ELIANA BERGAMIN

AUGUST 2021

SOCIAL NETWORK SITES ADDICTION BY DESIGN

A compatibilist free-will analysis towards policy intervention

SOCIAL NETWORK SITES ADDICTION BY DESIGN

A compatibilist free-will analysis towards policy intervention

A Master Thesis

by Eliana Bergamin

(s2414368)

First Supervisor:

Dr. Patrick T. Smith

Second Supervisor:

Dr. Nolen Gertz

MSc Philosophy of Science, Technology and Society (PSTS)

Faculty of Behavioural, Management and Social Sciences
University of Twente
August 2021

A Rosalia e Maria

Table of contents

Summary **4**

Acknowledgements **6**

Introduction **7**

1. What is Addiction? **11**

1.1 From Plato to Hume: freedom of action and the choice of the Good *12*

1.2 Contemporary compatibilism: valuation system and addiction *14*

1.3 Identity and society: from the sense of self to social identity *18*

1.5 Mediated addiction: what happens on social networks doesn't stay on social networks *22*

2. Are Social Network Sites designed for addiction? **24**

2.1 The taxonomy of social media: a working definition of Social Network Sites *24*

2.2 How do SNSs work? *27*

2.3 Addiction by Design *29*

2.3.1 Habitual performing 31

2.3.2 Detrimental consequences to users' life and well-being: impairment of free will 32

2.3.3 Social identity and self-categorization 33

Conclusion *35*

3. Taking action: philosophical scenarios for policy intervention **36**

3.1 SNS design for addiction: why is it politically relevant? *36*

3.2 Moral relevance: how SNS addiction by design is normatively objectionable *38*

3.3 Relevant public interest: how SNS addiction by design threats current regulations *40*

3.4 Policy intervention: GDPR 2.0 – General Design Protection Regulation *43*

3.5 Potential objections *44*

3.5.1 Technological determinism – an Ellulian objection 44

3.5.2 One must imagine the user happy: A case for non-addiction or for an addictive society 46

3.5.3 Individuals and tech companies: two alternatives to state regulation 48

Conclusions and final thoughts **50**

References **52**

Summary

There is a growing demographic that uses Social Network Sites (SNSs) daily. With more than one billion users, Instagram is considered one of the fastest-growing social media platforms in history. Together with this, an increasing demographic which complains about a lack of self-control when it comes to the amount of time they spend on these platforms. In my thesis, I investigate the concept of Social Networking Sites addiction. I ask: *Are Social Network Sites addictive? And, if yes, should their addictiveness be regulated?*

To answer this question, I first need to outline a philosophical definition of addiction. Through an introductory analysis of the concept of addiction as a lack of free will from a compatibilist perspective, I apply Hanna Pickard's framework and mediation theory to the topic. This aims to show how the typical (mis)conception of addiction does not take into consideration the sense of self and social identity built around the addict and the role technologies play in mediating addiction. Through this analysis, addiction is defined as a lack of alignment between motivation and evaluation over a specific action, fostered by a combination of habit, self-categorization, and social identity of the addict. SNS companies take advantage of the vulnerability of users building their platforms to normalize problematic SNS use. They can establish addictive mechanisms using strategic methods and algorithms aimed at weakening agents' reflective power. I suggest discussing the role of SNSs in addictive behavior through the lens of mediation theory. Mediation theory helps show how SNS technologies mediate the addictive patterns users develop through affordances embedded in them. This view complements the compatibilist definition of addiction, as the problematic design patterns can be directly linked to the singular aspect that characterizes addictive behavior.

The second chapter of my thesis aims at applying this theoretical framework to practical cases of SNS design. Here I show how the aspects that characterize addiction described in chapter one match with SNS design choices. Chapter One and Two aim at answering the first part of my research question.

Once I have shown that the design choices behind these platforms reflect the addictive theoretical framework, the final chapter of my thesis is dedicated to justifying the introduction of policy intervention that will mitigate the current way SNSs are designed. The choice of policy intervention is justified because Social Network Sites, the way they are currently designed, do not respect the political value of autonomy. I narrow my perspective to the analysis of deliberative democracies, as described by Samuel Freeman. He states that public interest in deliberative democracies is defined as the circumstances that can enable and maintain the freedom, independence, autonomy, and equality of citizens. By threatening political autonomy through addictive and exploitative design, SNS design n be regulated through government intervention. I outline a GDPR

2.0 – General Design Protection Regulation – that will ensure the protection and respect of users' autonomy through design choices. Finally, I address some objections that could be made to my argument.

Acknowledgements

I would like to express my gratitude to a few people who have contributed to the realization of this thesis. First, I would like to thank D. Patrick T. Smith for his constant guidance and support throughout the completion of my project. His insights and motivation have helped me immensely in writing my dissertation and my personal development as a young researcher. Secondly, I want to thank Dr Nolen Gertz for his valuable input and comments. His perspective allowed me to expand my philosophical horizon and gather the courage to ask uncomfortable questions. Thank you to all my fellow PSTS classmates who have put up with me in my endless discussions and debates. Your support and input have been fundamental in my research process. I want to thank all my friends, old and new, for constantly encouraging me and believing in me, especially when I hesitate to believe in myself. Finally, to the family I left back home in Italy and to Benedetta, *grazie*. In these unusual times and throughout my whole life, you have been and continue to be my source of strength and unconditional love. To all these people, *grazie*.

The narratives that people recount about us shape the identity of a person. Thank you because, without your narrative, I wouldn't be the person I am today.

Introduction

On March 31st, 2020, the Facebook website published an article written by Nick Clegg, VP of Global Affairs and Communication at Facebook. Its title, *You and the Algorithm: It Takes Two to Tango*, alludes to the fact that the manipulation of free will put into action by SNS companies is nothing but a dystopian depiction. Clegg stated that the personalized ads, content ranking and engagement increase are part of the dynamic relationship between users and algorithms. He claimed that users are active participants in creating their SNS experience and that personalization of content is now a key feature of internet services. One of the most striking sentences in his article states:

> "But ultimately, content ranking is a dynamic partnership between people and algorithms. On Facebook, it takes two to tango."
> (Clegg, 2021)

The message Clegg tries to convey is simple: Social Network users are making use of the platform and feeding the algorithm Therefore the fact that its mechanisms end up manipulating their behavior online is attributable to themselves. However, this impoverished counterattack does not consider the power imbalance that is at the center of the relationship between users and algorithms. The agency users exert on the platform is intrinsically limited by the mechanisms that underlie Social Network algorithms. Users are free to exert their power of choice within limits imposed by platforms. Clegg's article, in an attempt to do justice to the public, also mentions how SNS platforms should be more transparent towards users:

> "Companies like Facebook need to be frank about how the relationship between you and their major algorithms really works. And they need to give you more control."
> (Clegg, 2021)

This claim aims at opening a space for dialogue between users and the way SNS platforms are designed. However, when claiming that SNS companies need to give users more control, Clegg implicitly admits that users do not have enough control yet.

Facebook also replied to the accusations moved in the documentary *The Social Dilemma*, in a document titled *What the 'Social Dilemma' gets wrong* (Aghadjanian, 2020; *What 'The Social Dilemma' Gets Wrong*, 2020). The first point they address in the document is addiction, stating that:

> “Facebook builds its products to create value, not to be addictive.”
> (*What ‘The Social Dilemma’ Gets Wrong*, 2020)

The company affirms that its product teams are not encouraged to design features aimed at increasing screen time. As evidence of this, the company stated that in 2018 they modified their News Feed ranking to give less space to viral videos. This determined a reduction of 50 million hours a day spent on the platform. In Facebook’s perspective, this proves that the company does not design for addiction but aims to provide content with a positive impact on users’ well-being (*What ‘The Social Dilemma’ Gets Wrong*, 2020). In response to this, I shall provide two points. Firstly, many ex SNS companies employees – including Sandy Parakilas, former platform operations manager at Facebook, and Sean Parker, Facebook’s founding president – explicitly stated that these platforms are deliberately designed to establish addictive patterns (Andersson, 2018).

> "I don't know if I really understood the consequences of what I was saying, because [of] the unintended consequences of a network when it grows to a billion or 2 billion people and... it literally changes your relationship with society, with each other... It probably interferes with productivity in weird ways. God only knows what it's doing to our children's brains.”
>
> – Sean Parker, founding president of Facebook –
> (Allen, 2017)

In this context, Facebook’s words on a two-page document sound like a vain attempt to disguise what is actually happening. Secondly, giving their words the benefit of the doubt and accepting the fact that they are diminishing users screen time still does not mean that their products are not designed to be addictive. The fact that SNS platforms like Facebook have reduced users’ screen time does not per se determine the non-addictiveness of their product. Moreover, providing time management tools such as daily reminders, limits on notifications and activity dashboards does not increase users’ control on the product (*What ‘The Social Dilemma’ Gets Wrong*, 2020). If anything, it shows how little self-control people have on these platforms, to the point of needing an external barrier to stop them from performing the activity itself. The need for change that Facebook and SNS companies are advocating for should not shift the responsibility on the users’ side. Counting more than 2.7 billion active users per month, Facebook is currently the most used social network site in the world

(*Facebook by the Numbers*, 2021). Instagram comes right after that, with more than one billion monthly users (A. Post, 2020). Looking at these numbers, the words of Sean Parker become even more worrisome. In the interview, Parker mentions dopamine hits, feedback loops and exploitations. The vocabulary he uses is borrowed from the discipline of persuasive design. That digital technologies are designed to persuade and hook users appears, in everyday discourse, as an established fact.

This thesis originated from an observation that had become a pattern in my everyday life. Many people are using Social Networks, a lot of them seem to lack control over the way they interact with these platforms, and yet individuals appear powerless when it comes to limiting this behavior. From this initial observation, I decided to explore the topic of Social Network Sites addiction from a philosophical perspective. Thus, my thesis aims to answer the question: *Are Social Network Sites addictive? And, if yes, should their addictiveness be regulated?* My research focuses on proving how Social Network Sites are designed for addiction. Subsequently, I present a potential intervention that I have outlined to tackle the problem. To do so, I have divided my paper into three chapters.

In Chapter One, I present a philosophical analysis of the concept of addiction from a philosophical perspective. Here, I aim to answer the sub-question: *what is addiction from a philosophical perspective?* I start off my investigation by reviewing the concepts of addiction in early Western philosophy. I consider the Platonic concept of *akrasia* (lack of self-control), and I compare it to Hume's understanding of freedom of action. Subsequently, by referring to the works of Gary Watson and R. Jay Wallace, I outline a conception of addiction from a compatibilist perspective (Wallace, 1999; Watson, 1975b). This shows how, from the lack of alignment of motivational and valuation systems, together with other factors that characterize the philosophical conception of addiction, addicts find themselves having their ability to deliberate unduly burdened. This chapter subsequently introduces some concepts elaborated from the research on addiction carried out by Hanna Pickard (Pickard, 2020b). In her work, Pickard shows how the identity of the addict subject is shaped by the social contextualization they live in, which in turn shapes their own self-conceptualization. This focus on self and social identity results particularly useful to show how Social Network Sites shape the way users interact with themselves and within society, and how this shaping incentivizes addictive patterns. The discourse on addiction is linked to the realm of philosophy of technology through a mediation theory analysis. By introducing the concept of "mediated addiction", I suggest that framing addiction simply as an interaction dismisses the role of the technology in the addictive behavior. When addiction is technologically mediated, such as in the case of Social Network Sites, there is a mutual co-shaping of addict subjects and addictive substance. As mentioned by Pickard, and confirmed by Verbeek and Gertz, the subject is not a *user* only when they perform the

addictive behavior. Their whole identity and mode of life is shaped by it (Pickard, 2020b). In a mutually shaping relationship, the design of the technology plays a central role.

To research the role of design in Social Network Sites, Chapter Two will focus on specific design methodologies that SNSs adopt in order to instill addictive behaviors in users. This chapter aims to answer the sub-question: *Are SNSs designed for addiction?* Before describing each design component, I present a taxonomy of Social Network Sites (SNSs), emphasizing what distinguishes them from social media. After that, I focus on how SNSs work, describe what adaptive algorithms are, and how they can influence neural patterns. Finally, the last part of the chapter describes the design aspects that contribute to establishing addictive patterns in users, and the consequences that these have. My research considers design choices such as infinite scrolling, the intermittent variable reward and effects such as habitual performing and FOMO.

Finally, Chapter Three tackles the second part of my research question: *Should the addictiveness [of SNSs] be regulated?* Here, I answer the question of why SNS design for addiction is politically and morally relevant. I argue that their design undermines the political principle of autonomy. By adopting the theoretical account of deliberative democracy, I firstly show how SNS design for addiction is relevant from a public interest perspective and how this entails consequences on a normative level. Subsequently, I will argue for the need of policy regulations on SNS design, presenting ideas for a GDPR 2.0 – General Design Protection Regulation, based on the model of the GDPR. Finally, I address some potential objections that could be made to my argument and provide some inspiration on how to further investigate the topic based on these objections

1. What is Addiction?

In this chapter, I aim to answer the question: *What is addiction from a philosophical perspective?* As I will show, addiction has been a topic of interest in the philosophical discourse since the dawn of Western philosophy. From Plato to Hume to the contemporary compatibilist debate, many philosophers and scholars have investigated the nature of this concept. Often defined as a weakness of the will, the philosophical study of addiction relates to the realm of freedom of the will. How does the configuration of the will look like for an addict subject? Is an addict's will free? To answer these questions, the perspective I adopt in my thesis is based on the compatibilist free-will debate. By adopting a compatibilist perspective, I analyze the philosophical discourse on addiction and build on the existing research on the subject, to establish my own definition of the concept. Moreover, this theoretical point of view can answer potential determinist objections, as I will show in the last section of the chapter. The framing of addiction developed in this chapter will serve as a theoretical framework throughout this thesis.

Section 1.2 serves as an introduction to the earlier conceptions of freedom of choice and freedom of the will as presented by Plato and Hume. The Platonic concept of *akrasia* – namely lack of self-control – and the Humean definition of freedom of action introduce concepts such as control, will, and action, which will be central to the philosophical definition of addiction. Section 1.3 focuses on the compatibilist perspective on free will. My research centers on the works of compatibilist philosophers Gary Watson and R. Jay Wallace. The compatibilist perspective on free will affirms that freedom and responsibility can coexist in a deterministic world (Watson, 2003, p. 338). This thesis considers how and whether a subject's will is impaired the moment they suffer from addiction. Classic compatibilism affirms that freedom consists of the absence not only of external constraints on the subject, but also of internal constraints that could impede the subject's will (Pickard, 2019, p. 1). The main takeaway from this section is that addiction consists of a lack of alignment between a subject's valuation and motivational system. In section 1.4 I argue that the compatibilist framework is lacking, as it does not consider the social side of addiction. For this reason, I turn to the work of Hanna Pickard, who focuses on the influence of self and social characterization on addiction. This allows me to argue for a more holistic perspective on the philosophy of addiction related to the free-will debate. Finally, section 1.5 considers the technological aspect of addiction when it comes to Social Network Sites. Here I show insights from mediation theory and philosophy of technology can add to the definition outlined in the chapter, to contribute to a more holistic definition.

1.1 From Plato to Hume: freedom of action and the choice of the Good

The Protagoras is a Platonic dialogue where Plato depicts a dialogue between Socrates and the main sophists of the time (Plato, 2009, p. ix). The dialogue revolves around a fundamental question: How is it possible for somebody to achieve the ability to live the best possible life? In the dialogue, Socrates discusses the nature of desires with his interlocutors. Every agent is motivated to aspire to what is best for them, where best refers to the moral Good. Morality, in this case, is a necessary aspect of fulfilling one's nature (Plato, 2009, p. xi). It is the rational part of the soul that leads the agent towards what is Good for them, to achieve the good life. The Platonic depiction of desires that allow a subject to pursue the good is based on the hedonistic theory: good is what provides pleasure, bad is what causes pain. Following the Platonic line of reasoning, it is impossible to do something that somebody knows is wrong (Morris, 2006, p. 197):

> "No one freely goes for bad things or things the believes to be bad; it's not, it seems to me, in human nature to be prepared to go for what you think to be bad in preference to what is good."
>
> (*Protagoras*, 358d)

It is interesting to notice how Plato chooses to talk about freedom: nobody *freely* chooses to go for bad things, as if they chose to, they would not be free agents. This impossibility is tied to the concept of *akrasia*. From the Greek *a-* (without) and *-kratos* (power, force), akrasia is roughly defined as a lack of self-control. When an agent knows what the best course of action would be but nevertheless persists in choosing another course of action, that is defined as an *akratic* action. The source of motivation for action, in a Platonic framework, comes from the rational part of the soul (Watson, 2003, p. 339). The desires of the Reason that move an agent to action are intrinsically and by definition desires of the Good. At first glance, akratic actions have no space in a Platonic framework, as every time an agent knows what the better option is, that would be the one they would choose, unless they were coerced into doing otherwise. However, even in the case their action was coerced in practice, their will would still be free in rationally choosing the best option. The only case in which an agent would not *freely* choose the best option, would be when they were not aware of the fact that that was not the best option to pursue. The pursuit of the Good is in fact the ultimate goal of the wise person. An agent which follows the path of action that will lead to the Good is a free agent. In Platonic terms, freedom is then understood as self-mastery. The desires of the Reason (the rational part of the soul) are the desires of the Good, and once an agent is able to follow those, then that agent is free (Watson, 2003, p. 341). The agent who follows those desires that do not lead to the Good has their

will obscured, believing that they are following the Good when they are not. In Plato's perspective, this agent is not free.

In the same dialogue, Socrates also ponders the case of those people who surrender to pleasures which do not lead to the Good:

> Don't you maintain that it happens that in some circumstances, often for instance when you are conquered by the pleasures of food and drink and sex, you do things though you know them to be wrong? […]
>
> Do you suppose, Protagoras, that they [those who do these things] would give any other answer other than that they are bad not because they produce immediate pleasure, but because of what comes later, diseases and the like?
>
> (*Protagoras*, 354c-e)

This scenario could fit in the concept of *akratic* action. However, in the context of the Protagoras, Plato describes how it is impossible to have contradictory preferences. The agent who chooses to go for the worst option simply chooses to do so since they focus their attention on the immediate pleasure. However, as Michael Morris underlines, it would be extremely easy for an agent to switch their preferences and change their mind suddenly (Morris, 2006, p. 226). Overall, the account of *akrasia* presented in the Protagoras does not provide a full and complete understanding of freedom of the will.

Two thousand years later, the question regarding freedom of the will is still being investigated. Scottish philosopher David Hume explores the question of freedom of agency in the *Treatise of Human Nature* (Hume, 2003). The framework that Hume presents is more simplistic compared to the Platonic one and still leaves some space for questions regarding the free will debate. In a Humean perspective, unlike Plato, Reason is not a source of motivation for the agent (Watson, 2003, p. 339). An agent is free the moment they have the so-called "freedom to do otherwise". Freedom is framed as the absence of external constraints on the agent. The moment the subject had the possibility to do otherwise, and they could have chosen to do so, then they were considered free. Classic compatibilism resonates with this line of thought. The compatibilist framework aims at offering a solution to the free will problem when it comes to determinism. The general thesis of compatibilism argues that determinism is compatible with the ability to do otherwise, as an agent's action is taken into account as a conditional statement. While determinism argues that a certain action would have been unavoidable in the context of a certain past, classic compatibilism considers all the possibilities of

that course of action in the context of all different pasts. I am restricted by the scope of this thesis to discuss free will from a compatibilist perspective. Together with other classical compatibilists, Hume claims that the absence of external constraints is enough for an agent to be free the moment they act.

One of the main issues of this perspective is that the Humean conception gives an account of free action but not of free will. The common definition of freedom, stating that an agent is free to the point that they are able to do what they want, follows the Humean argumentation. It is true that the absence of external constraints guarantees the absence of external coercion. It does not, however, provide an account concerning the possible coercion of the will of the subject. Plato's conception of freedom, even if developed thousands of years before that, provides a more complete – even though not exhaustive – account of freedom of the will compared to the Humean perspective. In Plato's view, the Reason provides a source of motivation for the subject to act (Watson, 1975a, p. 339). In Hume's case, the Reason could be considered an evaluative tool to choose among different action options, but it represents in no way a motivational means.

1.2 Contemporary compatibilism: valuation system and addiction

The Platonic and Humean perspectives presented in the previous section provide an introduction to the free-will debate. Although setting the basis for the compatibilist debate, Hume's account of freedom of action does not account for the problem of free will that I aim to discuss in this paragraph.

Gary Watson also situates himself in the compatibilist debate. The author does not consider the truth or falsity of determinism when it comes to distinguishing between free and unfree action (Watson, 2003, p. 338). In this sense, he adopts a Platonic perspective on the free will debate but brings the Platonic dichotomy even further. Reason, in Platonic terms, serves as a catalyst for human action. At the same time, the rational part of the soul is in charge of regulating what has value and what does not, what follows the idea of the Good, so that the subject can act upon it. This distinction between *valuing* and *desiring* is what Watson brings into his theory (Watson, 2003, p. 341).

Stating that a subject is free to do something once they have done what they most *wanted* to do can be interpreted both in the sense of what they most *desire*, or what they most *value*. Here Watson's theory differs from the Platonic one, as there is no guiding Reason that always aligns with the Good. Those case scenarios where desires and passions generate certain desires that do not align with what the subject *wanted* to do, fit into the category of actions independent from the subject's judgement of the Good (Watson, 2003, p. 342). Watson rightfully introduces a fundamental and complex discrepancy between evaluation and motivation. Cases of actions where appetites – which are not linked to one's conception of the Good – *motivate* the subject to act independently of their

evaluation of a certain action constitute a problem of free action. To make a trivial example, a person could decide not to eat chocolate for a week because they deem this bad for their health. However, their desire to eat chocolate, driven by their appetite, could override their previous evaluation and make them end up eating chocolate. This scenario exemplifies the Watsonian distinction between *valuation system* and *motivational system.* To put it in Watson's term, an agent's valuation system is that series of considerations that allow them to formulate judgements such as "The thing for me to do in these circumstances, all things considered, is *a*" (Watson, 2003, p. 346). A subject who is can to make these kinds of judgement is considered to have free agency. In addition to the valuation system, another type of agency system moves the agent to action: the motivational system. The motivational system can be obstructed by desires driven by passions and appetites and lead the subject to act in such a way that overrides their own evaluation. This opens up the possibility for unfree action: when the two systems do not completely coincide, an agent's ability to act on their valuation judgements can be hindered by their own will.

Watson's framework can be traced back to Harry Frankfurt's conception of desires (Frankfurt, 2003). An unfree action is the one performed by what Watson calls the *unwilling addict*: when first order desires (the Watsonian motivational system) don't match second order desires (valuation system), the subject performing that action is comparable to an unwilling addict, meaning that they would not want to take the drug, but they still do. However, this definition is still too broad to constitute a proper definition of addiction. For instance, to make a similar example to the one mentioned in the previous paragraph, one could think of a case where somebody promised themselves not to eat pizza for a week (valuation system). Still, once they find themselves at a party with friends, they end up eating a slice of pizza that was ordered for everyone (motivational system). Should that person be considered an addict, as their values did not match their motivation in that scenario? Intuitively, this situation does not fit into the stereotyped conception of addiction.

Up until this point, Watson's account of agency systems distinguishes between valuation and motivational system. The lack of alignment between the two, however, does not distinguish between moral weakness and addiction. Suppose one can act upon their motivational system while ignoring their valuation system and that action is still not considered to be addictive. What are then the peculiar aspects that specifically characterize addiction? The author expands on this concept in his work *Disordered Appetites: Addiction, Compulsion, and Dependence* (Watson, 2010).

If one were to consider the vocabulary that is generally adopted when talking about addiction, this often includes terms such as "out of control", or "compulsive desire". Both lack of control and compulsion are ambiguous terms in this domain because they allow for discharge from responsibility. Compulsion, from Latin *cum-pellere* (go with the drive, with the force), affects an agent's

motivational system, in the sense that this last one overpowers the valuation system, which is responsible for one's account of the Good. But how far can one's desires overpower their evaluations? Can we talk about "lack of control" or "irresistible desires"? Strictly speaking, no desire is "irresistible", in the sense that, taking a close look at the term, in every case of irresistibility, if the agent had tried harder, they would have succeeded in avoiding that desire (Watson, 2010, p. 5). This is also confirmed by the fact that people are actually able to recover from addictions and moral weaknesses. If their desires were in fact irresistible, they could have never resisted them in the first place in order to start a recovery process. Watson refers to them as "recalcitrant desires" because they are very strong and hard to resist, but not impossible. With an incisive metaphor, the author describes addictive motivational obstacles as the forces that lead the exhausted climber to abandon their intent to reach the top of the mountain (Watson, 2010, p. 6). The lure of nicotine or alcohol does not work by brute force, like gravity would, or a big rock in the climber's path that physically impedes their climb. Going back to the etymological origin of compulsion, motivational obstacles invite the subject to go with the drive, in the sense of diverting them from their effective resistance (Watson, 2010, p. 7). This is why, when one gives up their climb or smokes the next cigarette, they tend to perceive a sense of shame: it is a sign of yielding to the appeal of the desire, giving up the better resolution of following the Good that the valuation system had set.

Together with the appealing character, another characteristic of seducing desires experienced as compulsive is their ability to capture someone's attention (Watson, 2010, p. 10). I have already discussed how being susceptible to counterincentives proves the non-irresistibility of compulsive desires. They possess, however, the capacity to break one's concentration while performing other activities. Sticking with the climber's example, the desire to quit the climb interrupts the undivided focus on the activity, firstly like a background noise that becomes more and more invasive over time. Smokers experience the same disposition towards smoking-related stimuli, resulting in being more prone to nicotine-related cues (Vollstädt-Klein et al., 2011, p. 223). The subject tends to apply techniques of mindfulness or self-control, in order to impede the motivational system to take over the valuation one and to refrain from such activities. This creates a contrast in the subject's systems, leading to what Watson defines as "fragmentation of consciousness" (Watson, 2010, p. 11). The amount of effort that the agent puts into trying to re-align the motivation and the evaluation already dominates the agent's attention, causing them to break their concentration from the main task. However, this aspect of compulsive desires still doesn't account for a necessary and sufficient condition for addiction. Fragmentation of consciousness does not determine complete incapacity in the subject. There are many other activities that interrupt our course of attention during the day, like the thought of a loved one, parents thinking about their children while they are at work, etc. (Watson,

2010, p. 18). However, these scenarios do not fall into the category of addiction, even though they comply with all the previously mentioned. conditions. To provide a more fitting account of addiction, the author suggests the idea of *acquired appetite* (Watson, 2003, p. 12). To become addicted is to become vulnerable to the temptation of a nonnatural appetite. This definition brings about many aspects that constitute the "nonnatural" part of the explanation, such as identity, culture, existential dependency. I will focus on them in the last section of the chapter.

Before turning to how addictive patterns influence one's identity and place in society, I wanted to analyze further conditions that help frame addiction, which R. Jay Wallace provides in his article *Addiction as a Defect of the Will* (Wallace, 1999). According to Wallace, one typical condition that characterizes addiction is automatism: routines that develop through habituation facilitate the possibility of one's valuation system to clash with their motivational system. On this regard, Wallace makes an important disclaimer: automatism due to habituation does not constitute an impairment to the volitional capacities of the subject (Wallace, 1999, p. 626). The fact that somebody is used to repeating a certain action every day doesn't necessarily imply that this person has lost control over that action. Daily activities such as eating, drinking or sleeping provide an example of this. Being a parent or being in a relationship also fit this characterization and all the previous ones, but still don't fall under the category of addictions.

To better frame what addiction is, Wallace introduces a normative element. Addictive behaviors are characterized by the fact that they are generally frowned upon. However, this does not necessarily link to the will or control an agent has upon their actions. This aspect will be further inquired in the next section. As long as the freedom of will is concerned, addictive behaviors typically influence the deliberative reflection of the agent, as they are not able to stop an action that the same agent would not want to perform (Wallace, 1999, p. 628). Again, this kind of definition does not seem to distinguish bodily appetites from addictive patterns properly. Bodily and addictive desires are both resilient in the sense that they are able to overcome the agent's deliberative reflection. For example, when a person has gone many days without eating, they are probably no longer in control of their actions when it comes to what they're going to eat, how they're going to obtain food, etc. This, however, does not make that person "addicted" to eating. This is because the agent's will matches their action. Eating was their aim in the first place. Therefore, there is no clash between their action and their will. However, the way and what they decide to eat can determine a clash between the two systems. This again can be challenged by cases of ordinary weakness: if a desire driven by someone's motivational system wins out over what their valuation system had decided should have been done, that indicates a defect of the will. When is a defect of the will considered to be addiction?

1.3 Identity and society: from the sense of self to social identity

So far, I have analyzed numerous aspects that generally characterize addictive behaviors. They are in most cases due to a lack of alignment between one's valuation and motivational system, they seduce by appeal, they tend to become automatic actions in the agent's life, and they could be defined as acquired appetites. However, as shown in the previous paragraph, framing addiction only from a freedom of the will perspective does not provide a full and satisfactory account of it. This is because actions characterized by a clash between one's motivational and volitional system are not necessarily considered addiction. For instance, if I promise myself not to drink Coca Cola for a week, and in that same week I go out with my friends and order one, my motivational system clashes with my volitional system. At that very moment, I am an akratic subject. However, that *akrasia* is generally not considered to be addiction. Why is that? Both Watson and Wallace provide further conditions that focus on a more normative-societal perspective.

One aspect that could help better frame Wallace's addiction is how addictive impulses tend to be frowned upon by society. However, the author quickly dismisses this aspect:

> This brings out a normative element in ordinary thinking about addiction. We label an impulse addictive only if its satisfaction is something that we tend to disapprove of – as being, for instance, difficult to reconcile with a worthwhile, dignified human life […]. This point about common classification, however, is not of much philosophical depth.
>
> (Wallace, 1999, p. 627)

The frowning upon the action steers the conversation towards a normative framework. Being situated in the free-will compatibilist debate, Wallace tends to shy away from societal influences of addiction. The focus of the compatibilist debate is, in fact, on the effects of addiction on freedom of the will. In this section and the following one I will show how both the work of Hannah Pickard, together with the one of Peter-Paul Verbeek can add to the compatibilist definition of addiction, contributing to a more holistic view of the topic from a philosophical perspective.

Both the philosophical discourse and popular cultural and societal contexts agree on the framing of the addict as somebody subject to irresistible desire (Pickard, 2019, p. 455). Not only are they framed as irresistible, but the action steered by these desires takes place even when the subject is aware of its detrimental consequences. These detrimental consequences can take many forms. In her work *The Sources of Normativity*, Christine Korsgaard investigates the sources of morality.

Ethical standards that lead society are normative, meaning they make claim on subjects (Korsgaard, 1996, p. 8). From this, individuals make claims on one another, and therefore institute standards and commands that ought to be followed. When society determines that an action is good, people are expected to perform it. On the same note, society does not disapprove of an action because it is bad. On the contrary, a certain action is in principle defined as bad because society disapproves of it[1] (Korsgaard, 1996, p. 50). If we were to combine the overall criteria that define addiction that have been analyzed so far, the resulting definition would be along the lines of: "Addiction is an acquired appetite that takes over when a subject habitually performs an action where their motivational and volitional systems most likely don't match, which leads to detrimental consequences in their lives and is generally frowned upon by society." On an intuitive level, this definition could cover up the general conception of "the addict" that frames the discourse around addiction.

However, if one wanted to challenge this definition even further, applying it to practical cases can lead to interesting results. Hanna Pickard's work focuses on philosophical definitions of addiction. In her research, she merges philosophical conceptions of addiction together with actual data resulting from medical and societal cases with addicts. Citing a passage from William James' *Principles of Psychology*, Pickard perfectly encapsulates the stereotypical definition of "The Addict"(Pickard, 2019, p. 454).

> "The craving for a drink in real dipsomaniacs, or for opium or chloral in those subjugated, is of a strength of which normal persons can form no conception. "Were a keg of rum in one corner of a room and were a cannon constantly discharging balls between me and it, I could not refrain from passing before that cannon in order to get the rum"; "If a bottle of brandy stood at one hand and the pit of hell yawned at the other, and I were convinced that I should be pushed in as sure as I took one glass, I could not refrain": such statements abound in dipsomaniacs' mouths."

People suffering from addiction are considered to be powerless, in the sense that their cravings overpower their own will, and therefore make them surrender in those cases when they aim at refraining from performing a certain activity. Addictive impulses are thus defined as irresistible,

[1] Korsgaard's view is further developed in her work *The Sources of Normativity*. Her final view does not claim that an action is badly judged because in principle society disapproves of it. To see her argument on the topic, see *The Sources of Normativity.*

meaning that, even when wanted to be resisted, addicts are unable to do so. However, as I have explained in the first paragraph of this chapter, no desire is actually irresistible. Advocates of this compatibilist framework support the view of recalcitrant and strong desires. It is important to underline that hard to resist does not automatically translate into irresistible, and that is also confirmed by methods such as contingency management (Pickard, 2019, p. 6). Addicts who are offered to take part in contingency management treatment have to provide three urine samples per week in exchange for 100$. Following the "irresistible desires" line of thought, contingency management should fail, as addicts are, by definition, unable to abstain from the addictive activity. What the experiment shows, however, is that CM treatment is extremely successful, even compared to other kinds of treatment – such as cognitive-behavioral therapy – which are not based on reward.

When linking back this practical definition of addiction to the philosophical account described in the previous paragraphs, the scenario doesn't add up. Addiction, it has been shown, cannot be simply characterized by the way someone's motivational and valuation systems are connected. Instead, the lack of control that the addict – not always – experiences needs to be relativized to the circumstances of the action (Pickard, 2019, p. 460). Contingency management treatment is only one practical example of how addicts in fact retain the capacity to exert control over their addictive behavior. It is not the case they are unable to exert control. It is more the case that they *decide not to* exert control.

This new scenario appears to contradict the Platonic concept. If an addict knows that the course of actions they are following will lead to detrimental consequences in their lives and that there is a better option – sobriety – to this course of action, what is stopping them from following it?

Both Watson and Pickard provide further conditions to bridge this gap. Gary Watson talks about *existential dependence*, meaning that the addict develops a series of practices that become crucial to the development of their own identity (Watson, 2003, p. 16). In order to break an addiction, the addict does not only need to re-align their valuation and motivational system, break a habit and unlearn an acquired appetite, but they also and most importantly need to get a new sense of their life and identity.

On the same line, Hanna Pickard incorporates Self-Categorization Theory and Social Identity Theory in the study of drug addiction (Pickard, 2020a; Tajfel, 1982; Turner, 1989). In her study, she portrays how addicts tend to self-categorize themselves as such, and are also identified as "addicts" by society. These two categorizations together contribute to forging their own identity, fostering their addictive habits. To a certain extent, addicts are addicts because they do not know what else they could be. In her book *Addiction by Design*, Natasha Dow Schüll analyzes the issue of problem gambling. The testimonies she present of ex-gamblers perfectly depict the way the identity of

"problem gambler" fits the description that gamblers give of themselves. The activity of the gambler can create a structure and a sense of purpose in their lives, which strongly defines their daily routine. Julie, an ex-gambler interviewed by Dow Schüll, perfectly describes the shift from non-addict to addict (Schüll, 2014, p. 204):

> "When the time comes to leave and the things I escaped from start crowding back into my brain, I find myself rationalizing, *Well, I don't really have to go today…* and I ask an attendant to hold my machine while I run to the payphone to call and buy myself more time, and then back to continue […] I'm thinking of how to arrange things so that I can stay there, *how to economize."*

The acquired addictive appetite is so much engrained in the addict life that their own identity revolves around it, making it challenging to even think about an alternative.

Pickard considers the reasons why addicts do not or cannot have an alternative social identity and provides three possible alternatives (Pickard, 2020a, p. 14). It could be the case that their addictive patterns started so early in life that they did not have an option to develop any kind of alternative identity. Alternatively, it could be the case that what constituted the main aspect of their previous identity no longer exists, so they are denied the possibility of going back to their "old self". A third alternative would be those cases where the person has been living like an addict for so long that going back to their past self represents an internal barrier so great that it cannot be overcome. The amount of work required to become someone else is sometimes too great of a challenge.

So far, this chapter has taken into account a definition of addiction from a compatibilist perspective. The final definition would consist in: "Addiction is an acquired appetite that takes over when a subject habitually performs an action where their motivational and volitional systems most likely don't match, which leads to detrimental consequences in their lives and is generally frowned upon by society. This course of action strongly shapes the identity of the individual who ends up identifying with that character."

This final definition I provide leaves an open possibility for further investigation and clarification. However, for the scope of this thesis, I will be using this theoretical tool to analyze whether SNS use can be considered addictive in nowadays society.

1.5 Mediated addiction: what happens on social networks doesn't stay on social networks

In the previous sections, I have outlined a philosophical definition of addiction. When it comes to platforms such as Social Network Sites, what is the role of technology in the addiction discourse? The current debate on persuasive and addictive technologies builds on medical and psychological insights, as the work of psychologists such as Mark D. Griffiths and Dr. Daria Kuss shows (Griffiths et al., 2014, 2016; Kuss & Griffiths, 2011). This section presents a brief introduction to mediation theory, to show how a technology such as SNSs mediates addictive behaviors.

In his paper *Beyond Interaction: a Short Introduction to Mediation Theory*, philosopher Peter-Paul Verbeek briefly outlines his perspective on the relation between humans and technologies (P. P. C. C. Verbeek, 2015). The way psychological research describes the relation between addict patients and technologies is framed as an interaction. The individual interacts with the technology, and the addictive behavior is narrated in terms of human subject that becomes enslaved to the technological object. Verbeek's theory goes beyond the two poles of the interaction, claiming that humans and technologies are not only part of this interaction but the result of it (P. P. C. C. Verbeek, 2015). As technologies and humans mutually shape each other, the addictiveness of the technology emerges as a way human agents live in the world. SNS platforms act as a medium between the user and the cyber world they are experiencing, which in turn cannot be confined to the experience on the platform itself. This is because, firstly, the platform is not merely a platform. The cyberspace where these interactions occur is a result of multiple interconnected technologies – screens, phones, laptops, keyboards, etc. (Gertz et al., 2019, p. 74). Secondly, as philosopher Nolen Gertz points out, what happens on SNSs doesn't stay on SNSs:

> We may log out of our apps and our devices, but our apps and our devices do not log out of us. This is why we must not try to flee from our technologies or try to somehow get outside of technological mediation, as the belief that such escape is possible merely reinforces the illusion that technologies only influence us so long as we are using them.

(Gertz, 2018, p. 209)

The same way a gambler's experience does not end the moment they step out of the casino, the experience of Social Network Sites does not end the moment users close an app. Patsy, a problem gambler interviewed by Natasha Dow Schüll, described her whole life as revolving around the gambling machine: her life ended up being "a machine life" (Schüll, 2014, p. 189). Schüll renders

this testimony by stating that, when Patsy became addicted to machine gambling, another *mode of life* emerged. One could say that what happens in Vegas doesn't stay in Vegas, but shapes the life a subject conducts even outside of the addictive activity itself. In the same way, what happens on SNSs does not stay on SNSs. Pickard's research frames this mediation from a societal perspective. The way the addict interacts with the substance – in this specific case, Social Network Sites – shapes the way they perceive themselves and others perceive them. Humans are intrinsically technological beings, as their existence is permanently mediated by technologies, the same way the subject of addiction is permanently shaped by their addictive behavior. Verbeek claims that we cannot be human without technologies (P. P. C. C. Verbeek, 2015). Can the technologies that make us human not be addictive? In these dynamics, technology designers play a fundamental role, as they do not merely design technological products, but through their work they shape the human experience of users. The result of their work translates into modes of being of those who use their product. Mediation theory shines a light on the role of technology design when it comes to shaping users' behaviors.

This raises the question: *Are SNSs designed for addiction?* The next chapter will apply this theoretical scenario to the relationship users have with SNS, to compare the characteristics of SNS use to those of addictive behaviors.

2. Are Social Network Sites designed for addiction?

This chapter will focus on the way users approach Social Networking Sites, applying the theoretical framework developed in the previous chapter. By merging my philosophical investigation with an analysis of SNS design features, I aim to answer: *Are SNSs designed for addiction?* My analysis will focus on the role of design, demonstrating that the addiction-generating features of SNSs are intentionally included in the design of the product. This design choice is dictated by the monetization of the content in the platforms, which has a direct correlation with the increase of users' engagement. The definition of addiction outlined in chapter one states:

> Addiction is an acquired appetite that takes over when a subject habitually performs an action where their motivational and volitional systems most likely don't match, which leads to detrimental consequences in their lives and is generally frowned upon by society. This course of action strongly shapes the identity of the individual who ends up identifying with that character.

Section 2.1 lays out the precise taxonomy of SNS platforms, describing how they work and what distinguishes them from digital and social media. This includes a short introduction of the history and design of the Instagram app, that I consider as a main example. Section 2.2 explains hos SNSs work. In section 2.3, I show how SNS mode of operation meets the criteria of the definition of addiction outlined in Chapter One. Here I touch upon three main features that characterize addiction and show how they are implemented in SNSs: habitual performing, detrimental consequences on an agent's well-being, and social identity and self-categorization.

This chapter, which aims at showing how addiction-generating design patterns are deliberately implemented in SNSs, will serve as a basis for Chapter Three. There, I investigate what specific consequences this design pattern determines in the potential regulation of addiction by design.

2.1 The taxonomy of social media: a working definition of Social Network Sites

Before dwelling deeper into the definition of Social Network Sites, I shall briefly define what digital media is, where social media stands in this category and how social network sites represent a smaller division of it.

As Stacy O'Neal defines it in her book *Digital media: Human-Technology connection*, digital media is the name of content and devices applied to the digital domain (O'Neal Irwin, 2016, p. 17). This broad term encompasses a set of digital components, together with the content they provide.

Users experience the content on digital devices, and the combination of the two parts constitutes digital media. Social media is a sub-category of digital media, and it indicates the application of digital media to the social sphere. Social media are characterized by some common distinctive features: they are *convergent*,[2] in the sense that they allow the use of multiple types of content on one single device (Biscaldi & Matera, 2019, p. 28). For instance, a mobile phone can be used to make phone calls, check emails, scroll through one's Instagram feed, ask Siri to set a timer, etc. They are *hypertextual*, as their content refers to other textual – or visual – content. They are, finally, *interactive*, as users themselves can decide what content to use or to produce. This last characteristic leads to the *social* aspect of it and differentiates them from digital media. Social media presuppose a multi-directional communication based on the interaction between users. Because of this last feature, unlike digital media, the term applies mainly to the content and not so much to the device.

The current definition of social media allows to include a wide range of social applications, such as blogs, virtual game worlds, Social Network Sites, etc. (Kuss & Griffiths, 2017, p. 2). For instance, collaborative projects – namely social media applications that allow for the collaborative creation of knowledge-related content by users – fall into this category (Kaplan & Haenlein, 2014). The most prominent example of this is Wikipedia. Being a collaborative online encyclopedia, Wikipedia allows its users to interact in the creation of entries. Another example of social media that does not fall into the SNS category are forums. Also included in collaborative projects, forums allow for social interaction in the form of posted messages on a web page (Kaplan & Haenlein, 2014). When referring to social media, however, the first thing that comes to mind are online platforms such as Facebook, Instagram, TikTok, ect. What makes them stand out of the social media category?

Social Network Sites (SNSs) are a type of social media platform. They are characterized by the possibility of creating a personal profile that allows the person to interact with other users. Biscaldi and Matera identify three main affordances that SNS make possible.

Firstly, they allow for *relational*[3] *use*. This is the main use that SNSs and platforms tend to advertise. Unlike collaborative projects, forums and blogs, whose main goal is that of producing knowledge digitally accessible to a larger public, Social Network Sites' main aim is the one of profiling users and portray their digital self on the platform. Of course, knowledge is still produced and shared on platforms such as Facebook, Instagram and TikTok. However, one of their central goals is to connect people to each other. On an SNS profile, the user has links or "friends" that can check on the content they upload. On this note, Instagram's main page recites "We bring you closer to the people and things you love"(*About Instagram's Official Site*, n.d.). The relational use is more

[2] My translation
[3] Ibid.

specifically what makes Social Networking Sites *social*, in the sense that they promise the user the opportunity to broaden their social circle. Again, this differentiates them from the main aim of forums and blogs. However, Boyd and Ellison cleverly point out how the "network-*ing*" part of SNS use merely represents a marginal aspect of it (Boyd & Ellison, 2007, p. 211). Networking generally refers to the creation of new social connection, establishing the beginning of a new relationship between strangers. Nevertheless, studies have shown how this is not the dominant process users tend to go through (Boers et al., 2019; Damico & Krutka, 2018; Rozgonjuk et al., 2020). The most common practice is to connect with either people they already know, the so-called "latent ties" (Haythornthwaite, 2005, p. 13). Networking, therefore, should rather be replaced by "network", as users do not engage in the activity of networking on the platform itself, but instead cultivate the social relationship they have already established in their life. This aspect distinguishes SNSs both from collaborative projects and from, for example, multi-player or co-op videogames. While both form part of social media, their main goal is not that of strengthening social latent ties. In a forum or in a multi-player video game, the medium allows for and encourages the social communication aspect. However, the main goal for which a user would decide to write on a blog or play a multi-player game is not the one of reinforcing pre-existing social connections. Therefore, the "network" part is exclusive to SNSs platforms.

The second aspect that characterizes SNS platforms is *expressive use*. The Instagram main page recites "Our teams inspire creativity around the world, helping over 1 billion people create and share"(*About Instagram's Official Site*, n.d.). This aspect of SNS platforms makes it possible for users to upload personalized content, which contributes to the creation of their online social profile. Users decide which parts of their liver or of themselves they want to share, constructing their own online identity (Mun & Kim, 2021). This aspect once again differentiates SNSs from other social media platforms. Forums, blogs, and collaborative projects allow for the creation of a personal profile. However, that profile is functional to the publication of content on the web page, but it is not meant to express the online identity of the user who owns it. It allows for the expression of their opinion or knowledge. Nevertheless, the visual expression of the user's identity does not constitute the central aspect of these platforms. Therefore, the – visual – expressive use of one's profile is limited to SNSs.

Lastly, the third aspect that Biscaldi and Matera identify is the *explorative use*.[4] This last characteristic refers to the possibility that SNS offers to analyze other users' profiles and identities on the platform. Instagram, not coincidentally, has a feature called "explore page", where users are suggested photos and videos from profiles of people they know or whose content they enjoy. The explorative use can be associated with forums, blogs, and collaborative projects. As their main goal

[4] My translation

is to produce and share knowledge, it is possible for users to explore the entries and pages. However, SNSs differentiate themselves, as they represent the merge of the two previously cited aspects – relational and expressive. On SNS platforms, users explore not only knowledge-related content but also and most prominently profiles of users whom they want to be in contact with, where their online identity is displayed. Therefore, even though they share the explorative use with other types of social media, SNSs apply it in a specific way that differentiates them from other social media.

2.2 How do SNSs work?

The two main definitions of SNSs are provided by Kuss & Griffiths and Boyd and Ellison. They describe SNSs as virtual communities where users can interact with people based on their shared interests, share connections and view and explore these connections based on these interests (Boyd & Ellison, 2007; Kuss & Griffiths, 2011). For instance, on Instagram, users can share pictures, post stories sharing their location, chat with other users, and foster pre-existing connections. Despite the fact that many SNSs main pages advertise their product on a qualitative basis, the value of each user's profile tends to be largely measured by quantitative tools (Coxon, 2017, p. 78). The number of followers, followed people, likes, and posts are all very visible and present in the user's main profile page. Reversely, the way these platforms work, even when being advertised as focusing on the quality of the content provided, is very much based on quantitative methods. Engagement rate and screen time are the two main factors that make a user or a designer successful in the SNSs world (Lanier, 2018, p. 12). Social network sites and platforms work based on feedback reward: once a user has liked or shared a post that contains a certain type of content, the algorithm will suggest similar content. This mechanism is called positive feedback, where the user receives similar content to the one they have already engaged with, in order to foster their own preferences. There is, however, another trick up the sleeve of social media algorithm programmers: negative feedback mechanisms. For instance, if one's feed is constituted by cat videos and landscape pictures, the algorithm would sometimes provide a little incongruity, like for example a picture of a strawberry cake. That person has probably never even looked at dessert pictures, and maybe does not even like strawberries. However, this is the exact way algorithms work: they are adaptive. They do not always provide the user with exactly what they would expect. They tend to add a touch of randomness here and there, to see whether the user is going to engage more with the new type of content (Lanier, 2018, p. 13). If they don't like strawberry cakes and don't click on the picture, after a while they won't see strawberry cakes picture on their feed anymore. But if they click on it out of curiosity, a new pattern will start because that specific output kept them on the app for five seconds longer than usual. The attention of the user was captured for a fraction of time above average, which is what makes the algorithm "better", in profit

terms. This process is also called optimization. It is a never-ending mechanism that allows the algorithm to better trace the user's preference and provide a "better" experience on the platform.

The algorithm adapts to the preferences of the user in the sense that was described above. Another mechanism that Lanier describes is the so-called leaping mechanism. It is the same system that characterizes human evolution and, not surprisingly, the way the human brain works. In his book *The Shallows*, Nicholas Carr dedicates a lot of space to the description of neuroplasticity (Carr, 2011, p. 38). From a scientific point of view, the way the human brain is wired is not purely determined by one's genes. Its functionality is not completely determined by a series of prefabricated variations but is able to adapt to different paths depending on the way someone acts and thinks. Once the brain has been exposed to a new activity – or to a new technology –, its neurocircuits start developing new maps that are then going to trace the path that needs to be followed to carry out that specific activity. The brain is "plastic", in the sense that it can be molded to the path of the new activity that is learned. This characteristic is known as neuroplasticity. However, for how useful neuroplasticity is, this moldability carries a risk. As Norman Doidge writes in *The Brain that changes itself*, any action that involves unvaried repetition is at the risk of getting stuck in a rigid pattern (Doidge, 2007, p. 170). It may appear as a paradox, as neuroplasticity suggests the idea of a more fluid system that is easily able to adapt itself. However, when a new map area for a novel activity is developed, the brain aims at keeping it working. Doidge describes it as a "use-it-or-lose-it" brain: the new patterns are either exploited as much as possible or abandoned. This is on one side extremely effective because it allows humans to easily adjust to new situations and conditions. On the other side, paradoxically, once a new pattern is set into action, it is quite hard to remove. Hence the reason why neuroplasticity is easily linkable to addictive patterns. Even a small modification in the quantity or quality of the drug – or in this case, a small modification in the algorithm – can lead to long-term modifications in the way the brain reacts to this substance or activity. Lanier's leaping mechanism mentioned earlier works by exploiting this specific aspect. Adaptive algorithms, together with following the preference patterns of the user, insert a bit of randomness here and there. In the cases when the user picks up on that randomness, the map that had previously been built based on that behavior strengthens. While this mechanism is generally used to pick up on details out of patterns in real life, such as a car suddenly jumping out of a narrow road or a sudden loud noise, in the case of algorithm randomness the user is trying to build up a pattern based on a mere illusion (Lanier, 2018, p. 14). This behavior modification led by randomized patterns shows how the user's preferences and well-being are only marginally taken into consideration – at least as long as they can lead to "better" performance, meaning more profitable.

In the next paragraph I will analyze how, together with positive and negative reinforcement, SNSs platforms are designed to generate addictive patterns. This generates behavioral modifications, which determines and impairment of users' free will.

2.3 Addiction by Design

In her book *Addiction by Design*, anthropologist Natasha Dow Schüll describes the history of machine gambling in Las Vegas. Her analysis focuses on whether the addictive behavior stems from the gambler, the machine, or the interplay between them. She depicts how machine gambling algorithms are tailored to exploit the users' weaknesses to generate the most profit possible (Schüll, 2014). Gambling machines work thanks to mathematical algorithms installed on their chip, which executes the game's scoring scheme matched with a random number generator (RNG) leading to the game's final result (Schüll, 2014, p. 82). When the gambler pulls the lever or hits the button, the algorithm activates the RNG, which generates a result that is then translated into what the player sees on the screen. While the gambler has the illusion of controlling the process, they only work as a catalyst. The leaping mechanism used in SNSs platforms is the same that is used in machine gambling algorithms. The leap that Lanier describes is rendered by Dow Schüll with a majestic paraphrase, as she talks about "risk within a dependable framework" (Schüll, 2014, p. 13). The small risk determined by the novelty of the uncertain feature perfectly fits in into a pattern where the aim is not the one to win but to continue. In the same way, SNSs algorithms build on a pattern of expected content – to provide a sense of predictability and stability – and subtly introduce a sense of occasional disorientation. Other aspects that contribute to this apparent sense of stability are exemplified by Bhargava and Velasquez (Bhargava & Velasquez, 2020, p. 6).

One of the most prominent techniques is the intermittent variable reward, which is precisely called the "slot machine effect". Tristan Harris, ex design ethicist at Google and founder of the Center for Humane Technologies, mentioned this specific aspect at his hearing in front of the US Senate on June 2019. There he discussed the power of persuasive technologies (Harris, 2019). Two of the key aspects of intermittent variable rewards are positive and negative feedback, which contribute to the feeling of a machine gambling experience. Gamblers interviewed by Dow Scüll describe the gambling experience as "suspended automation", a place in time where they can just dissociate and trust an automated system (Schüll, 2014, p. 13). The same happens to SNSs users who, after they have been spending a considerable amount of time on SNSs platforms, just find themselves in a loop where they are not conscious anymore about what they are looking at on the page. Baym et al.'s research shows how users who deactivated their Facebook account for a while, later realized the approach they had towards the platform (Baym et al., 2020). Participant 23355 states:

> "Now I realize how much time I spend on it and how much time is not spent productively but rather wastefully. I catch myself mindlessly clicking on the icon on my phone even after I just got out of it."

And again, participant 11631 says:

> "I realize now how often I find myself endlessly scrolling and wasting time more than I had realized before the deactivation period."

They are not enjoying the experience anymore; they are just searching for cues to continue staying on the platform.

Another design pattern that contributes to this loss of spatial and temporal perception, which already appears in the previous testimony, is the infamous endless scrolling. Also mentioned both by Bhargava & Velasquez and Harris, an endless screen page that does not provide the possibility of hitting rock bottom cotributes to eroding natural stopping cues. Stopping cues, however, are necessary to spatially and temporally circumscribe the experience (Bhargava & Velasquez, 2020, p. 6; Harris, 2019). Through endless scrolling, the search for content becomes a boundless exploration for a result that has no end. Mark Kingwell describes this phenomenon as a "deferral of desire satisfaction combined with a substitution of its mechanism for the original desires" (Kingwell, 2017, p. 46). From a theoretical perspective, this statement can be analyzed in Watsonian terms. In the case the user had willingly decide to access the platform in search for content, the beginning aspiration of their motivational system – finding content on the SNS platform – was matching the one of their valuation system – wanting to find content on the SNS platform. However, thanks to the endless scrolling, the interface makes it easier for the motivational system content to switch from "finding content" to "search for content". As the search page has no rock bottom that can be hit, it just continues endlessly. It is no coincidence that the main explore page on Instagram, where users can browse for content, is called "feed": users are fed an unlimited amount of content, which does not offer the chance to satisfy their beginning need (Kingwell, 2017, p. 46). The deferral of the desire represents an interesting distinction between the SNSs experience and the drug/substance addict experience. In her work, Pickard is keen to stress the added value that drugs represent in the addict life. In simple terms, the narrator in Trainspotting renders it perfectly: "What they forget is the pleasure of it. Otherwise, we wouldn't do it. After all, we're not fucking stupid." (Boyle, 1996). Kingwell highlights how the analogy between the value of drugs and the value of SNSs scrolling is imperfect. The endless scrolling

seems to annihilate the possibility of achieving the initial will. With drugs, taking a pill will do the job. With SNS scrolling, the user may never even encounter the pill they were looking for. This difference, however, does not exclude either experience from the addiction category. Whether the final desire is achieved in practice or not, what matters is the unalignment between motivational and valuation system. In both cases, the will of the subject is impaired, as there is either no correspondence between the systems or the valuation system is completely bypassed and not considered. As highlighted in Chapter One, however, a lack of alignment between the two systems is not by itself sufficient to determine an addictive behavior.

This last aspect is linked to another characteristic of addictive patterns analyzed in the first chapter: habitual performing. The bottomless scrolling, together with the leaping mechanism and positive and negative feedback, contribute to render the SNSs experience a routine where the user is sucked in. The next paragraph will analyze how SNSs use has become a habitual pattern in the life of users.

2.3.1 Habitual performing

As Kuss and Griffith mention in their work, mobile phone use determines changes in users' life habits in their everyday lives (Kuss & Griffiths, 2017, p. 9). People carry around their smartphones wherever they go, to the point that a new type of phobia – nomophobia, namely no mobile phone phobia – is now being studied and discussed (Kaviani et al., 2020). Mobile phones, the most common medium through which SNSs are accessed, are always with us, and users struggle to detach from them. Nomophobia has been linked to impulsive mobile phone use, as users tend to automatically check their phones without a specific reason, simply because they have developed a habit out of it (*Do You Obsessively Check Your Smartphone?*, n.d.).

Habits, together with attitudinal changes, are brought about by the constant repetition of a certain action. If users check their phones more than 40 times a day, that becomes an automatic reaction that is then harder to eradicate. But not only are smartphones always with us, facilitating and favoring impulsive checking and habitual SNSs use. There are completely new routines and practices that are being molded around SNSs platform use. Insta tours are one of those (*Amsterdam Instagram Tour*, n.d.; *Bali*, n.d.; *Insta Tours – Shoot Amazing Instagrams*, n.d.; "InstaBanff," n.d.). Travel agencies and companies organize trips which are specifically tailored around the most instagrammable places in a certain destination. Travelers pay to visit specific locations and generally a photography-bundle is included in the service. The pictures that are taken during the trip will then be uploaded to their profiles, increasing the attractiveness of their feeds. Not only are SNSs platforms

part of our daily routines, but they also enter previously consolidated practices and mold them creating a new sense of need.

When relating these cases to the theory developed by Wallace, analyzed in Chapter One, it is important to remember that automatism brought about by habituation doesn't necessarily constitute and impairment to the volitional system. However, automatism does represent one of the main characteristics of addiction. When a certain action is automatic, it can be the case that the motivational system simply bypasses the judgement of the agent, in this case the valuation system. This is why sometimes users, while waiting for the train or even when chatting with a friend may find themselves with their phones in their hands without even realizing why they actually checked it (Hosie, 2017). Or when someone picks up their phone to check the time, look at it, put it away, and realize they have no idea what time it is. The impulsive action has simply bypassed the conscious intention, and they find themselves mindlessly looking at the screen for three more seconds. Maybe checking for notifications, maybe just following the path of habituation. Even though habituation doesn't necessarily lead to addictive patterns, it can still get in the way between motivational and valuation systems. The next paragraph will analyze this relationship further, investigating how SNSs design and use tends to impede and twist the dialogue between these two systems.

2.3.2 Detrimental consequences to users' life and well-being: impairment of free will

Many studies have already shown how SNSs use is linked to users' dissatisfaction (Boers et al., 2019; Damico & Krutka, 2018; Rozgonjuk et al., 2020). For example, the University of North Texas together with the University of Central Florida conducted a study among pre-service. The principal reasons that would lead participants to engage with SNSs use was not related to their happiness nor to their well-being. Some of the main reasons provided were habit, curiosity, procrastination or boredom (Damico & Krutka, 2018, p. 113). Habitual users don't principally use SNSs platform to search for engaging content, but more so to avoid boredom or to escape everyday life.

A study carried out by Natalie Pennington also brings interesting insights. Here she show the effect on users who decide to quit the use of social media (Pennington, 2020). In her article *Quitting social media: a qualitative exploration of communication outcomes*, the researcher shows the results of the interviews conducted with individuals who decided to quit using Social Networks**.** After they stopped using social media, the participants affirm that they encountered communicative benefits when it comes to their personal relations in their lives. Without the distraction of Social Networks, their testimony describes an increase in meaningful interactions with close ties, as well as a decrease in the constant comparison with people they were not close with. However, these benefits were paired

with relapse symptoms. One of the main consequences they experiences was feeling "out of the loop", as if they were missing out on relevant experiences they could not have access to. Overall, this goes to show the conflict in the motivational and volition system of users: on the one hand, they were able to experience the benefits of not being on the platforms. On the other hand, they were still conflicted in judging whether these benefits could overcome the drawbacks. They experienced a conflict not only between their motivational and volition system. Their own volition was torn between the judgement of the situation, as either positive or negative.

Together with detrimental consequences on users' free will, SNS cause damaging effects also on users' well-being. Unhappiness with screen time is not the worst of the negative consequences SNSs use carries. Bullying and trolling, for instance, are only two negative side-effects that people who engage with SNSs usage may encounter (Kross et al., 2021, p. 61). Despite the rhetoric of SNSs companies who claim that they aim at connecting users, research shows how SNSs users tend to feel more and more isolated (Dunbar, 2016; McDool et al., 2016; Sabatini & Sarracino, 2014). I am not claiming that social media necessarily determines unhappiness in the user. There is a great deal of research that shows how the use of social media can actually generate an increase in the connection among its users, or how they appreciate spending time on these platforms (Hobbs & Burke, 2017; *Online Social Integration Is Associated with Reduced Mortality Risk | PNAS*, n.d.). Connections created on SNSs platforms have the potential to be positive and determine beneficial consequences. However, the mechanism that determines the fruition of content on SNSs website is not calibrated to provide the user with the "best" content, meaning the content that can most likely enhance the user's well-being. It is just aimed at providing content that matches with the user's interests, namely that content that generates the most likes, keeps the user on the screen for the longest time, and increases engagement rates (Warren, 2021).

Together with influencing the users' well-being and their habits in their lives, SNSs use also molds the way people on social media platforms perceive themselves and the people around them. The next paragraph will investigate how the self and social identity of people is influenced by SNSs use.

2.3.3 Social identity and self-categorization

The elements that constitute the addiction definition – namely acquired appetite, habitual performing, clash between motivational and volitional systems and detrimental consequences – have been analyzed and proven to be met in the previous sections. The last part of the definition of addiction outlined in the first chapter reads "This course of action strongly shapes the identity of the individual,

who ends up identifying with that character". Patterns of addiction shape the identity of the addict, in the sense that the life of the person revolves around the fact of them identifying – and being identifying – as such. In this paragraph I will exemplify how SNSs dynamics can shape the identity and self and social perception of users.

One of the main downsides of SNSs use is FOMO. FOMO – an acronym that stands for "fear of missing out" – is a condition that has started to be investigated in recent years. It indicates the perception of other people experiencing rewarding events that one is missing out. This is also linked to the desire to know what other agents are doing and to always stay connected with them, not to miss any chance of being notified of an eventful experience happening (Liu & Ma, 2019; Przybylski et al., 2013). Przybylski et al. link FOMO to self-determination theory and social media use. In their research, FOMO is framed as a phenomenon that arises when an agent's psychological needs are under the minimum bar of satisfaction. The fear of missing out on eventful happenings is not by definition linked to the use of SNSs. It is an intrinsic need of human beings that is not being met. What SNSs platforms do is amplifing this feeling exponentially. When someone opens the Instagram app and all of their friends' Instagram stories show them having fun at a party where they were not invited, the sense of loneliness and being left out is intensified. This perception can influence the way the user see and perceive themselves.

FOMO is a useful example of showing how SNSs use can influence self-perception. Another aspect that the use of SNSs platforms can also affect is the development of one's identity. This is particularly evident in young users, who find themselves in an early developmental stage of their personality and face the social pressure that SNSs behavior imposes (Robin, n.d.). The pressure of being available at all times, matched with FOMO, and the pressure to adapt certain behaviors only to fit in in a specific stereotype mold the identity of teenagers in a way that identifies them as SNSs users. Their routines revolve around social media use, to the point of waking up in the middle of the night to log onto their profiles, which harms their sleep schedule and quality (*Teens' Night-Time Use of Social Media "Risks Harming Mental Health,"* 2015). However, as Pickard underlines in her analysis of self and social-categorization and addiction, there is a value attached to this kind of behaviors. The same way the use of drugs for the user determines self-categorization as an "addict", SNS use determines self-categorization as a social media user. Using SNS platforms provides a way to enter a specific social circle, and keeping oneself out of it is becoming increasingly more difficult. Instagram, Facebook, TikTok users see the value attached to engaging in SNSs use, therefore decide to engage with it especially in order to consolidate their social identity. Being a SNS user molds the routine and the identity of the person who has an Instagram or Facebook account the same way the

identity of the drug addict shapes the way they perceive themselves and they interact in their social group and in society.

Conclusion

This chapter has exemplified how the concept of addiction outlined in the first chapter can be applied to the concrete case of SNSs platforms. I have shown how SNSs platforms are designed to instill an acquired addiction in their users through behavior-modification mechanisms that tend to instill habitual actions. This leads to changes in the habits and routines of the agents, as well as a modification in the perception of their own self and identity both from a personal and a societal perspective. SNSs companies deliberately design their product in order to instill these kinds of changes in the user. The next chapter will address the question of whether this type of design is politically and morally relevant and whether it needs to be regulated.

3. Taking action: philosophical scenarios for policy intervention

So far, I've outlined my working definition of addiction in Chapter One, and I have subsequently applied it to how SNS platforms are designed in Chapter Two. This chapter brings them together and aims at answering the following question: *Should Social Network Sites design for addiction be regulated?*

The aim of this chapter is to show how the evidence gathered in Chapter Two is politically relevant. In the political scenario of deliberative democracies, political autonomy represents one of the pillars of the possibility to deliberate. In this chapter I will show how SNSs, by designing for addiction, undermine the political value of autonomy. By undermining agents' deliberative abilities through design for addiction, SNSs pose a threat to democratic governments. Furthermore, I outline a potential political intervention I imagine in order to regulate the situation. My goal is to establish the political relevance of the topic and provide a valid intervention that will mitigate SNS design for addiction.

In section 3.1, I answer the question of why SNS design for addiction is politically relevant. In section 3.2, I present my argument on why designing SNS for addiction is morally unacceptable and which consequences this entails on a political level. In section 3.3 I show which relevant public interests are being put at stake by SNS addiction by design, grounding my research on values presented in the Universal Declaration of Human rights and liberal democratic values. In section 3.4, I outline the policy intervention I imagine, based on the evidence and concepts developed throughout the chapter. Finally, section 3.5 presents three main objections that could be moved to my argument, and provides some inspiration on how to further investigate the topic based on these objections.

3.1 SNS design for addiction: why is it politically relevant?

So far, my research has focused on developing my philosophical account of addiction, which I have subsequently applied to the way SNSs are designed. The consequences of these specific design choices have been proven to have addictive effects the individual user. This section will focus on proving how the effect on the individual user are relevant on a public and political level. To do so, I will introduce a working definition of public interest in accordance with a theoretical account of deliberative democracy (Freeman, 2000). I choose to consider deliberative democracy as a political ideal as it poses strong emphasis on values such as freedom, independence, autonomy, and equality of citizens. Moreover, it allows me to narrow the scope of my research as far as a potential intervention is concerned. The account of deliberative democracy I will use serves as a basis for justification of an EU-level political intervention I will later argue for. In this sense, the European

Union appears uniquely positioned, as it possesses a unified user-sovereignty understanding. I am not arguing for the absolute validity of the values deliberative democracies defend. Yet, I believe deliberative democracy to be a valid starting point for the debate.

Deliberative democracy, as an ideal of political relations, represents the political system of European countries where SNSs are used. In a deliberative democracy, citizens express their deliberation and judgements to implement measures that ensure the common good of citizens (Freeman, 2000, p. 382). For the scope of my thesis, the concepts of public and of common good will be used interchangeably. I am aware of the fact that there are subtle nuances through which the two terms can be distinguished (Douglass, 1980). However, for the scope of this research, such nuances do not determine any relevant impediment to my argument. In a deliberative democracy, citizens reflect on the interests of the individual and of the community. They evaluate and choose the proposal that most fits their needs and interests (Freeman, 2000). As Freeman describes it, the public interest in deliberative democracies is defined as the circumstances that can enable and maintain citizen's freedom, independence, autonomy, and equality. Before deliberating on any subject, citizens who belong to a deliberative democracy need to take into account these preconditions in order to make a decision. Among the primary features that constitute the political ideal of deliberative democracy, Freeman mentions that political agents need to be able to be free and equal participants in the civic life and need to be able to develop their own independent conception of the good (Freeman, 2000, p. 382). These different concepts of the good are constitutionally protected and are considered to be publicly legitimate. The possibility to develop individual conceptions of the good is part of what constitutes the common good. To develop their own individual conceptions of the good, citizens their deliberation system needs not to be impaired so that their deliberation can focus on maintaining and fostering that good. In this sense, citizens need to be autonomous in their process of deliberation. Deliberative democracy advocates for the right to autonomy as a political value. Autonomy, defined as the ability to independently develop one's own conception of the good, represents one of the core values that allows for citizens to develop their independent conception of the good. In the process of developing their own conception of the good, citizens need to be self-legislative (*auto-nomos*), give themselves the rule for deliberating about their own good. They have the right to live accordingly to their own reasons and motives, without the interference of external influences (Christman, 2003). Before that, it is crucial to be mindful of the distinction between freedom and autonomy. While the first refers to the freedom to act in a Humean sense, autonomy indicates the independence of will. I have already mentioned in chapter one, an agent's will is undoubtedly influenced by external factors – namely their own interests or the one of their community. However, to be autonomous, the agent needs to be able to reflect upon their interests and perform through an alignment of motivational and

valuation systems. I have already exemplified how motivational and valuation systems do not always need to coincide for the subject's will to be considered free. Nevertheless, as Gary Watson observes, the cases where desires generate passions that does not match with what the agent *intended to* do, belong to that domain of actions that are independent from the subject's judgement of the good (Watson, 2003, p. 342). Once an agent cannot judge for their own good or the one of their society, the values at the base of deliberative democracy – such as the autonomy of the citizens – are at stake. Addictive patterns tend to disrupt the connection between an agent's motivational and valuation system, excluding these kinds of judgements from the conception of the good. There are cases where individual discrepancies of this kind do not represent a threat to one's autonomy. For instance, when deciding to not eat chocolate cake for a week, and then ending up eating chocolate cake at a friend's birthday party. In that case, the motivational and valuation system are clashing, but this does not determine a case of addiction. Once the lack of alignment between the two systems becomes habitual, societally shaped, etc., that is when the addictive pattern emerges. SNS companies design for addiction, which means their platforms are intentionally designed to pose a threat to agents' autonomy. As political autonomy is one of the central values that deliberative democracies aim at defending, SNS design for addiction poses a threat to the public interest.

In this section, I have shown how SNS addiction by design is – and should be – politically relevant. I now turn to the normative side of the issue, highlighting how public political interest is linked to moral values.

3.2 Moral relevance: how SNS addiction by design is normatively objectionable

While the previous section focused on the political relevance of SNS addiction by design, this section will shift the attention to the normative weight this design choice carries. In order to argue for this, I shall refer to the argument raised by Bhargava and Velasquez in their paper *Ethics of the Attention Economy: The Problem of Social Media Addiction* (Bhargava & Velasquez, 2020). Linking back to what I have argued for in the previous section, the attitude SNS companies adopt when designing for addiction does not protect agents' autonomy. As I have shown in chapter two, SNS platforms are adaptive in the sense that the more users interact with them, the more the platform will shape around their interests (Bhargava & Velasquez, 2020, p. 14). Bhargava & Velasquez highlight how this aspect of the process is demeaning towards the subject. Not only SNSs create addictive patterns in the user, which lead to undermining their autonomy. Users themselves carry out that process. An interesting parallel can be drawn with paternalistic policies. Paternalistic policies implemented by governments ask the government to justify their intervention under the assumption that citizens don't know what

is best for them in that specific case (Cornell, 2015). This aspect, as argued by Cornell, does not go against the value of autonomy. As long as the government – or any power representative – is able to show that the individuals would not be able to collectively choose what is best for them, then the paternalistic intervention is justified and does not undermine agents' autonomy. Despite this, in her paper *Paternalism, Unconscionability Doctrine, and Accommodation*, Shriffin argues that paternalism – and paternalistic policies – still represent a lack of respect, even when the paternalistic intervention is justified (Shiffrin, 2000, p. 220). The disrespect Shriffin argues for is different from the insult of violation of autonomy rights that Cornell mentions. The right to autonomy may be disrespected in Shriffin's sense but not be violated the way Cornell describes it, if the authority in power provides a fair justification for the paternalistic policy put into place. What Shriffin argues for – and what I also subscribe to – is the fact that paternalism is prima facie morally problematic (Shiffrin, 2000, p. 221). In the case of SNS platforms, the distinction between violation and disrespect of autonomy fades into the background. The objections SNS representatives have made against those who accused these platforms to design for addiction have already been proven to be insufficient in the previous chapter. Therefore, the attitude that SNS companies have towards their own users' autonomy is morally objectionable. In addition to this, Bhargava & Velasquez add a last piece to the puzzle, the mockery in addition to the damage. Not only are SNS companies lacking respect towards their users' autonomy, but it is the users themselves who are contributing to the building of this paternalistic dynamic. This aspect is directly linked with another element that Bhargava & Velasquez regard as morally problematic, which is exploitation.

For the scope of my thesis, I define exploitation as taking unfair advantage of someone, by using their weaknesses or vulnerabilities to achieve one's personal gain (Wood, 1995; Zwolinski & Wertheimer, 2017). While the threat addiction by design poses to autonomy shows how this design choice is politically relevant, exploitation of weaknesses and vulnerabilities sheds light on its moral relevance. Moreover, this allows to expand the framework beyond the political realm of deliberative democracies.

In the he same way paternalistic policies do not always determine moral harm or injury, exploitation is not always morally objectionable. For instance, when in a game of soccer a player takes advantage of the goalkeeper weakness to score a goal, that is not considered a case of exploitation. An approach is exploitative when the the object of the exploitative act is not respected (Wood, 1995). The lack of respect has already been discussed in the previous paragraph, regarding the autonomy of users. Therefore, here I will focus on the pervasiveness and inescapable vulnerability that SNS addiction by design induce in the users. SNS users are vulnerable, as in the power relation against SNS self-adapting algorithms they are in an inferior position. Wood himself cites the addiction

case, mentioning how an addict is a clear case of a vulnerable subject that a pusher can exploit to advance their own ends (Bhargava & Velasquez, 2020; Wood, 1995). In the same way, SNS companies first instill the addictiveness in their product, influencing users' behavior patterns, and then exploit this vulnerability to achieve financial gain. Another aspect that Bhargava & Velasquez point out is how this exploitation of users' vulnerability happens on a global scale, on a daily basis (Bhargava & Velasquez, 2020). In order to end an addictive pattern, whether behavioral or substance-based, addicts are often prevented from having access to that specific substance or behavior. So much so that we speak of an addict being "clean" for a certain amount of time, meaning that they haven't used that substance or performed that action in a while. In the case of SNS platforms, being "clean" can prove more difficult than expected. Not only are SNS companies building an exploitative relation with users by leveraging a vulnerability that users themselves have helped in creating. They can also rely on the fact that exposure to the internet in contemporary society is virtually unavoidable. Of course, the internet plays a crucial – and positive – role in many different activities. It is indeed this pervasiveness that makes SNS companies behavior even more deceitful. As internet platforms are hardly avoidable in everyday life, this makes it extremely difficult for addicted users to establish a healthy behavior online. As Bhargava & Velasquez point out, this creates countless opportunities for SNS platforms not only to addict but, also and most importantly, to re-addict users. Therefre, the exploitative relation becomes hardly impossible to avoid, which makes it complicated to "just say no". I want to highlight how an exploitative relation does not need to be addictive. However, I believe addictive relations are exploitative by definition. As Wood argues, cases of addiction are by definition cases where one side of an unbalanced power relation takes advantage of the vulnerable side (Wood, 1995). Therefore, this shows how SNS addiction by design necessarily leads to exploitation, which is morally objectionable.

This section has shown how, through patterns of exploitation, SNSs addiction by design is morally objectionable. The next section will further flesh out how SNS addiction by design affects the relevant public interest, merging the normative results obtained in this last section.

3.3 Relevant public interest: how SNS addiction by design threats current regulations

Before outlining my concrete suggestions regarding the political intervention I imagine, this section will shine a light on how SNS addiction by design is threatening proclaimed public values. I have already shown in section one how this design choice hinders the public interest of deliberative democracies. This section further delves into how this design approach is affecting relevant values by analyzing those defended in the Universal Declaration of Human Rights and in the General Data

Protection Regulation (General Data Protection Regulation (GDPR) – Official Legal Text, 2018; Nations, 1948). The choice of these two documents is justified by the fact that my research is primarily focused on values held by deliberatively democratic countries. The majority of those happen to be the ones where the General Data Protection Regulation is applied. However, as the Universal Declaration of Human Rights is extended to a broader scope of ideals of political relations, I believe the results of this analysis can also be extended to different scenarios.

The (General Data Protection Regulation (GDPR) is a regulation put into place by the European Union on May 25th, 2018, to protect natural persons in the domain of personal data processing (*General Data Protection Regulation (GDPR) – Official Legal Text*, 2018). While the GDPR focuses on how data-processing choices are made and need to be communicated, for the scope of my research I will translate this into the way platform design choices are made and need to be communicated to users. Article six, titled "Lawfulness of processing", addresses the cases in which the processing of data is lawful and under which conditions. Instead of the data processing, I will focus my attention on the dynamics of SNS design choices ("Art. 6 GDPR – Lawfulness of Processing," 2018). Provision (f) is particularly relevant in this case, as it reads:

> "processing [*SNS design choices*[5]] is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child."

As discussed in the first and second section of this chapter, and more thoroughly in chapter two, the design choices carried out by SNS platforms have proven to be not in favor of the autonomy nor the well-being of the user, but of the companies themselves. As far as these stay within the legitimate interest pursued by corporations, this type of behavior is allowed. However, when fundamental rights and freedoms – such as the right to autonomy or the freedom to independently deliberate – are undermined, the design choice or process becomes unlawful. It is important to keep in mind that article six mentions that *at least* one of the six conditions need to be met in order for the data processing – in this case, design choice – to be considered lawful. However, the previous five provisions are either also not met by addiction by design or do not comply with design choices.[6]

[5] My modifcation
[6] For the scope of my thesis, I will not analyze all of the six provisions mentioned in article 6. However, they can be found and analyzed at https://gdpr-info.eu/art-6-gdpr/.

Another article that can be linked back to SNS design choices is article twenty-two, titled "Automated individual decision-making, including profiling" ("Art. 22 GDPR – Automated Individual Decision-Making, Including Profiling," n.d.). The first paragraph of the article states:

> "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."

As I have argued for in chapter two, the methodologies employed by SNS designers – such as personal profiling, infinite scroll, etc. – aim at increasing the engagement the user has with the platform. When creating this engaging relation, the user becomes the vulnerable part of an uneven power relation. Wood argues that cases of addiction are clear cases of exploitation of a vulnerable subject (Wood, 1995). For this reason, I argue that SNS users are subject to decisions based on automated processing which are significantly affecting them, which goes against what is stated in the article. It is important to shine a light on the difference between the GDPR effects and SNS effects. In this case, the statute talks about legal effects, as the GDPR has legal force when it comes to enforcing these principles. The position I am defending is based on effects on the individual user. Therefore, the word "effect" here does not have the same connotation. However, in the case a potential GDPR 2.0 was put in place, its effects would be legally enforceable as well. Similarly, article twenty-two also presents a series of provisions. Provision (c) states that the article does not apply in case the decision is based on data – here, design choice – the subject has given explicit consent. However, giving consent requires an agent to be able to independently deliberate. As I have demonstrated both in the two previous sections as well as in chapter one of this thesis, agents who suffer from addiction are prevented from deliberating. Therefore, they are unable to give their consent, as their ability to deliberate is impaired by their addiction to the platforms. It is important to keep in mind that, unlike in the case of the GDPR, not all users who use SNS have their ability to deliberate impaired. However, the design choices that SNS companies put in place are intended to generate addictive behavior. Therefore, every user when signing up on SNS platforms runs the risk of becoming addicted to it because of a design choice that was made. The other two provisions – namely (a) and (b) – cited in the article refer to the necessity of the corporation to complete the procedure and the obligation to safeguard the subject's rights, freedoms, and legitimate interests ("Art. 22 GDPR – Automated Individual Decision-Making, Including Profiling," n.d.). I have argued that the design mechanisms put in place by SNS corporations do not *need* to generate addictive patterns in users. I will further argue for this position in the next section. Consequently, provision (a) does not hold. As far as the safeguard of the subject's

rights, freedoms, and legitimate interests, I have proven that this type of design does everything but safeguard the private and public interest of the user. Therefore, provision (b) also does not apply. It follows from this that SNS design methods do not comply with article twenty-two of the GDPR, in case this was applied to design methods instead of data usage.

In this section, I have taken two of what I consider to be the most relevant articles in the General Data Protection Regulation and shown how, if translated into design regulation policies, they would not allow for SNS platforms to be designed the way they currently are. There are other articles in the GDPR – for instance, article twelve ("Art. 12 GDPR – Transparent Information, Communication and Modalities for the Exercise of the Rights of the Data Subject," n.d.) – or other types of regulation that could equally apply to this case. For the scope of my research, I limited myself to consider the ones that I consider most relevant to my investigation. However, this shall not preclude a further evolution of the research in this sense. While this section has dealt with showing how SNS design choices do not comply with design choices, in the next one I will present the policy intervention I am imagining.

3.4 Policy intervention: GDPR 2.0 – General Design Protection Regulation

Following the argument developed in the previous section, this section outlines an idea for the type of policy intervention I imagine is needed to regulate SNS design and mitigate SNS addiction by design.

So far, the measures that have been suggested in order to limit SNS design for addiction have relied on single users or designer's initiative. For instance, at the end of their paper, Bhargava and Velasquez suggest five methods on how to limit the damages of SNS addiction (Bhargava & Velasquez, 2020). Some involve individual users and scholars, who should take the initiative to educate younger generations on the danger that SNSs can generate. Other solutions involve corporations themselves "to empower users to have a healthier relationship with social media" out of their own initiative (Bhargava & Velasquez, 2020, p. 23). They also suggest that platforms should make explicit when addictive design choices are put into place, and users could willingly opt-in only when they wanted to consent to it. As far as the last option is concerned, I disagree with it, as I feel it would not be advisable to ask a recently recovered addict if they wanted – or not – to continue with their addictive habit. As I have previously argued, addiction impairs an agent's ability to deliberate, so leaving them the choice would only increase the risk of creating re-addicts. Regarding the other options they suggest, I think they could lead to positive outcomes, but they are not sufficient. As far as user responsibility is concerned, leaving the users to take action on their own initiative does not take into account the impairment of deliberation they are suffering from and the power imbalance

between SNS corporation and the single user. Expecting an SNS corporation to take the lead is utopian and naïve. As Sinan Aral points out, to succeed in regulating SNS there is a need to address design choices, as well as the social, economic and regulatory setting where they are made (Aral, 2020, p. 377). I agree with the author when he states that design choices are important, but their regulation should not come from those who are implementing the design themselves, let alone from the corporations who are generating financial profit from it. In this case, policy intervention is the only way SNS design can be regulated and which principles SNS designers are required to follow. I support Aral's idea of establishing and enforcing norms that are able to protect human agency and translate this into concrete action (Aral, 2020, p. 379). I will further develop these two points in section 3.5.3.

My proposal is to adopt a general regulation on design norms, similar to the General Data Protection Regulation to achieve this result. As this regulation aims to govern SNS design choices, I will name it GDPR 2.0 – General Design Protection Regulation. Similarly to GDPR, I imagine GDPR 2.0 being a regulation in EU law. It could be integrated with the principles that are currently being discussed in the *Declaration of Digital Principles – the 'European way' of doing society* (*Declaration of Digital Principles – the 'European Way' for the Digital Society*, n.d.). GDPR 2.0 would primarily apply to Social Network Sites, but its principles would be applicable to all the platforms that fall into the social media category. It is not my aim in this thesis to outline all the criteria that would constitute this regulation. The sub-question I aimed at answering so far in this chapter asked whether SNS addictive design should be regulated. I have shown how the way SNSs are currently designed is a threat to political and moral values in deliberative democracies. For this, I have argued for state intervention in the regulation of design choices. In the next section I will specifically address some limitations that the implementation of a design regulation could entail, and some of the main objections that could be made to my thesis.

3.5 Potential objections

In this final section I will consider and address to some objections that could be made to the argument I have outlined in Chapter Three, and to my thesis.

3.5.1 Technological determinism – an Ellulian objection

> […] An autonomous technology. This means that technology ultimately depends only on itself, it maps its own route, it is a prime and not a secondary factor, it must be regarded as an "organism" tending toward

> closure and self-determination: it is an end in itself. Autonomy is the very condition of technological development.
>
> (Ellul, 2014, p. 430)

In his work *The "Autonomy" of the Technological Phenomenon*, French philosopher Jacques Ellul defines technology as being autonomous. The technological system, which is ascribed an independent agency, is described as a closed organism that aims to expand and survive. This view fits in the definition of technological determinism or, in the words of Smith and Marx, hard technological determinism (Marx & Smith, 1994, p. xiii). In the view of hard technological determinism, technology possesses an agency in itself, and it follows a pre-established path that cannot be subject to change. Borrowing from Ellul's vocabulary, technology presents itself as an intrinsic necessity. The desires and intentions of human beings play no role in shaping of the path of technological development but are themselves shaped by this necessity. There are two main pillars that support the techno-deterministic Ellulian view: firstly, the technological sphere and the human sphere are seen as two separate realms. Secondly, the first one possesses its own agency, which is autonomous and pre-determined. The second one is not only subject to technological influence but is powerless when facing it. Furthermore, Ellul explicitly refers to the role of governments and regulating bodies when it comes to shaping the path of technological development (Ellul, 2014, p. 432). The state cannot help intervening. However, state intervention merely follows the technological path. Its actions are unable to steer the path of the technological system. Even when governments try to change the path of technological development, their actions fit in the technological purpose. In this scenario, imagining a policy intervention appears not only as a futile attempt to stop what is unavoidable. It is yet another confirmation of humanity bending under technological rules. The GDPR 2.0 would serve as the umpteenth example of humans taking upon technological goals and assimilating them as their own. In this pre-determined plot, Ellul does not offer a solution. In *The Technological Society*, the philosopher states that his aim was to sound "a call to the sleeper to awake" (Ellul, 1967, pp. xxxi–xxxiii). But once awake, what should the sleeper do?

In his essay *Resistance Is Futile: Toward a Non-Modern Democratization of Technology*, Peter-Paul Verbeek proposes a different takeaway on the issue (P.-P. Verbeek, 2013). As mentioned in chapter one, his philosophy builds on mediation theory. Human beings are technological. This definition challenges one of the two pillars of Ellulian's theory: humans and technologies should not be considered two opposite poles in a dialectic relationship. Verbeek rightfully points out that conceptualizing this co-shaping relation as an oppressive one is comparable to trying to resist gravity. I agree with Verbeek when he states that:

> […] the model of oppression and resistance might not be the most productive model if one wants to change undesirable configurations of humans and technologies.
>
> (P.-P. Verbeek, 2013, p. 77)

Seeing the technological system as an oppressor and humanity as the oppressed subject leads to a dead end. I will take the case of Instagram as an example. Similarly to the Facebook example Verbeek provides, Instagram functions as a media through which humans experience friendship, photo-sharing, conversations. Reducing its function to a mere cog in the wheel of the technological system means depriving it of the complexity it plays in the socio-cultural scenario. Verbeek highlights how the functions of mediating technologies are not always desirable. Nevertheless, mediation theory answers the call to action that technological determinism is unable to respond to. Where technological determinism sees a unidirectional oppressive relationship, mediation theory offers a co-shaping answer. By stating this, I do not mean to dismiss the importance of technological determinism. I agree with Sally Wyatt when she states that it is detrimental and futile to ignore the warnings of technological determinism (Wyatt, 2008). She compares this neglect to ignoring "the equivalent of a thundering herd of elephants" (Wyatt, 2008, p. 171). However, I agree with her when she states that one of its main issues is the inability to account for human choice and intervention. Even more worryingly, it absolves humans from the responsibility they have towards the development of technologies. In this sense, my research is situated more on the pole of soft technological determinism. Here technology is situated in a complex cultural, economic, social and political realm (Marx & Smith, 1994, p. xiii). After being awaken from their sleep, the role of humans is to participate in the shaping of this realm.

3.5.2 One must imagine the user happy: A case for non-addiction or for an addictive society

Someone could argue that, overall, Social Network Sites are not addictive after all. Ultimately, a lack of alignment between motivational and volitional systems, merged with detrimental consequences for the individual, are characteristics that can be ascribed to other activities besides SNSs. Some of them match with what are generally considered to be addictions in everyday discourse: alcohol abuse, drug abuse, problem gambling. However, the same definition could be stretched to encompass other activities performed daily, such as going to work. Are people addicted to their job the moment they admit they don't like it and yet choose to go to work every day? I have two answers to this question.

To the reader who may claim that my definition of addiction makes it too easy for too many activities to fit in it, I want to remind the words of Gary Watson. Watson talks about the lure of

seducing desires. Motivational obstacles invite the agent to go with the drive of beguiling desires, deceiving them in a misleading way. The desire becomes the object of fascination, and the subject cannot do anything but surrender to that fascination. In this sense, the relationship to the addictive behavior can be compared to a toxic romantic relationship. The power relation is not put in place through brute force. It is a sort of enchantment that seduces the agent to believe that the toxic behavior they are adopting is in fact what they want to perform. However, to certify the toxicity of the activity, one needs to refer to a third observer. To follow the metaphor, t is generally a doctor, a family member, friends, partners who are able to attest the toxicity of the relation. The person who is held captive by the lure of the toxic relationship is rarely able to assert the toxicity of it. One reason for it is the fact that their motivational system and their valuation system systematically do not match. Therefore, they are unable to deliberate about the toxicity of their addiction. Nevertheless, there is another option I want to consider.

It could be the case that my definition is correct *and* not too broad. Does this mean that most activities humans perform on a daily basis are addictive? This raises philosophical and societal questions which deserve further research and investigation. Nonetheless, I will not be addressing them in this thesis. For now, I shall allude to the Myth of Sisyphus, as described in Camus' homonymous philosophical essay (Camus, 2013). In his work, French author and philosopher Albert Camus inquires about the absurdity of existence. He states that living is never easy, as the numerous contradictions of life follow one another in an endless spiral of absurdity. To bear the burden of the absurdity of life, the only viable solution Camus proposes is suicide. Not in a physical sense but in a spiritual sense. The need for clarity that haunts the humankind can be overcome by the *homme absurde* (Camus, 2013, p. 58). The absurd man is the one who has accepted the absurdity of life and knows there is no more space for hope. At the end of his essay, he describes the pagan myth of Sisyphus as the leitmotif that can best describe his existentialism. Sisyphus dared to challenge the Gods, who punished him by sentencing him to push a very heavy rock up a mountain for eternity. But his fate is tragic only in those spare moments when he realizes his condition.

> Si ce mythe est tragique, c'est que son héros est conscient. […] L'ouvrier d'aujourd'hui travaille, tous les jours de sa vie, aux mêmes tâches et ce destin n'est pas moins absurde. Mais il n'est que tragique qu'aux rares moments où il devient conscient.[7]

[7] "This myth is tragic only when our protagonist is conscient. […] The workman of today works every day in his life at the same tasks, and this fate is no less absurd. But it is tragic only at the rare moments when it becomes conscious."

(Camus, 2013, p. 165)

Camus compares Sisyphus to the worker who goes to work every day. Their existence is not less absurd than the one of Sisyphus forced to push the rock up the mountain for eternity. For the reader who is more interested in the existential meaning of addiction in everyday life, this may be an interesting starting point of philosophical interrogation. Unlike in the myth of Sisyphus, we may accept to live a life of absurdity – or addiction – and be happy with it.

> Lui aussi juge que tout est bien. […] La lutte elle-même vers le sommets suffit à remplir un cœur d'homme. Il faut imaginer Sisyphe heureux.[8]
>
> (Camus, 2013, p. 168)

3.5.3 Individuals and tech companies: two alternatives to state regulation

The conclusion I reached in this chapter is that SNS design for addiction threatens the political value of autonomy. In addition, SNS design companies exploit user vulnerabilities in a paternalistic dynamic where users themselves co-design their own addiction. There are two other possible solutions to this matter that I decided to consider.

The first option would be for the users to combat the addictiveness of platforms autonomously. In this sense, they could either abandon the platforms altogether, completely avoiding the risk of SNS addiction, or taking responsibility for their own behavior and regulate their SNS use. The first option is indeed pursuable. There are countless stories of people who have stopped using SNS, and are benefitting from it (Youn, 2021). For instance, Sharon Baldessari, who gave up both Instagram and Facebook despite working in the realm of technology, says she feels much better after leaving the platforms and after she started reflecting on her motives. The same goes for Meaghan Connaire, who, after giving up Instagram and Facebook, states that she has more time for activities that really spark her interest. There are individual stories that prove that leaving SNS is possible. However, as Bhargava and Velasquez point out, SNSs are becoming increasingly necessary to navigate today's world (Bhargava & Velasquez, 2020). Not everyone can have the privilege to not use them. Some people may need LinkedIn to create networks that will allow them to start their next career. Some others may need Facebook to keep in contact with family members who live in a different country. Moreover, when users quit the platforms these do not stop being designed for addiction. Therefore, the fact that SNS companies keep designing to impair people deliberative capacities and to

[8] "He also judges that all is well. The fight towards the top is sufficient to fill a man's heart. One must imagine Sisyphus happy." Both tranlsations are made by me.

threatening their autonomy still holds. For that reason, individual action can lead to beneficial individual consequences, but does not provide a valuable solution as far as political and moral relevance is concerned. The second option, namely the one that states that user can self-regulate their own behavior on the platform, is not viable because of the nature of design for addiction itself. As I have demonstrated in Chapter One and Two, design for addiction determines an impairment of users' deliberative capacity. It is important to highlight that, like any other addictive substance or behavior, it isn't the case that each user becomes addicted. Therefore, some users may be able to successfully regulate their own presence on social media. It is nonetheless true that, being designed *for* addiction, SNS platforms are aimed to instill that behavior in users. Cases of users who are able to self-regulate their relationship with SNSs are the exception to the rule, but their ability to deliberate is still at risk of being impaired. Therefore, users regulation may be a solution to the symptom, but not to the cause of addiction.

The second option is companies' self-regulation. SNS companies could individually adhere to internal standards which state that design for addiction should not be pursued. There are two pragmatic reasons why I think this option is not viable. Firstly, as shown by the declarations made by Clegg at the beginning of this thesis, SNS companies such as Facebook are already claiming that their design is not intended for addiction (Clegg, 2021). Nevertheless, as I have shown throughout this thesis, this is not the case. SNS current design choices are aimed at impairing users' deliberative capacities and are therefore a threat to their autonomy. On top of that, Clegg's statements create a precedent to not trust tech companies in creating internal standards to adhere to when it comes to combating design for addiction. The second and most relevant point is the fact that, when it comes to regulating design for addiction, SNS companies find themselves between two conflicting values. On one side, they should aim at protecting users' autonomy and not exploit their vulnerabilities to gain financial profit from it. On the other side, they have to gain financial profit. The words of Natasha Dow Schüll perfectly capture this, when she says: "In the online economy, revenue is a function of continuous consumer attention – which is measured in clicks and time spent" (Busby, 2018). There is a positive correlation between time spent on the platform and financial income SNS tech companies gain from it. Therefore, the desire they may have to design their platforms following methodologies that do not determine addiction is hindered by their desire of financial gain. For this reason, I argue that there is a need for an impartial third party to step in the regulation of design features. I argue for government intervention as the state represents an impartial third party when it comes to the conflict of interest that tech companies are facing. Moreover, the main value that addiction by design is hindering is one of the underlying principles of deliberative democracies. To sum up, I believe government intervention to be the most coherent solution to tackle the issue of addiction by design.

Conclusions and final thoughts

The goal of my thesis has been to prove that Social Network Sites are currently designed to instill addictive behavior patterns in users and, consequently, to show if they should be regulated. This was to show that addiction interferes with deliberative capacities. Therefore, users cannot take individual action to mitigate the situation. Demanding users to act upon their lack of control over their social media behavior would be like asking an addict to exert control over their addiction. For this reason, I have argued for the need for a centralized and political intervention that provides guidelines and regulations for SNS design choices. This last part of my research will bring together the findings I have gathered throughout the writing of this project.

When I first started thinking about the topic of my final project, what mainly motivated me to pursue this research line was a combination of curiosity and anger. How is it possible that platforms that we use daily are profiting from addict users' vulnerability? And why are people addressing this issue only on a superficial level? When starting to delve deeper into all the aspects that surround this issue, I directed my research toward a more complete understanding of what addiction is.

Considering Watson and Wallace's conception of will and addiction, I was able to trace a distinction between motivational and valuation systems. Through the intricacy of multiple understandings of the philosophical concept of addiction, this perspective allowed me to create a framework where to put addictive desires and to show how they impair an agent's deliberative capacity. The contribution of Pickard's work was also invaluable: combining the impairment of deliberation with self and social categorizations of addiction enriched my perspective and strengthened my argument. It also showed a less prevailing side of addiction, which is often disregarded in the compatibilist philosophical understanding of addiction. The addict's identity is shaped by their action, but also by the way they are perceived by society and by themselves. This last aspect allowed me to merge my definition of addiction with mediation theory. Through the mediation lens, I was able to show how the addictiveness of the technology does not only reveal itself in the interaction between user and technology, but it shapes the human form of life also when SNSs are not in use. This chapter showed how the compatibilist approach to addiction can benefit from insights from both mediation theory and self and social characterization of addiction. These two allow to pinpoint the specific aspects in the human realm that can then be traced back to the co-shaping relationship between human and technology.

The theoretical framework that I built in the first chapter allowed me to show how many of the current SNS design choices actually match with the symptoms that characterize addiction. Thanks to the insights identified both in neuroscience and in psychological studies, I could show a strong connection between the multiple aspects that characterize addiction – impulsivity, impairment of

choice, habitual patterns, self and social characterization – and those that are typical of SNS platform design and consequences – endless scrolling, adaptive algorithms, FOMO, etc.

The final part of my thesis was the one I strived towards during my whole process of research. I approached the last chapter from the perspective of liberal democracies, as this is the type of government that rules in many of the countries where I took scientific and psychological findings from. This allowed me to narrow my scope and point out one of the main values that underlie liberal democratic societies: political autonomy. One of the very underpinnings of political autonomy is the ability to deliberate. In Chapter One I have shown how this same ability is undermined by addictive behavior, as the subject is systematically unable to line up their motivational and volitional system. By designing for addiction, Social Network Sites pose a threat to the autonomy of the user, and therefore need to be regulated. I must admit it was not easy to realize not only how much these platforms shape us, but also how hard it is to merge all these insights in a single, all-encompassing regulation. The options that I took into account besides state intervention were less satisfactory than this one. Both individual users' action and SNS tech company action were shown to be inapplicable in this scenario. Finally, it was particularly hard to respond to the technological-deterministic objection moved by Ellul. I agree with Sally Watt when she states that technological determinism should not be dismissed for its conceptual crudeness (Wyatt, 2008). Ellul provides an insightful take on the way technologies shape the development humanity. However, it fails to recognize the opposite direction of the relation. In this, mediation theory offers a viable alternative. In the words of Verbeek, resistance is futile. Political philosophy can play a central role in the regulation of technology design (P.-P. Verbeek, 2013).

As a final remark, I started this project with the hope of ending on a hopeful note. I wanted to outline a code that could once and for regulate SNS design. Yet, the more the project took shape, the more I realized that the complexity of technological addiction does not only lie in the behavior humans adopt when they are using the technologies. The more pervasive SNSs will become, the more complexities will arise when it comes to the shaping of the human-technology relation. In this complex scenario, coming up with exhaustive and complete regulations may be an obstacle too complicated to overcome. Nevertheless, disciplines and methodologies such as Value Sensitive Design or capabilities approaches bring successful examples of embedding moral and ethical values into technologies (Brey, 2015; Friedman & Hendry, 2019). I believe this research path can lead to SNSs designed not for addiction, but for autonomy. May scholars and fellow and future researchers build up on my investigation to further add to the human shaping of technologies, and may their recalcitrant desire not rest in their research path.

References

About Instagram's Official Site. (n.d.). Retrieved November 8, 2020, from https://about.instagram.com/

Aghadjanian, N. (2020, December 8). *A Second Look: Facebook Addresses What Netflix's 'The Social Dilemma' "Gets Wrong."* AList. https://www.alistdaily.com/technology/facebook-refutes-claims-made-in-netflix-documentary-the-social-dilemma/

Allen, M. (2017, November 9). *Sean Parker unloads on Facebook: "God only knows what it's doing to our children's brains."* Axios. https://www.axios.com/sean-parker-unloads-on-facebook-god-only-knows-what-its-doing-to-our-childrens-brains-1513306792-f855e7b4-4e99-4d60-8d51-2775559c2671.html

Amsterdam Instagram Tour. (n.d.). Top-Rated Tours of Amsterdam and Holland. City Walks, Museums, Countryside and Free Walking Tours. Retrieved May 5, 2021, from https://www.amsterdamclassictours.com/amsterdam-instagram-tour.html

Andersson, H. (2018, July 3). Social media apps are "deliberately" addictive to users. *BBC News*. https://www.bbc.com/news/technology-44640959

Aral, S. (2020). *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health--and How We Must Adapt*. Currency.

Art. 6 GDPR – Lawfulness of processing. (2018). *General Data Protection Regulation (GDPR)*. https://gdpr-info.eu/art-6-gdpr/

Art. 12 GDPR – Transparent information, communication and modalities for the exercise of the rights of the data subject. (n.d.). *General Data Protection Regulation (GDPR)*. Retrieved July 21, 2021, from https://gdpr-info.eu/art-12-gdpr/

Art. 22 GDPR – Automated individual decision-making, including profiling. (n.d.). *General Data Protection Regulation (GDPR)*. Retrieved July 21, 2021, from https://gdpr-info.eu/art-22-gdpr/

Bali: Tour Instagram di 1 giorno dei luoghi più suggestivi. (n.d.). Retrieved May 5, 2021, from /bali-l347/bali-tour-instagram-tra-i-luoghi-piu-belli-dell-isola-t162114/

Baym, N. K., Wagman, K. B., & Persaud, C. J. (2020). Mindfully Scrolling: Rethinking Facebook After Time Deactivated. *Social Media + Society*, *6*(2), 2056305120919105. https://doi.org/10.1177/2056305120919105

Bhargava, V. R., & Velasquez, M. (2020). Ethics of the Attention Economy: The Problem of Social Media Addiction. *Business Ethics Quarterly*, 1–39. https://doi.org/10.1017/beq.2020.32

Biscaldi, A., & Matera, V. (2019). *Antropologia dei social media. Comunicare nel mondo globale*. Carocci.

Boers, E., Afzali, M. H., Newton, N., & Conrod, P. (2019). Association of Screen Time and Depression in Adolescence. *JAMA Pediatrics*, *173*(9), 853. https://doi.org/10.1001/jamapediatrics.2019.1759

Boyd, danah m., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Boyle, D. (1996, February 23). *Trainspotting*. PolyGram Filmed Entertainment.

Brey, P. (2015). Design for the Value of Human Well-BeingWell-being. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains* (pp. 365–382). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_14

Busby, M. (2018, May 8). Social media copies gambling methods "to create psychological cravings." *The Guardian*, 08.05.2018.

Camus, A. (2013). *Le mythe de sisyphe: Essai sur l'absurde*. Sodis.

Carr, N. (2011). *The Shallows: What the Internet Is Doing to Our Brains*. W. W. Norton & Company.

Christman, J. (2003). *Autonomy in Moral and Political Philosophy*. https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/

Clegg. (2021, March 31). You and the Algorithm: It Takes Two to Tango. *About Facebook*. https://about.fb.com/news/2021/03/you-and-the-algorithm-it-takes-two-to-tango/

Cornell, N. (2015). A Third Theory of Paternalism. *Michigan Law Review*, *113*(8), 1295–1336.

Coxon, K. (2017). Attention, Emotion, and Desire in the Age of Social Media. In *Social media and your brain: Web-based communication is changing how we think and express ourselves*. http://site.ebrary.com/id/11293205

Damico, N., & Krutka, D. G. (2018). Social media diaries and fasts: Educating for digital mindfulness with pre-service teachers. *Teaching and Teacher Education*, *73*, 109–119. https://doi.org/10.1016/j.tate.2018.03.009

Declaration of Digital Principles – the 'European way' for the digital society. (n.d.). Have Your Say. Retrieved July 22, 2021, from https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13017-Declaration-of-Digital-Principles-the-%E2%80%98European-way%E2%80%99-for-the-digital-society/public-consultation_en

Do you obsessively check your smartphone? (n.d.). Retrieved May 5, 2021, from http://www.cnn.com/2011/HEALTH/07/28/ep.smartphone.obsessed.cohen/index.html

Doidge, N. (2007). *The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science* (Reprint edition). Penguin Books.

Douglass, B. (1980). The Common Good and the Public Interest. *Political Theory*, *8*(1), 103–117.

Dunbar, R. I. M. (2016). Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science*, *3*(1), 150292. https://doi.org/10.1098/rsos.150292

Ellul, J. (1967). *The Technological Society*. Vintage Books.

Ellul, J. (2014). The "Autonomy" of the Technological Phenomenon. In R. C. Scharff & V. Dusek (Eds.), *Philosophy of Technology: The Technological Condition: An Anthology* (2nd edition). Wiley-Blackwell.

Facebook by the Numbers (2021). (2021, January 4). http://www.omnicoreagency.com/facebook-statistics/

Frankfurt, H. G. (Ed.). (2003). Freedom of the Will and the Concept of a Person. In *Free Will* (2nd edition, pp. 322–336). Oxford University Press.

Freeman, S. (2000). Deliberative Democracy: A Sympathetic Comment. *Philosophy & Public Affairs*, *29*(4), 371–418.

Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.

General Data Protection Regulation (GDPR) – Official Legal Text. (2018). General Data Protection Regulation (GDPR). https://gdpr-info.eu/

Gertz, N. (2018). *Nihilism and Technology*. https://rowman.com/ISBN/9781786607027/Nihilism-and-Technology

Gertz, N., Verbeek, P.-P., & Douglas, D. (2019). Cyberwar and Mediation Theory. *Delphi*, *2*(2), 72–78. https://doi.org/10.21552/delphi/2019/2/5

Griffiths, M., Kuss, D., & Demetrovics, Z. (2014). *Social networking addiction: An overview of preliminary findings*. https://doi.org/10.1016/B978-0-12-407724-9.00006-9

Griffiths, M., Kuss, D., Pontes, H., & Billieux, J. (2016). *Where do Gambling and Internet 'Addictions' Belong? The Status of 'Other' Addictions (volume 2)*. Sage. https://doi.org/10.4135/9781473922143

Harris, T. (2019, June 25). Optimizing for Engagement: Understanding the Use of Persuasive Technology on Internet Platforms. *U.S. Senate Committee on Commerce, Science, & Transportation*. https://www.commerce.senate.gov/2019/6/optimizing-for-engagement-understanding-the-use-of-persuasive-technology-on-internet-platforms

Haythornthwaite, C. (2005). Social networks and Internet connectivity effects. *Information, Communication & Society*, *8*(2), 125–147. https://doi.org/10.1080/13691180500146185

Herman, D. (2000). Introducing short-term brands: A new branding tool for a new consumer reality. *Journal of Brand Management*, *7*(5), 330–340. https://doi.org/10.1057/bm.2000.23

Hobbs, W. R., & Burke, M. K. (2017). Connective recovery in social networks after the death of a friend. *Nature Human Behaviour*, *1*(5), 1–6. https://doi.org/10.1038/s41562-017-0092

Hosie, R. (2017, February 10). The psychological reason you can't stop checking your phone. *The Independent*. https://www.independent.co.uk/life-style/why-keep-checking-phone-psychology-smartphone-notifications-social-media-a7572916.html

Hume, D. (2003). *A treatise of human nature*. Everyman.

Insta Tours – Shoot Amazing Instagrams. (n.d.). Retrieved May 5, 2021, from https://insta.tours/

InstaBanff: Banff Instagram Tour. (n.d.). *Discover Banff Tours*. Retrieved May 5, 2021, from https://www.banfftours.com/activities/banff-instagram-tour/

Kaplan, A., & Haenlein, M. (2014). Collaborative projects (social media application): About Wikipedia, the free encyclopedia. *Business Horizons*, *57*(5), 617–626. https://doi.org/10.1016/j.bushor.2014.05.004

Kaviani, F., Robards, B., Young, K. L., & Koppel, S. (2020). Nomophobia: Is the Fear of Being without a Smartphone Associated with Problematic Use? *International Journal of Environmental Research and Public Health*, *17*(17). https://doi.org/10.3390/ijerph17176024

Kingwell, M. (2017). Bored, Addicted, or Both: How We Use Social Media Now. In *Social media and your brain: Web-based communication is changing how we think and express ourselves*. http://site.ebrary.com/id/11293205

Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511554476

Kross, E., Verduyn, P., Sheppes, G., Costello, C. K., Jonides, J., & Ybarra, O. (2021). Social Media and Well-Being: Pitfalls, Progress, and Next Steps. *Trends in Cognitive Sciences*, *25*(1), 55–66. https://doi.org/10.1016/j.tics.2020.10.005

Kuss, D. J., & Griffiths, M. D. (2011). Online Social Networking and Addiction—A Review of the Psychological Literature. *International Journal of Environmental Research and Public Health*, *8*(9), 3528–3552. https://doi.org/10.3390/ijerph8093528

Kuss, D. J., & Griffiths, M. D. (2017). Social Networking Sites and Addiction: Ten Lessons Learned. *International Journal of Environmental Research and Public Health*, *14*(3). https://doi.org/10.3390/ijerph14030311

Lanier, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. Henry Holt and Co.

Liu, C., & Ma, J.-L. (2019). Adult Attachment Orientations and Social Networking Site Addiction: The Mediating Effects of Online Social Support and the Fear of Missing Out. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02629

Marx, L., & Smith, M. R. (1994). Introduction. In *Does Technology Drive History?: The Dilemma of Technological Determinism*. MIT Press.

McDool, E., Powell, P., Roberts, J., & Taylor, K. (2016). *Social Media Use and Children's Wellbeing* (SSRN Scholarly Paper ID 2886783). Social Science Research Network. https://papers.ssrn.com/abstract=2886783

Morris, M. (2006). Akrasia in the "Protagoras" and the "Republic." *Phronesis*, *51*(3), 195–229.

Mun, I. B., & Kim, H. (2021). Influence of False Self-Presentation on Mental Health and Deleting Behavior on Instagram: The Mediating Role of Perceived Popularity. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.660484

Nations, U. (1948). *Universal Declaration of Human Rights*. United Nations; United Nations. https://www.un.org/en/about-us/universal-declaration-of-human-rights

Newport, C. (2020). *Digital Minimalism: Choosing a Focused Life in a Noisy World*. Penguin Business.

O'Neal Irwin, S. (2016). *Digital Media: Human–Technology Connection*.

Online social integration is associated with reduced mortality risk | PNAS. (n.d.). Retrieved May 7, 2021, from https://www.pnas.org/content/113/46/12980

Pennington, N. (2020). Quitting social media: A qualitative exploration of communication outcomes. *Qualitative Research Reports in Communication*, *0*(0), 1–9. https://doi.org/10.1080/17459435.2020.1817140

Pickard, H. (2019). Addiction. In *The Routledge Companion to Free Will* (pp. 454–467). Routledge & CRC Press. https://www.routledge.com/The-Routledge-Companion-to-Free-Will/Timpe-Griffith-Levy/p/book/9780367869977

Pickard, H. (2020a). Addiction and the self. *Noûs*, *n/a*(n/a). https://doi.org/10.1111/nous.12328

Pickard, H. (2020b). Addiction and the self. *Noûs*, 1–25. https://doi.org/10.1111/nous.12328

Plato. (2009). *Protagoras* (C. C. W. Taylor, Trans.; 1st edition). Oxford University Press.

Post, A. (2020). *The Fastest Growing Social Media Platforms in 2020*. Social Media Data. https://socialmediadata.com/fastest-growing-social-media-platforms/

Post, T. J. (n.d.). *Why putting away your phone before bed will help you live longer*. The Jakarta Post. Retrieved May 5, 2021, from https://www.thejakartapost.com/life/2019/11/06/why-putting-away-your-phone-before-bed-will-help-you-live-longer.html

Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, *29*(4), 1841–1848. https://doi.org/10.1016/j.chb.2013.02.014

Put the Phone Away! 3 Reasons Why Looking at It Before Bed Is a Bad Habit. (2019, April 22). Health Essentials from Cleveland Clinic. https://health.clevelandclinic.org/put-the-phone-away-3-reasons-why-looking-at-it-before-bed-is-a-bad-habit/

Robin. (n.d.). *Teens and Social Media: Constant Pressures And Social Media Stresses*. Retrieved May 7, 2021, from https://netsanity.net/teens-social-media/

Rozgonjuk, D., Pruunsild, P., Jürimäe, K., Schwarz, R.-J., & Aru, J. (2020). Instagram use frequency is associated with problematic smartphone use, but not with depression and anxiety symptom

severity. *Mobile Media & Communication*, *8*(3), 400–418. https://doi.org/10.1177/2050157920910190

Sabatini, F., & Sarracino, F. (2014). Online networks and subjective well-being. *ArXiv:1408.3550 [Cs]*. http://arxiv.org/abs/1408.3550

Schüll, N. D. (2014). *Addiction by Design: Machine Gambling in Las Vegas* (Reprint edizione). Princeton Univ Pr.

Shiffrin, S. V. (2000). Paternalism, Unconscionability Doctrine, and Accommodation. *Philosophy & Public Affairs*, *29*(3), 205–250.

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, *33*(1), 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245

Teens' night-time use of social media "risks harming mental health." (2015, September 10). The Guardian. http://www.theguardian.com/society/2015/sep/11/teens-social-media-night-risk-harm-mental-health-research

Turner, J. C. (1989). *Rediscovering the Social Group: A Self-Categorization Theory*. Blackwell Pub.

Verbeek, P. P. C. C. (2015). Cover story: Beyond Interaction: a short introduction to mediation theory. *Interactions (ACM)*, *22*(3), 26–31. https://doi.org/10.1145/2751314

Verbeek, P.-P. (2013). Resistance Is Futile: Toward a Non-Modern Democratization of Technology. *Techné: Research in Philosophy and Technology*, *17*(1), 72–92. https://doi.org/10.5840/techne20131715

Vollstädt-Klein, S., Loeber, S., Winter, S., Leménager, T., von der Goltz, C., Dinter, C., Koopmann, A., Wied, C., Winterer, G., & Kiefer, F. (2011). Attention Shift towards Smoking Cues Relates to Severity of Dependence, Smoking Behavior and Breath Carbon Monoxide. *European Addiction Research*, *17*(4), 217–224.

Wallace, R. J. (1999). Addiction as Defect of the Will: Some Philosophical Reflections. *Law and Philosophy*, *18*(6), 621–654. https://doi.org/10.2307/3505095

Warren, J. (2021, January 4). This is How The Instagram Algorithm Works in 2021. *Later Blog*. https://later.com/blog/how-instagram-algorithm-works/

Watson, G. (1975a). Free Agency. *Journal of Philosophy*, *72*(April), 205–220. https://doi.org/10.2307/2024703

Watson, G. (1975b, August 1). *Free Agency*. The Journal of Philosophy. https://doi.org/10.2307/2024703

Watson, G. (Ed.). (2003). Free Agency. In *Free Will* (2nd edition, pp. 337–351). Oxford University Press.

Watson, G. (2010). Disordered appetites: Addiction, compulsion, and dependence. In *Agency and Answerability: Selected Essays*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199272273.001.0001/acprof-9780199272273-chapter-4

What 'The Social Dilemma' Gets Wrong. (2020). https://about.fb.com/wp-content/uploads/2020/10/What-The-Social-Dilemma-Gets-Wrong.pdf

Why It's Time to Ditch the Phone Before Bed. (n.d.). Retrieved May 5, 2021, from https://www.sclhealth.org/blog/2019/09/why-it-is-time-to-ditch-the-phone-before-bed/

Wood, A. W. (1995). Exploitation. *Social Philosophy and Policy*, *12*(2), 136–158. https://doi.org/10.1017/s0265052500004702

Wyatt, S. (2008). *Technological Determinism is Dead; Long Live Technological Determinism* (pp. 165–180).

Youn, S. (2021, February 10). "I get better sleep": The people who quit social media. *The Guardian*. http://www.theguardian.com/lifeandstyle/2021/feb/10/people-who-quit-social-media

Zwolinski, M., & Wertheimer, A. (2017). Exploitation. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2017/entries/exploitation/