component of a good life.

I propose the following suggestions for further interdisciplinary research: 1. devising rules for acceptable and non-acceptable uses of algorithmic paternalism, 2. empirically measuring the impacts of hypernudges on society, 3. critically assessing the choice environment designs of different social media platforms and 4. devising a platform design that is structurally not manipulative and overcomes the power asymmetry between platform and user.

Bibliography

BBC. 2020. “India Bans TikTok, WeChat and Dozens More Chinese Apps,” June 29, 2020, sec. Technology. https://www.bbc.com/news/technology-53225720.

Bentham, Jeremy. 1970. *An Introduction to the Principles of Morals and Legislation.* Greenwood Press.

Bovens, Luc. 2008. “The Ethics of Nudge.” https://philarchive.org/rec/BOVTEO-8.

Brown, Peter. 2020. “Exploring Information Literacy in the Context of the Attention Economy: A Heuristic-Based Framework for Attitude Certainty.” Lubbock: Texas Tech University. https://ttu-ir.tdl.org/bitstream/handle/2346/86642/BROWN-DISSERTATION-2020.pdf?sequence=1.

Carville, Olivia. 2023. “TikTok’s Algorithm Keeps Pushing Suicide to Vulnerable Kids.” *Bloomberg*, April 20, 2023. https://www.bloomberg.com/news/features/2023-04-20/tiktok-effects-on-mental-health-in-focus-after-teen-suicide.

Crisp, Roger. 1997. *Mill on Utilitarianism*. Second Edition. Routledge Philosophy Guidebooks. London: Routledge.

———. 2021. “Well-Being.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2021. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2021/entries/well-being/.

Davenport, Thomas H., and John C. Beck. 2001. *The Attention Economy: Understanding the New Currency of Business*. Boston: Harvard Business School Press.

Dworkin, Gerald. 1971. “Paternalism.” In *Morality and the Law*, edited by Richard A. Wasserstrom. Belmont, Calif., Wadsworth Pub. Co. https://archive.org/details/moralitylaw00wass/page/106/mode/2up.

———. 2020. “Paternalism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2020. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2020/entries/paternalism/.

Gao, Michelle. 2020. “Facebook Makes More Money per User than Rivals, but It’s Running out of Growth Options.” CNBC. November 3, 2020. https://www.cnbc.com/2020/11/03/facebooks-average-revenue-per-user-leads-social-media-companies.html.

Hill, Kashmir. 2014. “Facebook Manipulated 689,003 Users’ Emotions For Science.” *Forbes*, June 28, 2014, sec. Tech. https://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/.

Husak, Douglas N. 1981. “Paternalism and Autonomy.” *Philosophy and Public Affairs* 10 (1): 27–46.

Isaac, Mike. 2021. “Facebook Nearly Doubles Its Profit and Revenue Rises 48 Percent, as Tech Booms.” *The New York Times*, April 28, 2021, sec. Business. https://www.nytimes.com/2021/04/28/business/facebook-earnings-profit.html.

MacGowan, Amalie. 2020. “The TikTok Algorithm Knew My Sexuality Better Than I Did.” *Repeller* (blog). July 8, 2020. https://repeller.com/tiktok-algorithm-bisexual/.

Mackenzie, Catriona, and Natalie Stoljar. 2000. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. Oxford University Press.

McCradden, Melissa D., and Roxanne E. Kirsch. 2023. “Patient Wisdom Should Be Incorporated into Health AI to Avoid Algorithmic Paternalism.” *Nature Medicine*, February, 1–2. https://doi.org/10.1038/s41591-023-02224-8.

Merrill, Jeremy B. 2016. “Liberal, Moderate or Conservative? See How Facebook Labels You.” *The New York Times*, August 23, 2016, sec. U.S. https://www.nytimes.com/2016/08/24/us/politics/facebook-ads-politics.html.

Merrill, Jeremy B., and Will Oremus. 2021. “Five Points for Anger, One for a ‘like’: How Facebook’s Formula Fostered Rage and Misinformation.” *Washington Post*, October 26, 2021, sec. Technology. https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/.

Morrison, Sara. 2023. “TikTok’s Master Plan to Win over Washington.” Vox. January 17, 2023. https://www.vox.com/recode/2023/1/17/23552716/tiktok-ban-cfius-bytedance.

Oremus, Will. 2012. “One Facebook Banner Ad Caused 60,000 More People To Vote in the 2010 Elections.” *Slate*, September 13, 2012. https://slate.com/technology/2012/09/facebook-voting-study-online-friends-influence-voter-turnout-in-elections.html.

Raz, Joseph. 1986. *The Morality of Freedom*. Clarendon Press.

Smith, Ben. 2021. “How TikTok Reads Your Mind.” *The New York Times*, December 6, 2021, sec. Business. https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html.

Susser, Daniel, Beate Roessler, and Helen Nissenbaum. 2019. “Technology, Autonomy, and Manipulation.” *Internet Policy Review* 8 (2). https://policyreview.info/articles/analysis/technology-autonomy-and-manipulation.

Thaler, Richard H, and Cass R Sunstein. 2003. “Libertarian Paternalism.” *American Economic Review* 93 (2): 175–79. https://doi.org/10.1257/000282803321947001.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven, CT, US: Yale University Press.

“TikTok App Store Page.” 2023. App Store. February 28, 2023. https://apps.apple.com/us/app/tiktok/id835599320.

Weale, Sally. 2023. “‘We See Misogyny Every Day’: How Andrew Tate’s Twisted Ideology Infiltrated British Schools.” *The Guardian*, February 2, 2023, sec. Society. https://www.theguardian.com/society/2023/feb/02/andrew-tate-twisted-ideology-infiltrated-british-schools.

Yeung, Karen. 2017. “‘Hypernudge’: Big Data as a Mode of Regulation by Design.” *Information, Communication & Society* 20 (1): 118–36. https://doi.org/10.1080/1369118X.2016.1186713.

Zuboff, Shoshana. 2015. “Big Other: Surveillance Capitalism and the Prospects of an Information Civilization.” *Journal of Information Technology* 30 (1): 75–89. https://doi.org/10.1057/jit.2015.5.