Personalizing the Lifeworld Away: Nudging and Algorithmic

Personalization in Habermas' Theory of Communicative Action

Kjell Dewald / 453639kd

Supervisor: M. Wehrle

Advisor: J.A. van Ruler

Main study: Psychology

Date of completion: 15/06/2021

10987 words

Table of Contents

Introduction

Shocking to some and affirming the beliefs of others, the Netflix documentary "The Social Dilemma" (2020) describes how users of social platforms are directed in their behavior through the design of a personalized digital environment. Starring important voices of the Big Data debate such as Shoshana Zuboff, the documentary explains in detail how machine learning and algorithms help detect the content that engages users. In the course of the documentary, Tristan Harris, a former employee of Google who focused on the ethical use of technology, makes an interesting point on the widely shared fear of technology ending up overtaking human ability:

> ….we're all looking out for the moment when technology would overwhelm human strengths and intelligence — when is it going to cross the singularity, replace our jobs, be smarter than humans? This is the thing everyone's talking about.
>
> But there's this much earlier moment when technology actually overwhelms human weaknesses. ("Humane: A New Agenda for Tech," Tristan Harris)

The documentary goes on to illustrate how social media utilizes basic principles of psychology, such as positive reinforcement, to make its users log in more often and for longer amounts of time. In short, the documentary demonstrates that technology, in some ways, has already succeeded in overwhelming human weaknesses. Interestingly, one psychological concept, the theory of nudging, seems to miss in the discussion.

Nudging, which was first introduced by economists Sunstein and Thaler in 2008[1] describes the process of designing the environment to guide people towards a specific behavior. To illustrate the concept, the authors gave the example of a cafeteria being designed. In this process, the designer of the cafeteria has to make decisions that will affect the customers of the cafeteria, such as what food to place at eye level. Such designed environments are also referred to as choice architectures while nudging is defined by its originators as

> (…) any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. (…) Putting the fruit at eye level counts as a nudge. Banning junk food does not. ("Introduction", *Nudge: Improving Decisions About Health, Wealth, and Happiness*, 6)

The example of putting fruit at eye level exemplifies the potential that nudging has in guiding the recipients to behaviors beneficial to their well-being. However, nudges may also be used

---

[1] Richard H. Thaler, and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth. And Happiness* (New York: Penguin Press, 2008).

to further the goals of a nudge's architect, for instance, to make customers buy the more expensive products, regardless of their benefits to the consumer.[2]

When introducing the internet and the possibilities it gives to the choice architects, the debate gets more complex. I argue that nudging first and foremost is an enabler, a gateway so to speak, that opens the door to the collection of user data. By consulting data, online environments and nudges can be personalized by using algorithmic logic. Considering the importance online environments carry in modern societies, it is necessary to investigate the possible influences that this data-consulted personalization has on an individual and societal level.

Considering Habermas' theory of communicative action, this paper argues that algorithmic personalization is problematic as it evades public scrutiny and imposes its algorithmic logic on the general public. To illustrate the argumentation behind this, the paper is first going to introduce the Theory of Communicative Action. Subsequently, nudging and its role in algorithmic personalization is framed in Habermas' account. Here, the normative issues are made apparent.

Since algorithmic logic is closely tied to user behavior, I consider Latour's concept of technical mediation as a meaningful addition to the discussion that could help point towards a possible solution. Latour's concept of black-boxing illustrates that users engage with algorithms mostly non intentionally, leading to a passive personalization procedure of online experiences. A solution could be to counter this passive personalization procedure with an active customization process that enables users to regain agency over the personalization of online environments. This more procedural outlook on the issue is helpful to identify possible ways out of the normative issue.

Chapter 1 - A Theory of Communicative Action

Forming a major part of his work, The Theory of Communicative Action, comprises Habermas' observations and criticisms regarding the modernized society. As opposed to earlier members of the Frankfurt School, such as Adorno and Horkheimer, Habermas is much more positive about the project of Enlightenment. An especially criticized characteristic of the modernized society is the prevalent use of rationality in a means to an end process (called instrumental rationality). In Horkheimer and Adorno's perspective, this type of rationality leads to several such as the homogenizing of society, and a view on nature that is only

---

[2] Thaler and Sunstein, *Nudge,* 242.

interested in its usefulness.[3] While Habermas does not content that there are problems in the project of enlightenment, he argues that next to instrumental rationality, there is another form of rationality implicit in any communication.[4]

Habermas's theory is based on communication and argues that in any form of communication there are validity claims made with relation to the outer world, the social world, and the inner world of a communicator.[5] To Habermas, the enlightenment enabled the use of communicative rationality within society, leading to a discourse on social norms which are no longer presumed to be valid but can be criticized.

This communicative focus of Habermas is especially important in later argumentation as it is suggested that nudges communicate norms and values implicitly. As such they too can be framed with Habermas's theory. Nudging is a good example to show how the design of environments can carry communicative meaning. This communicative claim is also argued to be present in algorithmically personalized environments.

While Habermas has been connected to developments of the internet, such as social media, through the public sphere and the interaction of users with other users[6] the interaction between user and provider seems to be overlooked. As is argued in subsequent sections, Habermas contributes important ideas to this debate that point out a lack of communicative rationality towards the practices of online personalization.

Lastly, Habermas not only conceives of a communicative theory that underlines important structures of a functioning modern society but also points out possible consequences when said structures are disturbed and undermined. This is helps discuss the consequences of algorithmic personalization.

To take a closer look at this criticism, a clear idea of Habermas's framework is necessary. In the following chapter, I aim to explicate central notions relevant to further argumentation. First, the concept of Lifeworld and Communicative Action is going to be summarized before explicating the System as well as Communicative Relief.

1.1 Lifeworld and communicative action

---

[3] Jospeh Berendezen "Max Horkheimer," in *Stanford Encyclopedia of Philosophy* (June 2009, substantiative revision June 2017): Accessed June 9, 2021 on https://plato.stanford.edu/entries/horkheimer/.
[4] Jürgen Habermas, *Theory of Communicative Action Vol I* (henceforth referred to as TCA I)*, trans. Thomas McCarthy (Cambridge, UK: Polity Press, 1984), 10.
[5] Jürgen Habermas, *TCA I*, 99.
[6] Axel Bruns, and Tim Highfield, "Is Habermas on Twitter? Social Media and the Public Sphere," in *The Routledge Companion to Social Media and Politics,* ed. Axel Bruns, Gunn Enli, Eli Skogerbø, Anders Olof Larsson, Christian Christensen (New York: Routledge, 2015) 56 – 70.

To Habermas, all communication is teleological in the sense that communication is used by actors to pursue an aim.[7] In communicating, the actor can enter relations to three distinct worlds: The social world, the object world, and the subjective world.[8] Communication now differs in the ways that aims can be achieved as well as the number of relations a communication manifest. A strategic communicator for instance relates primarily to the object world and tries to achieve his aim with instrumental rationality.[9] In contrast to strategic, normative, and dramaturgical action, communicative action relates to all three worlds at once, intending to reach an understanding.[10]

By relating to all three worlds at once, communicative action makes three validity claims:

1.   On the truth of the statement (accurate representation of the object world).

2.   On the normative context (that the statement is appropriate).

3.   On the sincerity of the speaker (that the statement is in line with the intentions of the speaker).

These validity claims can either be criticized or accepted. If they are accepted, an understanding is reached that affirms the three validity claims.[11] If, however, one or more of the claims are criticized, a discourse will develop.[12] Here, the validity claims raised will be discussed further based on justifiable reasons for making a certain claim. This process of discourse is explained in detail by Habermas's theory of discourse.[13] For this thesis, the process of discourse is not further relevant.

When considering this process of evaluating validity claims, the question arises for the point of reference in the discourse. After all, the speaker and listener need to be able to investigate the claims according to common background knowledge. Here, Habermas introduces the concept of the Lifeworld, which he defines as this common backdrop of implied values and knowledge.[14] In general, this backdrop is not questioned, but in cases of truth claims it is subject to scrutiny.

The Lifeworld is itself comprised of three structural parts namely, culture, society, and personality[15]. These structural parts are integral to assure the possibility of communicative

---

[7] Habermas, *TCA I*, 101.

[8] Habermas, *TCA I*, 99.

[9] Habermas, *TCA I*, 87.

[10] Habermas, *TCA I*, 94.

[11] Habermas, *TCA I*, 114-115.

[12] James Bohman, and William Rehg, "Jürgen Habermas," in The Stanford Encyclopedia of Philosophy (May 2007, minor correction Fall 2017): Accessed June 9, 2021 on https://plato.stanford.edu/entries/habermas/.

[13] Ibid.

[14] Habermas, *TCA I*, 100.

[15] Jürgen Habermas, *Theory of Communicative Action Vol II* (henceforth referred to as TCA II), trans. Thomas McCarthy (Cambridge, UK: Polity Press, 1987), Chapter VI, Section 1C, Kindle Version, position 3222.

action. Society, for instance, also deals with norms that stabilize communicative action; While no law that would forbid me from going back on my promise to meet someone at a specific time without good reason, it would reduce my future (social) credibility because of established norms. Habermas defines the three different structural components as follows:

> I use the term culture for the stock of knowledge from which participants in communication supply themselves with interpretations as they come to an understanding about something in the world. I use the term society for the legitimate orders through which participants regulate their memberships in social groups and thereby secure solidarity. By personality I understand the competences that make a subject capable of speaking and acting, that put him in a position to take part in processes of reaching understanding and thereby to assert his own identity. (Chapter 6, Section 1C, *Theory of Communicative Action Vol. II*)

Communicative action can be observed in a multitude of aspects of modern life, such as the basic interactions that we have with family and friends (i.e. deciding on where to eat), but also in political discussions held in the private and public sphere. These interactions depend on a shared understanding of the world we live in. At the same time, this world depends on the interactions to continue its existence as a backdrop of shared understanding.[16]

The Lifeworld needs to be reproduced so that it persists over time. This process is achieved through communicative action. More specifically, by reaching an understanding with others, the validity claims made are accepted and thereby affirm the Lifeworld.[17] Employing discourse, the lifeworld can change but it needs to be changed based on communicative action through discourse since at the end of that discourse there is understanding which affirms the Lifeworld. The three structural components of the Lifeworld each have a corresponding reproductive process: Culture is reproduced using cultural reproduction, society by social integration, and personality is reproduced through socialization.[18] Habermas identifies a threat to society in the disturbances of these reproductive processes. This could happen when communicative action is no longer performed sufficiently, for instance, because other forms of action are employed instead (such as strategic action).

To illustrate the main source for this threat, the next section will focus on the System and its use of steering mechanisms.

1.2 The System

---

[16] Habermas, *TCA II*, Chapter VI, Section 1C, Kindle Version, position 3267.
[17] Habermas, *TCA II*, Chapter VI, Section 1A, Kindle Version, position 2868.
[18] Habermas, *TCA II*, Chapter VI, Section 1C, Kindle Version, position 3222.

While there is symbolic reproduction to assure the existence of culture, society, and personalization over time, material reproduction is necessary to ensure that there is a physical world in which social interaction can take place.[19] This material reproduction depends on purposive action.[20] Thus, it can be achieved by communicative action or strategic action. In early tribal societies, material reproduction was possible utilizing communicative action.[21] However, with the increasing complexity of society, communicative action was no longer an effective way of coordinating the process of material reproduction. In its stead, modern societies developed institutions that are rooted in the lifeworld which coordinate action systematically.[22] These institutions form what Habermas terms the system. In modern society, a state is organized which is guaranteed by legal means instead of kinship or divine interventions. Its steering media, that is, its way of coordinating action, is power backed by means of enforcement.[23] In a state-organized society, a second steering media is introduced with money, through the developments of markets. This steering medium is backed by gold.[24]

The use of these steering mechanisms is that they are more efficient in deciding as they can circumvent the validity claims of communicative action. Alongside the steering mechanisms, there exist other forms of communication relief, that succeed in easing the process of communicative action. Habermas specifically mentions influence and value commitment as two further mechanisms. He terms this form of communication "generalized communication". [25]

1.3 Generalized communication

Influence is of importance whenever the objective validity claims are too complex to argue about[26]. A good example may be the far-reaching influence of medical experts exuded in their recommendations during the COVID-19 pandemic. Value commitment on the other hand is performed by opinion leaders.[27] Here the normative validity claims are targeted, and listeners are inspired to an increased commitment to values. Again, a current example may be the

---

[19] Habermas, *TCA II*, Chapter VI, Section 1D, Kindle version, position 3236.
[20] Ibid.
[21] Habermas, *TCA II*, Chapter VI, Section 2A.
[22] Habermas, *TCA II*, Chapter VI, Section 2C.
[23] Habermas, *TCA II*, Chapter VII, Section 2C, Kindle version, position 5998.
[24] Habermas, *TCA II*, Chapter VII, Section 2C, Kindle version, position 6061.
[25] Habermas, *TCA II*, Chapter VII, Section 2C, Kindle version, position 6168.
[26] Thomas Hove, "Understanding and Efficiency: Habermas's Concept of Communication Relief," *Communication Theory* 18, no. 2 (18. April 2008): 249. https://doi.org/10.1111/j.1468-2885.2008.00321.x
[27] Hove, "Understanding and Efficiency," 248.

number of celebrities (that can be considered role models to some), who tried to inspire people to follow appropriate norms during the pandemic. As summarized by Thomas Hove in his discussion on the concept of communication relief:

> On one hand, when people interact through steering media, they orient themselves toward one another's relative power to issue rewards and punishments. On the other hand, when people interact through generalized communications, they orient themselves toward one another's inferred ability to provide legitimate reasons for believing or doing something. ("Understanding and Efficiency", 248)

In this regard, influence and value commitment differ to power and money, as they are still interested in performing communicative action while simplifying the acceptance of certain validity claims. Importantly, they do not depend on an empirical notion of motivation (i.e. punishment and reward).

The detailed account of these methods of communication relief (steering media and generalized communication) are important as they help frame classical nudging as a form of communication in the later argumentation. Importantly, it is argued in later sections that classical nudging resembles a form of generalized communication, while the consultation of data in algorithmic personalization resembles the characteristics of steering mechanisms.

Lastly, Habermas is not oblivious to the fact that generalized communication may be used with a strategic focus. Doing so would also mean that generalized communication becomes an instrument of money and power; non-manipulative goods are misappropriated for manipulative use.

1.4 The colonization thesis

This dichotomous account of either communicating by influence and value commitment to promoting communicative action on the one hand, and using general communication in a strategic, manipulative way, on the other hand, is something that is also reflected in the general conception of the "System" (i.e., state government and market): Habermas argues that a danger of the System is, that communicative action is more and more suppressed by steering mechanisms such as power and money.[28] If communicative action is suppressed in this way, the reproduction of the lifeworld is at stake. This is widely known as the colonization thesis.[29] However, according to Hove, it would be a misreading if one were to assume that the only possible effect is unidirectional from the System to the lifeworld.[30] On the contrary, the

---

[28] Habermas, *TCA II*, Chapter VI, Section 2F. Kindle version, position 4454.
[29] Bohman, and Rehg, "Jürgen Habermas".
[30] Hove, "Understanding and Efficiency," 244.

lifeworld may also influence the system, leading to an expansion of the lifeworld rather than a limitation of it as per the colonization thesis.

> Both are conceivable: the institutions that anchor steering mechanisms such as power and money in the lifeworld could serve as a channel either for the influence of the lifeworld on formally organized domains of action or, conversely, for the influence of the system on communicatively structured contexts of action. In the one case, they function as an institutional framework that subjects system maintenance to the normative restrictions of the lifeworld, in the other, as a base that subordinates the life-world to the systemic constraints of material reproduction and thereby "mediatizes" it. (Chapter VI, Section 2F, *Theory of Communicative Action Vol. II*)

Chapter 2 - Nudging, Algorithms, and Personalization

Chapter 1 introduced Habermas's theory of communicative action. The aim of chapter 2 is to apply this framework to the practice of nudging and explicate how the characteristics of nudging and algorithmic personalization differ. More specifically, I hope to illustrate that classical nudging can be conceived of as a delinguistifed form of generalized communication. In contrast, the consultation of data in algorithmic personalization rather resembles a steering mechanism similar to money and power. Classical nudging is important to algorithmic personalization since it opens the door to increased collection of user data. To effectively argue my case, I will take a closer look at nudging and its underlying mechanisms. Subsequently, the introduction of the default nudge explains how classical nudging opens the door to algorithmic personalization.

2.1 Nudging

Explicated by Sunstein and Thaler in 2009, nudging rests on the psychological theory of Kahneman and Tversky. The key idea is that human decision-making is guided by two systems, one of which makes impulsive split-second decisions, while the other is a more deliberate system that takes more effort but is also more accurate.[31] Kahneman and Tversky, build on this conceptualization by arguing that biases and heuristics in decision making are rooted in the impulsive system.[32] Often, these thought processes may lead to decisions, not in the best interest of the decision-maker. In other words, they are weaknesses of human decision-making. A quick example makes this point clearer: The availability heuristic states

---

[31] Daniel Kahnemann, "*Schnelles Denken, Langsames Denken*," trans. Thorsten Schmidt (München: Pantheon Verlag, 2011), 33.
[32] Kahnemann, *Schnelles Denken, Langsames Denken*, 136.

that our decisions are influenced by the immediacy of examples.[33] Thus, a person thinking of taking the plane for a vacation may change his original plan when he is confronted with a recent example of a plane crash. Instead, he opts to take the car, even though it is considered to be a riskier means of transportation.

Sunstein and Thaler, recognize that the weakness present in decision making often does not work towards one's advantage. However, if we know of the weakness, why not design an environment, that allows people to decide impulsively without it resulting in negative consequences for the decider? This is the original scope that the theory of nudging aimed to address. Generally, nudging is seen as a chance to improve the decisions one makes (as is apparent from the title of the book by Sunstein and Thaler).

Further, an important consideration in nudging is that it is inevitable.[34] No matter how you order the food in the cafeteria, or what you show on television, it is inevitably going to have an influence. The question, therefore, is not whether to nudge but rather in what direction to nudge. This is vital to the leading principle behind nudging: liberal paternalism.[35] An oxymoron at first sight, the core idea is that no matter the nudge, everyone should still have all choice possibilities (Liberalism). Nonetheless, and this is the paternalistic account, one of the possibilities is inevitably going to be appeal more than the others. Take the example of the cafeteria; Paternalism as such would likely just ban junk food from the cafeteria, thereby subjecting everyone to a normative decision. Liberal paternalism on the other hand still keeps the junk food but it is going to ease eating fruit over junk food. The nudge, therefore, is a piece of advice rather than an order.

Here, I see a parallel to Habermas and his concept of generalized communication. More specifically, I think that value commitment, much like nudging can be considered a piece of advice first and foremost. If an opinion leader is going to appeal to the population, he is going to have the desired influence on some, but not all of those listening. Much in the same way, a classical nudge is going to work for some of the population, and for others, it may not.

Further parallels can be drawn between the two concepts. Nudges, as well as appeals to value commitment, are normative. The celebrity urging people to limit their contacts during the pandemic is doing so out of the normative idea of solidarity. Similarly, the speed bump is

---

[33] Kahnemann, *Schnelles Denken, Langsames Denken*, 164 – 184.
[34] Sunstein and Thaler, *Nudge*, 10.
[35] Cass R. Sunstein., and Thaler, Richard T. "Libertarian Paternalism," *The American Economic Review* 93, no.2, (May, 2003): 175-79. Accessed June 10, 2021 http://www.jstor.org/stable/3132220.

urging drivers to drive slower through densely populated areas, with regard to the norm of solidarity towards others. Lastly, and importantly for future discussion much like generalized communication, nudging may well be misappropriated for manipulative use. Examples of such misappropriated nudges can be encountered in everyday life – think for instance of sweets being presented next to the cash register. Working the opposite way as the example with fruit and junk food, this nudge even has a designated word in German; "Quengelware" refers to the notion that children will beg their parents to buy the sweets when waiting in line.[36] Nudging is misappropriated here, as children should be encouraged to eat healthily. Unfortunately, modern society is not yet at the point where children desperately beg their parents to buy some more broccoli while standing in the line of the supermarket cash register. Nudges rooted in the private interest of the market will be employed strategically and are therefore misappropriated.

The important point of my argumentation is that nudging possesses communicative characteristics that appear similar to a form of generalized communication. Due to the normative focus of value commitments, this form of generalized communication most closely resembles the form of communication present in nudging. Nonetheless, the two do not equate. Most importantly, value commitment depends on the explicit linguist expression of advice. This explicitness is not present in nudging, since the concept depends on being implicit in choice architectures. When thinking of nudging in a form of speech, it might therefore best equate a suggestive question rather than an explicit appeal to values. Still, due to its normative nature, nudging possesses an important characteristic of generalized communication since it does not depend on empirically motivating resources (i.e., punishment and reward) but rather on rationally motivated resources such as reasons, evidence, and justifications. Choice architects can be criticized for their designs, and when they are, they can point towards rationally motivated resources, such as the explanation that they wish to encourage healthy eating over junk food. The main argument to be made in this section is that nudging is a form of communication occurring in the design of choice architectures.

As such I consider classical nudging a delinguistified form of generalized communication. To make the case that in contrast to classical nudging, algorithmic personalization can be characterized as a steering mechanism, the paper is first going to take a look at how classical nudging opens the door to algorithmic personalization.
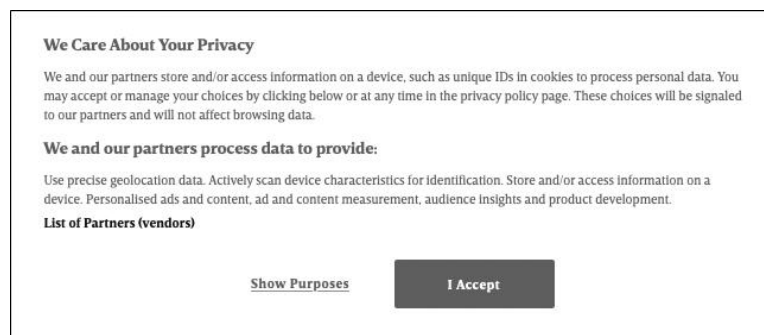
---

[36] Duden entry „Quengelware," accessed June 10, 2021 on https://www.duden.de/rechtschreibung/Quengelware

2.2 The default nudge

Turning to nudges in an online context, windows popping up asking the user to accept the cookie preferences of the content provider has become a common occurrence. Giving consent and asking for it have been important issues in policies on consumer protection.[37] Nonetheless, it seems a large majority of customers do not read the terms and services.[38] The recent scandal of a Facebook study on emotional contagion that only required its participants (close to 700,000) to have signed the terms and services of Facebook illustrates the importance of informed consent.[39] While the EU General Data Protection Regulation privacy regulations implemented in 2018 made content providers request active consent from their visitors to web tracking via cookies, it is still up to the provider how to ask for this consent.[40] As has become apparent in the previous section, interactive design is susceptible to nudging. In this case, a so-called default nudge is employed to make it more likely that visitors accept cookie preferences rather than reject them.

Default nudges build on the assumption that users are more likely to accept a default rather than to change it, especially if the benefits of changing the default are considered rather low.[41] [42] Consider the following example of a window asking the user to accept the cookie preferences:[43]

We Care About Your Privacy

We and our partners store and/or access information on a device, such as unique IDs in cookies to process personal data. You may accept or manage your choices by clicking below or at any time in the privacy policy page. These choices will be signaled to our partners and will not affect browsing data.

We and our partners process data to provide:

Use precise geolocation data. Actively scan device characteristics for identification. Store and/or access information on a device. Personalised ads and content, ad and content measurement, audience insights and product development.

List of Partners (vendors)

Show Purposes   I Accept

---

[37] Michiel Rhoen, "Beyond consent: Improving data protection through consumer protection law," *Internet Policy Review* 5, no. 1 (March 2016): 2. https://doi.org/10.14763/2016.1.404

[38] Jonathan A. Obar., and Anne Oeldorf-Hirsch, "The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services," *Information, Communication & Society* 23, no. 1 (June 2018): 140. https://doi.org/10.1080/1369118X.2018.1486870

[39] Adam D. I. Kramer, Jaime E. Guillory, and Jeffrey T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Sciences* 111, no. 24 (2014): 8788. https://doi.org/10.1073/pnas.1320040111

[40] Deloitte, "Cookie Benchmark Study," (2020): 6. Accessed on June 10, 2021 on https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/risk/deloitte-nl-risk-cookie-benchmark-study.pdf

[41] Craig M. Smith, Daniel G. Goldstein, and Eric J. Johnson, "Choice Without Awareness: Ethical and Policy Implications of Defaults," *American Marketing Association* 32, no. 2 (Fall 2013): 163. https://doi.org/10.1509%2Fjppm.10.114

[42] Sunstein and Thaler, "Nudge," 85.

[43] Examples are screenshots of Wired.com's cookie notification.

While users cannot ignore this banner, they are invited to click "I Accept" (note how "I Accept" is designed more invitingly in comparison to the "Show Purposes" link). Only by clicking on "Show Purposes", does the user get the possibility to personalize what information the website is allowed to collect. Here, the subsequent window shows up:

After the second click on "confirm my choices", the user successfully opted out of web tracking (for this particular site). Thus, all the characteristics of a default nudge are there. The default is to accept the cookie tracking and most user will do exactly that since the benefits of changing the settings are not judged to be high enough.

Naturally, this nudge is not only used by individual websites but also by the largest search engine, Google. When signing up for a Google account, it is necessary to first confirm the personalization settings and cookies. Only when finishing signing up for google is it possible to alter the settings.[44] Again, this is a classic form of the default nudge since after a pre-selection has been made, only those that place a high value on their privacy settings will go through the trouble of changing them.

Default nudges, therefore, play an important role to allow content providers to collect and use personal data. As such, default nudges as applied in the given example are an example of the misappropriation of nudges since their use is based on economic self-interest.

---

[44] Tested by the author in May 2021.

　　　　To summarize, this section explicated how a classical default nudge is employed to obtain a maximized amount of data on users. In the following sections, I aim to argue that the obtained data enables the algorithmic personalization of online environments. In contrast to nudging and its communicative claim, algorithmic personalization resembles a steering mechanism rather than generalized communication. Since steering mechanisms such as money and power are backed by gold and means of enforcement, I would like to propose that algorithmic personalization is backed by data. To do so, I am first going to look at how data structures are utilized within algorithmic logic.

2.3 Algorithmic logic

In its broadest sense, an algorithm is just a sequence of steps that will solve a problem, and this still rings true even when problems are getting more and more complicated.[45] An algorithm may be the sequence of steps taken to solve a Rubik's cube, but it can also be the solution to complicated tasks such as increasing user engagement. It is difficult to explain the exact way online algorithms work as they often are developed by private companies which proclaim them a trade secret.[46] Nonetheless, there are certain steps and goals of algorithms that can be generalized. Gillespie offers an overview of these in his essay on the relevance of algorithms.[47] The key point to understand here is that even though the mathematical process of an algorithm can be described to be objective, they are based on assumptions that cater to the needs of those designing the algorithms. Thus, even though they can be perceived to be objective, algorithms are based on subjective goals.

　　Gillespie argues that this subjective reality can already be seen in the selection of data included in the application of an algorithm and continues in the analysis as well as application of algorithmic conclusions on the design of users' personalized environment.[48] In that form, an algorithm communicates the premises on which it is programmed to the user. An example of this may be the idea of Google ranking search outcomes according to relevance. But what is relevance exactly? How can you measure it? For Google, relevance can be measured by user behavior. If users for instance were happy with the entries provided by Google, they would not perform another search entry that is worded differently. Google concludes the

---

[45] Merriam-Webster entry, "algortihm," accessed June 10, 2021 on https://www.merriam-webster.com/dictionary/algorithm

[46] Tarleton Gillespie, "The Relevance of Algorithms," in *Media Technologies: Essays on communication, materiality, and society,* ed. Tarleton Gillespie, and Pablo J. Bozkowski (MIT Press, 2014): 185.

[47] Gillespie, "The relevance of algorithms," 168.

[48] Gillespie, "The relevance of algorithms," 179.

original outcomes provided were enough to satisfy this customer, hence they were relevant.[49]
Naturally, this conception of relevance can be contested. And here we arrive at another
possible link to Habermas.

When conceiving of algorithms, much like nudges, as a form of delinguistified
communication, we can also identify validity claims made by the algorithm-based design of
our environment. However, due to the increased complexity of the way in which algorithms
are constructed, the original assumptions of algorithms are difficult to identify. Only in certain
moments, mistakes in the matrix if you will, do the claims on which algorithms are based,
identify themselves. One example may be the categorization mistake made by one of
Amazon's algorithms in which several thousand books with gay-friendly contents were
categorized as "adult".[50]

The issue that the workings of algorithms cannot be explained by laypersons and are
further considered trade secrets now symbolizes a key problem as they evade communicative
rationality and thereby public scrutiny, which could object to the claims on which algorithms
are based. In a sense, it can be said that this evasion of communicative rationality is also a
process of evading responsibility for the claims. Thereby, another parallel can be drawn to
Habermas, who stated:

> Delinguistified media of communication such as money and power, connect up interactions in space and
> time into more and more complex networks that no one has to comprehend or be responsible for. (Chapter VI,
> Section 2E, Theory of Communicative Action Vol. II)

2.4 Algorithms and nudges

The line between nudges and algorithms gets increasingly blurry. Thus, to clear this
distinction up a bit, let us take a closer look at their relation.

As already explained earlier, algorithms as such are not nudges. Most often they work
far from user input and make work processes easier. One frequent example of such an
algorithm is a so-called PID controller.[51] In essence, these controllers sense the current state
of a specific state of affairs (say, temperature), compare this state with what the state should
be like, and make according adjustments. They can be found in thermostats [52], but also

---

[49] Gillespie, "The relevance of algorithms," 175.
[50] Allison Flood, "Amazon apologises for 'ham-fisted' error that made gay books 'disappear'," in The Guardian (April 2009), accessed June 10, 2021 on https://www.theguardian.com/books/2009/apr/14/amazon-gay-sex-rankings-apology
[51] NI, "PID Theory Explained," accessed June 10, 2021 on https://www.ni.com/de-de/innovations/white-papers/06/pid-theory-explained.html
[52] Ibid.

autonomous driving cars.[53] Because of the broad definition, however, algorithms can also be found directly involved with user input. Some examples are algorithms of social media platforms, deciding the order of content, and the type of content presented to users. I want to focus on this latter type of user-involved algorithm as they are encountered in online interaction.

Generally, I argue that these algorithms are a means employed to nudge users towards greater user engagement. After all, I can freely spend my time, the nudging objective of user-involved algorithms is to increase my tendency of engaging with the respective platforms. The distinction between algorithm and nudging in detail however is more intricate.

Take for instance a news outlet presenting different headlines to users. The news outlet has two problems when presenting their news to the world. Firstly, they only have a limited amount of space on their website, not everything can be included. Secondly, much like the cafeteria example, something has to be the first headline a user sees. Thus, there exists an issue of selection of news, and an issue of the ordering of news. The first problem is not a nudging problem. In line with the cafeteria example, the exclusion of junk food is not a nudge since the choice is made for someone. The second problem however is one in which the algorithm takes over the role of a choice architect. It has to decide which type of content will most likely engage the user – how it can get the user to not simply exit the website. Hence, while both problems of the news website may be solved by algorithms consulting data and personalizing their choices of inclusion and order, only the latter problem has a nudging scope.

This practice of individual, data-informed nudging is also referred to as personalized nudging.[54] Overall, there is a distinction between choice personalization and delivery personalization.[55] While the first refers to the direction one wants to nudge someone, the second is concerned with the way this nudge is performed. Thus, delivery personalization in a news context may concern itself with whether images grab the user's attention more or less than sensationalist headlines. Choice personalization in contrast would be concerned with the direction of news (economy, politics, entertainment, etc.) suggested to the user through its order. Personalized nudges are more potent than their classical counterparts as they can take into account the nuances between recipients instead of employing a one size fits all approach. The algorithmic inference is of key importance here. To link back to Habermas, if classical

---

[53] Jeremy Cohen, "Control Command in Self-Driving Cars," in towards data science (August 2018), accessed June 10, 2021 on https://towardsdatascience.com/the-final-step-control-783467095138
[54] Stuart Mills, "Personalized Nudging," *Behavioural Public Policy*, (2020): 1. doi:10.1017/bpp.2020.7.
[55] Mills, "Personalized Nudging," 5.

nudges are considered a form of generalized communication, user-involved algorithms could be considered a form of a steering mechanism. Where the steering mechanism of power and money are backed by means of enforcement and gold, algorithmic personalization is backed by data. Personalized nudging then, is a manifestation of this new steering mechanism.

In summary, the relationship between nudging and algorithmic personalization is threefold. Firstly, algorithmic personalization can be conceived of as a way to nudge me towards an increased engagement with certain platforms. Secondly, classical nudges relate to algorithmic personalization in so far as they enable it (see section 2.2), Lastly, personalized nudging is a manifestation of algorithmic personalization.

2.5 Algorithmic personalization and algorithmic culture

Understanding algorithmic personalization as a steering mechanism is helpful because it enables a clear distinction to classical nudging. While it is possible to opt-out of it, just like the classical nudges, it is significantly more difficult. Whereas a classical nudge would not change depending on the individual behavior of any, one recipient, algorithmic personalization is capable of doing so. If an individual user would not engage with a news website that has a one size fits all approach, this is no reason to change the layout as long as the majority of users show the desired behavior. If, however, a personalized layout does not effectively nudge an individual towards the desired outcome, this is important information that will inform the algorithm for its next attempt to personalize the layout.[56] [57] Thus, in contrast to classical nudging, algorithmic personalization gets more potent with any interaction, no matter if successful or not.

Another point of distinction implied here is the persistence of a personalized online environment. It is important to keep in mind that nudging is built on the human weakness of impulsive decision-making. If confronted with constantly new attempts of nudging behavior (for instance the obligatory default nudges explained earlier), at some point, the impulsive decision making will engage with it. Escaping personalized nudges may be easier said than done.

The objection could be made here that, users are still able to opt-out of the personalized online experience. While true in principle, this would require a certain amount of knowledge that most users do not possess, and more effort than users would want to invest.

---

[56] Gillespie, "Relevance of Algorithms," 183.

[57] Balázs Bodó, Natali Helberger, Sarah Eskens, and Judith Möller, "Interested in Diversity," *Digital Journalism* 7, no. 2 (December 2018): 217. https://doi.org/10.1080/21670811.2018.1521292

Further, it needs to be pointed out that opting out also means losing out. Even when contesting the way user-involved algorithms are utilized, their helpfulness in certain situations need not be denied. Recommendations of products or content can lead to a broadening of the horizon, to a book, a song, or a service that adds value to one's life.[58] By opting out of the personalized algorithmic online experiences, one also loses out on these perks.

Naturally, the use of algorithmic personalization does not only come with perks but also dangers that result out of their misappropriation. One of the most prominent examples of misappropriation is the use of algorithmic personalization within the 2016 U.S. election. Here, the general objective of algorithmic use can be attributed to nudging. Since it is not possible to decide the voting choice of a citizen, the objective was to increase the likeliness of some not-yet convinced voters to vote republican.[59] This was achieved by selecting a personalized advertisement that best aligns with the Facebook profiles of the recipients.

Additionally, the problem of an evaded communicative rationality is still an issue, no matter the perks that algorithmic personalization provides. To elaborate further on the example of the news outlet from earlier, Gillespie observes that the task of selecting the news and deciding on the order of news presented was initially performed by editors.[60] Now, he argues that algorithms are increasingly taking their place. To Gillespie, the main difference between editors and algorithms is that editors are working based on an elaborate ethical codex as journalists.[61] The same is not true for algorithms. Instead, algorithms operate on a form of technologically assured objectivity. However, the concept of an objective (news) algorithm is questionable at best, and impossible at worst.

In contrast to Gillespie, the difference between algorithm and editor identified by this paper is that the editor is capable of being the object of communicative rationality while the algorithm is not. It is much harder to criticize the choices of an algorithm than those of an editor. With algorithmic logic being ingrained more and more into the everyday life of society, a form of algorithmic culture develops, as identified by Striphas:[62]

> What one sees in Amazon, and in its kin Google, Facebook, Twitter, Netflix and many others, is the enfolding of human thought, conduct, organization and expression into the logic of big data and large-scale computation, a move that alters how the category culture has long been practiced, experienced and

---

[58] Andy Cush, "Uneasy Listening: My Year of Surrendering to the Strange, Soothing Power of the YouTube Algorithm," in SPIN (December 2018): accessed June 11, 2021 on https://www.spin.com/featured/youtube-algorithm-music-essay-ambient-hiroshi-yoshimura-midori-takada/

[59] Scott Detrow, "What Did Cambridge Analytica Do During the 2016 Election?" in NPR (March 2018): accessed June 11, 2021 on https://text.npr.org/595338116

[60] Gillespie, "The Relevance of Algorithms," 192.

[61] Ibid.

[62] Ted Striphas, "Algorithmic culture," *European Journal of Cultural Studies* 18, no. 4-5 (2015): 396.

understood. This is the phenomenon I am calling, following Alexander R Galloway (2006), 'algorithmic culture'.

To Striphas, those actors that are the main drives behind algorithmic development in the social context, such as search engines, social platforms, and online retailers are becoming the apostles of said algorithmic culture.[63] By evading a communicative discourse, the world once *entzaubert* (disenchanted) is *verzaubert* (enchanted) by a form of subjective, instrumental rationality that cannot easily be subjected to communicative reason.

Concerning Habermas, this subjective, instrumental rationality that forms the basis of algorithmic culture is dangerous for the reproduction of the lifeworld. After all, while algorithmic culture may be confined to a technological sphere, its consequences can reach far beyond it. As already indicated by Habermas, if the reproduction of the lifeworld is disturbed, a crisis can develop.[64] Some may already be observed: Psychopathologies for instance develop if members of a society are not properly socialized. In line with this, studies on the effects of social media on the human psyche have shown that social media use can lead to an increased tendency to engage in social comparison, which in turn can lead to a decrease in well-being.[65] Further, a disturbance in the reproduction of culture is symptomized by a loss of shared meaning.[66] This can also be argued for, especially when considering the occurrence of filter bubbles. Eli Pariser first argued for this concept, suggesting it occurs, when personalizing algorithms personalize the selection of content presented to the consumer.[67] With time, an information bubble develops around the user, that only provides information judged to be relevant to the algorithms' approximation of the user. This is problematic when considering the nature of public interaction based on shared meaning. If the content is personalized, the development of shared meaning is hindered, and communicative action is more difficult to achieve. This is exacerbated by the possibility that opinions are more polarized by algorithmic recommendations.[68] Another problem arises out of the recommendation of misinformation, for instance, conspiracy theories.[69] Since the

---

[63] Striphas, "Algorithmic culture," 407.
[64] Habermas, *TCA II*, Chapter VI, Section 1D, Kindle version, position 3282.
[65] Erin A. Vogel, Jason P. Rose, Lindsay R. Roberts, and Katheryn Eckles, "Social Comparison, Social Media, And Self-Esteem," *Psychology of Popular Media Culture* 3, no. 4 (2014): 216. http://dx.doi.org/10.1037/ppm0000047
[66] Habermas, *TCA II*, Chapter VI, Section 1D, Kindle version, position 3284.
[67] Eli Pariser, „Beware online 'filter bubbles'" filmed Marcch 2011 on Ted 8:48, accessed June 14, 2021 on https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles#t-9848.
[68] Paul R. Resnick, Kelly R. Garrett, Travis Kriplean, and Sean A. Munson and Natalie Jomini Stroud, "Bursting your (filter) bubble: strategies for promoting diverse exposure," in *Proceedings of the 2013 conference on Computer supported cooperative work companion* (2013): 96. https://doi.org/10.1145/2441955.2441981
[69] Marc Faddoul, Guillaume Chaslot, and Hany Farid, "A Longitudinal Analysis of YouTube's Promotion of Conspiracy Videos," preprint (2020): 6. https://arxiv.org/abs/2003.03318

recommendation is based on the algorithm's approximation of the user, some users might mistakenly be introduced to the rabbit hole of online conspiracy contents, while others, albeit correctly identified, are encouraged in their search for misinformation.

While the literature on the filter bubble is unsure to what extent this phenomenon already takes place,[70] the potential for the development of a filter bubble is there, and its consequences need to be considered. For now, the general problem of algorithmic use to personalize has been sufficiently elaborated upon. Its consequences are clear and, in some cases, already observable. Thus, an important question now is whether there is a solution and if so, what this solution looks like?

If we remain in Habermas's conception of society, a solution might be found in the legal system. While it is true that Habermas also despised a process, he called juridifcation which describes an increasing bureaucratization of the lifeworld through written law and formal processes,[71] the law is also mentioned as a protective measure against increasing systemization.[72] However, simply proposing that this is a problem of jurisprudence is not a constructive solution. To address some key points that should be considered in a legal solution of the problem, a more technical view on the relationship between algorithm and user can be helpful. To do so, Latour's essay on technical mediation is considered in the next chapter.

Chapter 3 – Latour and the Algorithmic Blackbox

In his essay on technical mediation, Latour explains how subject, and object are intertwined in technology.[73] A key concept of his theory is the so-called black-boxing of technology. Both, the intertwined relationship of humans with technology and the idea of black boxing are useful to identify the problem on a technical level. In addition to the normative account already explicated through Habermas, this technical perspective can more concretely identify how we could reapproach the topic of algorithmic logic to suit societal needs.

3.1 The subject-object dichotomy

Latour identifies a key problem in the proposed distinction between object and subject. To illustrate, he gives a famous example of the still prevalent discussion on guns in America. The

---

[70] Marko Haim, Andreas Graefe, and Hans-Bernd Brosius, "Burst of the Filter Bubble? Effects of personalization on the diversity of Google News," *Digital Journalism* 6, no. 3 (2017): 333. https://doi.org/10.1080/21670811.2017.1338145
[71] Habermas, *TCA II*, Chapter VIII, Section 1B, Kindle version, position 7030.
[72] Habermas, *TCA II*, Chapter VIII, Section 2C, Kindle version, position 7946.
[73] Bruno Latour, "On Technical Mediation," *Common Knowledge* 3, no. 2 (Fall 1994): 29-64.

two positions according to Latour can be distinguished as either materialist ("Guns kills people"), or sociological ("People kill people").[74] The mistake of both is that they assume a possible distinction between objects (i.e., guns) and subjects.[75] Instead, the relationship between object and subject according to Latour is bidirectional. Whenever an individual picks up a gun, both are transformed into a new composition, a "citizen-gun" in Latour's words.[76] Thus, if the gun, as well as the person, are two different agents, their combination can produce a third agent, that has a distinguishably different goal. This process, termed translation, is the first of four meanings of technical mediation.[77]

3.2 Four meanings of mediation

Translation is the creation of different composite agents and in consequence different goals also find their relevance in an algorithmic setting. User and algorithm are two parts of a feedback loop that often leads to the creation of new goals. A student intending to watch a YouTube video on mathematics might end up longer on the platform than he intended to since new videos are constantly being recommended that are personalized to the student's interests (probably not math-related, however). The second meaning of mediation is termed composition.[78] Here Latour suggests that action is unjustifiably attributed to subjects, even though objects are just as necessary to performs said actions. If humans travel to the moon, the rockets are just as necessary as the astronauts. Again, a link to algorithmic logic can be established. When the student inevitably realized that he ended up spending a lot more time on the platform than he intended to, the realization could also be rephrased in that the platform spent a lot of time with the student, gaining new data, learning about how the student can be engaged to stay on the platform.

The origin for this misappropriation of action can be found in the third meaning of mediation, reversible black boxing.[79] Technology, if it performs as it is intended tends to become opaque. One does question the functioning of technologies so long as they do what they are intended to do. Only when technologies break down, there is awareness of this ignorance towards them. This is illustrated by the example of an overhead projector breaking down during a lecture. Suddenly, the overhead is brought from the peripheral of our attention to the focus. Now, relations that were initially concealed in the black-boxed nature of the

---

[74] Latour, "On Technical Mediation," 31.
[75] Latour, "On Technical Mediation," 33.
[76] Latour, "On Technical Mediation," 32.
[77] Latour, "On Technical Mediation," 32.
[78] Latour, "On Technical Mediation," 35.
[79] Latour, "On Technical Mediation," 37.

projector are apparent (e.g., the professor's dependence on a functioning overhead or the repairman's dependence on a broken overhead projector). With algorithmic logic, there is a prevailing ignorance of users towards the working of an algorithm and its assumptions. This status of a black-boxed algorithm is further emphasized by the complexity of the algorithm (in some cases even the originators have difficulty explaining the working of an algorithm), and the non-transparent policies of companies. Lastly, the most important meaning of technical mediation is the ability of objects to act on behalf of subjects, called "delegation".[80] Interestingly, Latour uses the example of a speed bump here, a notorious nudging example. Latour posits that speedbumps transform the situation from "negotiable" speed signs to "non-negotiable" speed bumps, offering a parallel to the argument made on the algorithmic personalization of news and the difference between humans and algorithmic editors. However, next to shifting a situation from negotiable to non-negotiable, intentions are also shifted. Whereas the intention of a driver taking care to respect the speed limit was to drive safely for the sake of others, a speedbump transforms this solidarity-focused intention into an egocentric intention to not damage the care by driving too fast. The speedbump is an expression of delegation since by it the system (i.e. the city, politics, laws) is acting without being present, delegating its intent to be fulfilled by the object. A similar point is made earlier when discussing the inability to question the subjective basis of algorithmic functioning. When considering the earlier example of a student ending, losing track of time on YouTube even though the initial aim was to watch a video on mathematics, the translation from the original goal to a new goal, gets a normative dimension within the context of delegation. I argue that as already identified within Habermas, this problem is in large part due to the black-boxing of algorithms, that does not allow for negotiation.

3.3 Black-boxing algorithms

Black boxing happens in a succession of steps and should not be regarded as an intentional product but rather as a natural consequence of technology. In his essays, Latour describes black boxing as:

> The way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become. ("Glossary", *Pandora's Hope: Essays on the Reality of Science Studies*, 304)

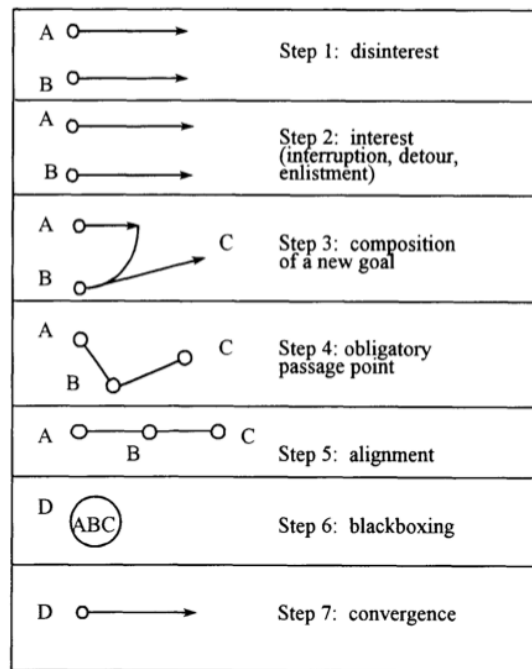
[80] Latour, "On Technical Mediation," 39.

Figure 1. Steps of reversible black-boxing.[81]

Figure 1 shows the different stages of black-boxing according to Latour. To illustrate the process of the different steps, think of a primate trying to reach fruit, eventually enlisting a stick as a means to his end. At first, he might try to reach the fruit himself, demonstrating his disinterest in the stick (Step 1). However, he quickly realizes that he is not able to reach the fruit, it is too far away. Thus, he searches for a stick long enough to help reach the fruit (Step 2) and might realize that this technique could also help in reaching termites, forming a new goal (Step 3). By some trial and error, he eventually learns the best technique to reach termites and fruits with his rudimentary tool. This technical skill is what Latour refers to in Step 4. When technical skill, technical tool, and goal are aligned (Step 5), black-boxing is the next step. The primate now knows he reaches termites and fruits beyond his reach. What he should be thinking however is, that he *and* the stick can reach termites and fruits beyond his reach. Admittedly, the stick is a rather rudimentary example of a technical tool. Nonetheless, it shows how black-boxing follows naturally: the stick, in its success, seems to remove itself from the cognitive equation of the primate, much like we disregard the technical means of space travel when saying that mankind travels to the moon.

However, once a technology has reached the status of a black box, it does not mean that it will be disregarded no matter what. Instead, it could be argued that after convergence a new cycle may start a new, leading to ever more complicated technologies, i.e., black boxes within black boxes. Further, black boxes may enlist our interest when they no longer work

---

[81]Figure from Latour, "On Technical Mediation," 37.

like they are supposed to. Latour made this point by illustrating a situation in which a lecture is interrupted because the overhead projector broke, which brings technology out of the periphery and into the focus of our attention. This process of enlisting interest can also be observed in algorithmic culture. While the general status quo may be an unawareness in algorithmic processes,[82] this quickly changes when there are scandals (e.g., Cambridge Analytica) or general errors in the algorithmic working (consider the already mentioned categorization error of the amazon algorithm).

Curiously, this is not the only way public interest may be enlisted. An interesting example to consider from the algorithmic literature are so-called whoa-moments.[83] These occur whenever there is an uncanny hit by the algorithm such as getting ads for things mentioned in an earlier discussion. In a sense then, users become aware of the black-boxed algorithm when it is working too well when it offers a glimpse of the cards it is playing with. Thus, for algorithms, black boxing occurs, when there are no errors in its functioning and when its success does not arouse the suspicion of its users.

Another characteristic of algorithmic logic is, that while usually, the prime mover is the actor (i.e. the primate picking up the stick, developing a new goal along the way), this may not be the case in algorithmic logic. Here, it is not entirely clear who enlists who. The student approaches the platform intending to learn something. As soon as this happens, the algorithm enlists the user intending to create a new composed goal that keeps the user engaged for longer. The user, however, is in the first instance enlisting the website, not the algorithm. Thus, from a user perspective, there exists the tendency that algorithmic structures act as the prime mover, enlisting actors for their means. Contrary to the usual process of a new goal composition, however, the aim of the algorithm never changes after enlistment. It is always trying to fulfill whatever objective it was programmed to fulfill. Sure, the goals of the actor might change (e.g., from watching for educational purposes to watching for entertainment purposes), but the same is not true for the algorithm.

As opposed to other techniques, algorithms are depending on user input to give a certain output, they constitute a feedback loop. As such, each member of the feedback loop has at least some agency in the eventual outcome. This has been realized increasingly by users of algorithms. And algorithms are now more and more explicitly engaged with. This leads to

---

[82] Kevin Hamilton, Christian Sandvig, Karrie Karahalios, and Motahhare Eslami, "A Path to Understanding the Effects of Algorithm Awareness," *CHI 2014* (Spring 2014): 632. http://dx.doi.org/10.1145/2559206.2578883
[83] Tania Bucher, "Neither Black Nor Box: Ways of Knowing Algorithms," in *Innovative Media and Communication Research*, ed. Sebastian Kubitschko, and Anne Kaun (Switzerland, Cham: Springer International Publishing AG, 2016): 91.

what Gillespie calls a "domestication" of technology.[84] Algorithms may be acted with in ways that the originator did not account for. While some implement the algorithmic logic into their daily life, for instance by explicitly searching for music with YouTubes recommendation system, others instead take the opposite route and try to separate algorithmic logic as much as possible from their sense of self.

With increasing public attention to data collection, another route to protect against a possibly detrimental effect of black-boxed algorithmic culture becomes apparent. Increasingly privacy is becoming a marketable good. Cynics may point out that this is yet another instance of the market selling the solution to a problem that was created by it in the first place. Nonetheless, recent examples of Apples' non-tracking function or Firefox's decision to set the default to "non-acceptance" of not necessary cookies exemplify possible consequences of increased public awareness.[85] [86] First data shows that an overwhelming majority are using Apple's function.[87] Naturally, there is opposition to these consequences as Facebook argues that businesses will suffer from this, including small businesses trying to get more attention for their product.[88] There is some truth to this claim, as it points out that personalization is not per se a problematic idea – most users will inevitably be introduced to some product that may have a positive impact on their lives, be that a new book or a new type of musical genre.

While the public awareness of data tracking and personalized algorithmic nudging may lead to a reversal of the default nudge or the use of anti-tracking software, these practices lead to a disengagement with the black-boxed algorithm. Instead of changing the black box, they abandon it. Therefore, I think this should not the be-all-end-all solution to the problem, as it denies the positive potential algorithmic personalization has.

4. A Possible Solution

[84] Gillespie, "The Relevance of Algorithms," 185.
[85] Sergiu Gaitan, "Firefox now blocks cross-site tracking by default in private browsing" on BleepingComputer (June 1, 2021), accessed June 11, 2021 on https://www.bleepingcomputer.com/news/security/firefox-now-blocks-cross-site-tracking-by-default-in-private-browsing/
[86] Press release Apple, "Apple advances its privacy leadership with iOS 15, iPadOS 15, macOS Monterey and watchOS 8" on apple.com (June 7, 2021), accessed June 11, 2021 on https://www.apple.com/newsroom/2021/06/apple-advances-its-privacy-leadership-with-ios-15-ipados-15-macos-monterey-and-watchos-8/
[87] Margaret Taylor, "How Apple screwed Facebook," on Wired (May 19, 2021), accessed June 11, 2021 on https://www.wired.co.uk/article/apple-ios14-facebook
[88] Ibid.

The problem that was emphasized by the Theory of Communicative Action is that algorithms and nudging can be misappropriated, which leads to a disturbed reproduction of the lifeworld. However, it's helpful to re-introduce an earlier quote from Habermas:

> Both are conceivable: the institutions that anchor steering mechanisms such as power and money in the lifeworld could serve as a channel either for the influence of the lifeworld on formally organized domains of action or, conversely, for the influence of the system on communicatively structured contexts of action. (Chapter 6, Section 2F, *Theory of Communicative Action Vol. II*)

In much the same way, nudges and algorithmic personalization can help instill the lifeworld more within the system. The increasing demand for privacy-related features such as already described by Apple and Firefox shows that this influence of the lifeworld on the system is gaining momentum. As indicated by black-boxing, users were not intentionally engaging with the algorithm but rather with the website. The unawareness of the black-boxed algorithm and the effort needed to opt-out of it were the main reasons for engaging with the algorithm at all. However, as already discussed, algorithmic logic may also broaden horizons and therefore be of positive value to users.

In yet other cases, personalization and nudging practices can improve social interaction. One example is the problem of misinformation. A recently published study in Nature showed that the practice of priming may reduce the spreading of misinformation.[89] This type of nudging is performed by providing environmental cues that guide decision-making.[90] In this case, researchers primed participants to be more attentive regarding the perceived accuracy of headlines. While the researchers tested their hypothesis by messaging users who spread misinformation to rate the accuracy of specific headlines, the nudge could prove more effective if personalized. Users who often spread misinformation might for instance receive a report on how accurate their shared headlines are, according to fact-checkers. Further, algorithmic personalization could also be used to challenge clearly manifested opinions instead of reinforcing them, enabling an open online discussion. While the literature on personalized nudging for public policy is scarce, there is a huge potential in nudging. How then, are we able to find a solution that includes both, the upsides of personalization and the agency of the user?

Giving users a transparent choice could be a start. A comprehensive description of what is collected, used, and what it is used for is already a feature of any cookie notification. However, a description falls short when aiming at a deliberative solution in which users have

---

[89] Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand, "Shifting attention accuracy can reduce misinformation online," *Nature* 592 (March 2021): 590-595. https://doi.org/10.1038/s41586-021-03344-2

[90] Thaler, and Sunstein, *Nudge*, 70-73.

agency. Instead, a possible option may be that browsers offer a mandatory quiz for users in which a personal default can be customized that is automatically applied to any visited content provider. A default setting in this approach need not be black or white, all the cookies or none of them. Instead, a quiz might ask if personalization should be performed selectively. Maybe a user decides that book advertisements should be personalized, while political content should not be personalized. This option would give users their agency, providers their data, and companies depending on personalized advertisement their customers. Algorithms would constitute black boxes that lose their opaqueness when users are asked for their preferences and regain it afterward, disappearing in their successful working. This model would enable an online environment in which algorithms can remain trade secrets while their subjective basis is up to the consumer.

Another question is how this solution should be applied. In line with an earlier mentioned conception of law that protects the lifeworld, lawmakers should recognize their responsibility and start making customization obligatory over the current, passive personalization. Other propositions in law are also possible, for instance, the banning of fake news content or categorically not allowing personalization to be applied for scientific, and political content.

However, it can also be argued that the market is already self-correcting. The earlier examples of Apple and Firefox make apparent that privacy is increasingly marketable. To retain a viable business model, companies relying on personalization such as Facebook may soon need to adapt to the privacy awareness of their customers. Customization thus could be a natural consequence arising out of increased demand for an agency.

5. Conclusion

This paper set out to frame nudging and algorithmic personalization within the Theory of Communicative Action. Illustrating nudging from Habermas's perspective offered the idea that design has the potential to communicate with those interacting with it. This concept was subsequently transferred to algorithmic personalization. The problem identified in algorithmic personalization is that by imposing it on users, the agency is not shared evenly between algorithm and user. Because of this imbalanced agency, communicative rationality is difficult to achieve, rendering algorithmic personalization a steering mechanism that depends on the extensive amount of data collected through the default nudge. This practice of a passive personalization process is dangerous as it risks the development of a shared lifeworld. In

search of a solution to this dilemma, the different meanings of technical mediation by Latour were consulted. Here, the concept of black-boxing was helpful to identify that users are not actively engaging with algorithms but rather, that algorithms are passively imposed on users. To avoid this problem, the proposed solution focused especially on the concept of active customization to oppose passive personalization.

It is important to point out once again that not algorithms themselves are criticized but rather the process in which these algorithms are imposed on the user. Considering the amount of data that is present online, the use of algorithms is without alternative. However, when employing algorithms in a personalizing context, it is important to realize their influence on an individual and societal level. As can be taken from Latour's second meaning of technical mediation, it is necessary to recognize that objects have active responsibility. We need to start holding algorithms to this standard and recognize them as communicative agents in our society. Per Habermas, we then also must actively start to criticize their validity claims, the information that is implied in every algorithmically determined decision. Not doing so could prove fatal for the reproduction of a lifeworld. At the same time, it is important to remind oneself that it is not possible to design a value and validity claim-free online environment. This paper proposes to make customization mandatory to actively engage in a communicative process that gives users agency in the design of their environment. Future research is encouraged to investigate more options to enable this communicative process and investigate the feasibility of the already proposed process.

Another domain of future deliberation is to limit the possibly detrimental effects through lawmaking. While currently a trade secret, algorithmic personalization should more closely be monitored by lawmakers, to ensure that subjective underpinnings are not working against a pluralistic opinion formation. Especially with consideration of filter bubbles, there exists a potential of an increasingly polarized opinion formation that will greatly inhibit the public discourse.

Bibliography

Apple, press release. "Apple advances its privacy leadership with iOS 15, iPadOS 15, macOS Monterey and watchOS 8." On apple.com (June 7, 2021). Accessed June 11, 2021 on https://www.apple.com/newsroom/2021/06/apple-advances-its-privacy-leadership-with-ios-15-ipados-15-macos-monterey-and-watchos-8/

Berendezen, Joseph. "Max Horkheimer." In *Stanford Encyclopedia of Philosophy.* Stanford University, Summer 2021. Article published June 2009; substantiative revision June 2019. https://plato.stanford.edu/entries/horkheimer/

Bodó, Balázs, Natali Helberger, Sarah Eskens, and Judith Möller. "Interested in Diversity," *Digital Journalism* 7, no. 2 (December 2018): 206-229. https://doi.org/10.1080/21670811.2018.1521292

Bohman, James, and William Rehg. "Jürgen Habermas." In *Stanford Encyclopedia of Philosophy*. Stanford University, Summer 2021. Article published May 2007; minor correction Fall 2017. https://plato.stanford.edu/entries/habermas/.

Bruns, Axel and Tim Highfield, ed. "Is Habermas on Twitter? Social Media and the Public Sphere." In *The Routledge Companion to Social Media and Politics,* edited by Axel Bruns, Gunn Enli, Eli Skogerbø, Anders Olof Larsson, Christian Christensen, 56-70. New York: Routledge, 2015.

Bucher, Tania. "Neither Black Nor Box: Ways of Knowing Algorithms." In *Innovative Media and Communication Research*. Edited Sebastian Kubitschko, and Anne Kaun. Switzerland, Cham: Springer International Publishing AG, 2016: 81-98.

Cohen, Jeremy. "Control Command in Self-Driving Cars." In towards data science (August 2018). Accessed June 10, 2021 on https://towardsdatascience.com/the-final-step-control-783467095138

Cush, Andy. "Uneasy Listening: My Year of Surrendering to the Strange, Soothing Power of the YouTube Algorithm." In SPIN (December 2018). Accessed June 11, 2021 on https://www.spin.com/featured/youtube-algorithm-music-essay-ambient-hiroshi-yoshimura-midori-takada/

Deloitte. "Cookie Benchmark Study," (2020): 6. Accessed on June 10, 2021 on https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/risk/deloitte-nl-risk-cookie-benchmark-study.pdf

Detrow, Scott. "What Did Cambridge Analytica Do During the 2016 Election?" In NPR (March 2018). Accessed June 11, 2021 on https://text.npr.org/595338116

Duden entry ,"Quengelware," accessed June 10, 2021 on https://www.duden.de/rechtschreibung/Quengelware

Faddoul, Marc, Guillaume Chaslot, and Hany Farid. "A Longitudinal Analysis of YouTube's Promotion of Conspiracy Videos." Preprint (2020): 1-8. https://arxiv.org/abs/2003.03318

Flood, Allisson. "Amazon apologises for 'ham-fisted' error that made gay books 'disappear'."
    In The Guardian (April 2009). Accessed June 10, 2021 on
    https://www.theguardian.com/books/2009/apr/14/amazon-gay-sex-rankings-apology

Gillespie, Gillespie. "The Relevance of Algorithms." In *Media Technologies: Essays on
    communication, materiality, and society,* edited by Tarleton Gillespie, and Pablo J.
    Bozkowski. MIT Press, 2014: 167-193.

Habermas, Jürgen. *Theory of Communicative Action Vol I.* Translated by Thomas McCarthy.
    Cambridge, UK: Blackwell Publishers Ltd., 1984. Kindle.

Habermas, Jürgen. *Theory of Communicative Action Vol II.* Translated by Thomas McCarthy.
    Cambridge, UK: Polity Press, 1987. Kindle.

Haim, Marko, Andreas Graefe, and Hans-Bernd Brosius. "Burst of the Filter Bubble? Effects
    of personalization on the diversity of Google News." *Digital Journalism* 6, no. 3 (2017):
    330-343. https://doi.org/10.1080/21670811.2017.1338145

Hamilton, Kevin, Christian Sandvig, Karrie Karahalios, and Motahhare Eslami. "A Path to
    Understanding the Effects of Algorithm Awareness." *CHI 2014* (Spring 2014): 631-642.
    http://dx.doi.org/10.1145/2559206.2578883

Harris, Tristan. "Humane: A New Agenda for Tech with Tristan Harris – Presentation and
    Transcript". In Ethical (May 2019). Accessed June 15, 2021 on
    https://ethical.net/ethical/humane-new-agenda-tech-tristan-harris/

Hove, Thomas. "Understanding and Efficiency: Habermas's Concept of Communication
    Relief." *Communication Theory* 18, no. 2 (April 2008): 240-254.
    https://doi.org/10.1111/j.1468-2885.2008.00321.x

Kahnemann, Daniel. *Schnelles Denken, Langsames Denken*. Translated by Thorsten Schmidt.
    München: Pantheon Verlag, 2011.

Kramer, Adam D. I., Jaime E. Guillory, and Jeffrey T. Hancock. "Experimental evidence of
    massive-scale emotional contagion through social networks." *Proceedings of the National
    Academy of Sciences* 111, no. 24 (2014): 8788-8790.
    https://doi.org/10.1073/pnas.1320040111

Latour, Bruno. "On Technical Mediation." *Common Knowledge* 3, no. 2 (Fall 1994): 29-64.

Latour, Bruno. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA:
    Havard University Press, 1999.

Merriam-Webster entry, "algortihm," accessed June 10, 2021 on https://www.merriam-
    webster.com/dictionary/algorithm

Mills, Stuart. "Personalized Nudging." *Behavioural Public Policy*, (2020): 1-10.
    doi:10.1017/bpp.2020.7.

National Instruments Corporations. "PID Theory Explained." Accessed June 10, 2021 on
    https://www.ni.com/de-de/innovations/white-papers/06/pid-theory-explained.html

Obar, Jonathan A., and Anne Oeldorf-Hirsch. "The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services." *Information, Communication & Society* 23, no. 1 (June 2018): 128-147. https://doi.org/10.1080/1369118X.2018.1486870

Pariser, Eli. „Beware online 'filter bubbles'." Filmed March 2011. Video, 8:48. Accessed June 14, 2021 on https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles#t-9848.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand. "Shifting attention accuracy can reduce misinformation online." *Nature* 592 (March 2021): 590-595. https://doi.org/10.1038/s41586-021-03344-2

Resnick, Paul R., Kelly R. Garrett, Travis Kriplean, and Sean A. Munson and Natalie Jomini Stroud. "Bursting your (filter) bubble: strategies for promoting diverse exposure." In *Proceedings of the 2013 conference on Computer supported cooperative work companion* (2013): 95-100. https://doi.org/10.1145/2441955.2441981

Rhoen, Michael. "Beyond consent: Improving data protection through consumer protection law." *Internet Policy Review* 5, no. 1 (March 2016): 1-15. https://doi.org/10.14763/2016.1.404

Sergiu Gaitan. "Firefox now blocks cross-site tracking by default in private browsing." On BleepingComputer (June 1, 2021). Accessed June 11, 2021 on https://www.bleepingcomputer.com/news/security/firefox-now-blocks-cross-site-tracking-by-default-in-private-browsing/

Smith, Craig M., Daniel G. Goldstein, and Eric J. Johnson. "Choice Without Awareness: Ethical and Policy Implications of Defaults." *American Marketing Association* 32, no. 2 (Fall 2013): 159-172. https://doi.org/10.1509%2Fjppm.10.114

Striphas, Ted. "Algorithmic culture." *European Journal of Cultural Studies* 18, no. 4-5 (2015): 395-412.

Sunstein, Cass R., and Richard Thaler. "Libertarian Paternalism." *The American Economic Review* 93, no.2, (Spring, 2003): 175-79. http://www.jstor.org/stable/3132220.

Taylor, Margaret. "How Apple screwed Facebook." On Wired (May 19, 2021). Accessed June 11, 2021 on https://www.wired.co.uk/article/apple-ios14-facebook

Thaler, Richard H., and Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth. And Happiness.* New York: Penguin Press, 2008.

Vogel, Erin A., Jason P. Rose, Lindsay R. Roberts, and Katheryn Eckles. "Social Comparison, Social Media, And Self-Esteem." *Psychology of Popular Media Culture* 3, no. 4 (2014): 206-222. http://dx.doi.org/10.1037/ppm0000047