Voter Manipulation in the Cybernetic Age
Olierhoek, Alex

Citation

Olierhoek, A. (2021). *Voter Manipulation in the Cybernetic Age*.

Version:	Not Applicable (or Unknown)
License:	[License to inclusion and publication of a Bachelor or Master thesis in the Leiden University Student Repository](#)
Downloaded from:	https://hdl.handle.net/1887/3274303

Note: To cite this publication please use the final published version (if applicable).

Title: Voter Manipulation in the Cybernetic Age

Student: Olierhoek, A. (1284843)

Supervisor: Dr. Kalf, W. F.

Field of Study: Philosophy of Political Science

Submitted: 30-08-2021

Voter Manipulation in the Cybernetic Age

The subject of voter manipulation has gained a lot of attention recently. Involvement of foreign countries in democratic elections such as the Russian involvement in the election of Donald J. Trump in 2016 and attempts of foreign involvement in the European elections of 2018 showed the vulnerability of democratic systems to undesired influences. In addition, the abuse of big-data to influence voters, as was the case with the Cambridge Analytica scandal, and growing polarization originating on digital platforms has spurred a debate on how to approach voter manipulation and new technologies that exacerbate this in democratic societies.

This dissertation is an attempt to make sense of *voter manipulation*. This is manipulation with the goal to alter or bolster one's political preferences, the ultimate end of which is exerting control over votes and therewith influencing the democratic practice of voting.

In the first chapter, I will investigate the manipulation literature to come up with a definition of manipulation with can be used to identify specific instances of it under a common denominator, namely *unstraightforwardness*, which I take to be characteristic of all instances of manipulation.

Hereafter, in chapter two, I will show what makes voter manipulation specifically morally problematic for democratic systems. Voter manipulation undermines *individual autonomy*, which is the fundamental democratic value. This being the case, the outcome of a voting process tainted with voter manipulation fails to generate public justification for the coercive power of the state. By undermining individual autonomy, voter manipulation undermines democracy.

In the third chapter, I will argue that this undermining effect on democracy means that democratic citizens and the democratic state both have a duty to resist voter manipulation. For citizens, I will call this 'the duty of suspiciousness', which means that citizens ought to have an attitude of suspicion towards political information they receive and must critically reflect on the possibility that they are being manipulated when receiving such information to fulfill this duty.

The corresponding duty of the democratic state is 'the duty to alleviate suspicion'. The state ought to alleviate the suspicion of citizens *vis a vis* political information by giving citizens the means to detect manipulation and provide non-manipulative information alternatives if none are provided by the market.

In the final chapter, I will treat cybernetic technologies that are currently being utilized for voter manipulation or have the potential to be utilized in this way. In addition, I will give policy options to combat the voter manipulation (potentially) resulting from these technologies that democratic citizens and the democratic state may choose to fulfill their respective duties.

The aim of this dissertation then is, to build from the ground up, an account of voter manipulation that contains a moral appreciation of voter manipulation in democracies and apply it to contemporary developments in the voting process of democratic systems and therewith adding to the existing literature on manipulation in four ways.

First, the account of (voter) manipulation as *unstraightforwardness* addresses weaknesses in existing accounts of manipulation, specifically how to deal with manipulation that does not bypass reason[1] and where no deceitful intent is to be found in the manipulator[2].

Secondly, It will add to the debate on the desirability of nudging, where I will argue that there is no qualitative difference between nudging and manipulation when it comes to influencing political preferences in liberal democracies[3].

In addition, I develop an account of duties for liberal democratic states and their citizens that

---

[1] See Chapter 1. Section II, where I treat Gorin's (2014) critique of *bypassing reason* accounts of manipulation and why this critique does not apply to the specific case of *voter manipulation.*

[2] See Chapter 1 Section II, specifically the remark on Whitfeld (2020) and the closing remarks on agency.

[3] See Chapter 2 Section I.

is related to the specific wrongness of voter manipulation in liberal democracies and necessary for maintaining healthy liberal democracies in the cybernetic age[4].

Lastly, I come up with novel policy proposals that are designed to tackle the most significant problems related to voter manipulation in according with these duties[5].

Chapter I: Identifying Voter Manipulation

In an illuminating article Gregory Whitfield (2020, p.2) analyzes *political manipulation* because *"much liberal-democratic political thought has concerned itself primarily – even exclusively – with coercive influences in citizen's lives"* ignoring a *"wide range of cases not well described as coercion, persuasion,, interference or alternative takes on manipulation"*. Whitfield (2020, p.3) argues that these cases, being morally disapprobative to some extent, and concerning the public, *"call for justification in much the same way that coercion does"*. I am in agreement with Whitfield in holding to this view. Whereas Whitfield (2020) has an ambitious scope for his approach, conceptualizing political behavior in the broad sense, covering instances of gerrymandering, political alliances and strategic voting, my goal in this thesis will be somewhat narrower. I will limit myself to a specific form of political manipulation captured in the concept *voter manipulation.* I take *voter manipulation* to be an instance of manipulation that aims to alter or bolster one's political preferences therewith exerting control over the outcome of one's vote. Although more modest in scope, there exists a wide range of manipulative acts that may influence someone's political preferences. From peer-pressure, emotional blackmailing and charming, to guilt-tripping and deceit, manipulation is a many-headed beast, and although no new head will grow when we chop one off, instead of covering each instance individually it is prudent to turn to the manipulation literature and see what these seemingly disparate acts might have in common. This will be the objective of the first section of this chapter. In the second section we will use the insights gained from the literature in the first section, to argue for a conception of *manipulation* and *voter manipulation* specifically.

Chapter I Section 1 : Conceptual Analysis of Manipulation

The earliest in-depth treatment of manipulation stems from medical ethics and treats manipulation as a form of *pressure*. Often in the situation where a doctor applies pressure to a patient to make the patient agree to a certain treatment the patient would not take otherwise. The authors (Faden, et al., 1986, p.258) argue that there are in-between cases on a pressure spectrum with coercion on one pole (i.e. overwhelming pressure), and rational persuasion (i.e. no pressure at all) on the other. Manipulative acts, according to them, are all acts that fall between these poles. Joel Feinberg (1989, p. 37) agrees with this assessment, stating that "*the line between forcing to act and merely getting to act is drawn somewhere in the manipulation or persuasion part of the scale*". This sentiment is echoed by Kligman and Culver (1992, p.178) who claim that "*the attempt to influence one's behavior takes on a manipulative character when*" normal persuasion fails in generating assent and one "*procures or engineers the needed assent by brining pressure to bear […]*". The conception of manipulation as a form of pressure is insightful and definitely covers quite some instances of manipulative acts, however, intuitively we would describe more subtle acts that do not apply pressure, such as deception, as instances of manipulation too. The doctor-patient cases often found in medical ethics, where the pressure account of manipulation originates, fails to describe those instances of acts we would

---

[4] See Chapter 3 Section I & II.
[5] See Chapter 4.

intuitively describe as being manipulative that do not use the application of pressure[6]. The doctor never uses deceit in these examples, never tricks the patient. She pressures the patient into taking a certain decision through her position as doctor, as authority that is being authoritative. The idea, put forward by Faden et al. (1986, p. 261), that manipulation violates the conditions for informed consent, although a bit narrow in their own application, nevertheless captures an important characteristic of manipulation that we will treat in the second section of chapter II.

I have already hinted that the manipulation as *pressure* account fails to capture deceptive manipulative acts. In these instances of manipulation, no pressure is applied at all and these are described as manipulation as *trickery*. The conceptualization of manipulation as trickery suffers from the inverse deficiency of the pressure accounts in that they fail to include pressuring acts as manipulative acts. The thought behind *manipulation as trickery*, as put forward by Claudia Mills (1995, p.101) is that the *"manipulator tries to change another's beliefs and desires by offering her bad reasons, disguised as good, or faulty arguments"*. The manipulator thus generates a faulty mental state in the manipulated, resulting in an act or judgment that is in accordance with the desire of the manipulator. This covers the deceptive part of manipulation, and Mills' insight (Ibid. p. 97) that manipulation aims at altering "*certain internal or subjective features of one's choice situation"*, rather than changing the "*external and objective features"*, as is the case with coercion neatly distinguishes these instances of influencing behavior.

The omissions of the trickery and pressure accounts have led Felicia Ackerman (1995, p.337) to claim that the concept of manipulation suffers from "*combinatorial vagueness*". Similarly Cass Sunstein (2016, p.81) doubts the feasibility of capturing manipulation in a single coherent concepts for which necessary and sufficient conditions can be given. Some authors, such as Joel Rudinow (1978, 341) seem to have foreseen this difficulty and proposed hybrid conceptions where both pressure and trickery are included in the concept, avoiding the hardships of conceptualizing manipulation altogether. However, these accounts fail to capture a similarity between the *pressure* and *trickery* account that gives us important insights in how manipulation, as a single concept, operates.

One such commonality between the *pressure* and *trickery* accounts that has gained much traction, is the idea that a manipulative act always *bypasses reason*[7]. Manipulation as *bypassing reason* succeeds in capturing both the use of *trickery* and *pressure* as instances where the capacity to adequately use one's reason is omitted by the introduction of non-rational influences in the thought process. These non-rational influences, according to Joseph Raz (1988, p. 117) *"pervert the way that a person reaches decisions, forms preferences or adopts goals".* Both pressure and trickery can be thought to have this *perverting* effect on one's reason. Although promising, there are two difficulties with this conception of manipulation. First, it may be too broad as non-rational influences that bypass reason can come in many forms, not all of which we would find manipulative or in any sense immoral[8] and we would rather conceptualize as instances of *nudging*. Imagine the well-known fly in the urinal that makes men urinate in a more controlled and less messy fashion, or presenting healthy food on the top of a shelf in a cafeteria to direct people towards healthier decisions. In addition, there seem to be at least one convincing case of manipulation where the manipulator uses rational influences to get his or her way and does not bypass reason at all. In this case, put forward by Gorin (2014, p. 56), an atheist wishes the death of a Christian person and knows that if this person stops believing in God, he will find life worthless and kill himself. The atheist proceeds to give the Christian reasonable arguments

---

[6] e.g.: Making someone doubt about one's own judgement (i.e. gas lighting) or feeding someone with false information to get a pre-determined action or judgment.

[7] The *bypassing reason* account is identical to the *non-rational influence* account that can be found in Wood, A. (2014). Coercion, Manipulation, Exploitation. In C. Coons & M. Weber (eds.), *Manipulation: Theory and Practice*. Oxford; New York: Oxford University Press.

[8] This last point is unproblematic if we hold the view that not all instances of manipulation are morally wrong.

that counter the existence of God. The Christian becomes convinced of the truth of the atheist argument and proceeds to kill himself. Contra to Mills (1995, p.101) account where manipulation offers *bad reasons or faulty arguments,* or Raz's account where it bypasses reason, no such thing happens in Gorin's case. The atheist gives perfectly reasonable and scientifically grounded arguments and appeals to the Christians' sense of reason rather than bypassing it. Intuitively however, we would think this act is an instance of manipulation as the atheist is not sincere about his intentions and is, as Mills (ibid. p.101-102) would put it "*interested in reasons not as logical justifiers but as causal levers".* Withstanding this critique, an important upshot of manipulation as bypassing reason is its gradual move away from more particular instances of manipulation, as in the trickery and pressure conception, towards a broader definition that is convincing and works for almost all and only cases of manipulation.

Another attempt worth mentioning is Robert Noggle's (1996, p.45) conceptualization of manipulation as an act that "*causes a response in the influenced that he or she deems non-ideal".* Herewith, Noggle seems to escape the difficulties of trying to find commonalities between all instances of the manipulation on the basis of the means employed by manipulators, such as the *bypassing of reason, trickery* and *pressure*, by defining manipulation purely on the basis of its result. However, he misses the fact that a non-ideal response may also result from coercion, rendering his account to capture instances of manipulation only. An important upshot of Noggle's conception however, is its ability to distinguish between manipulation and nudging, as the latter may be seen as an influence on one's behavior that is ultimately deemed ideal[9]. Noggle's conception can ultimately be criticized for the same reason as Raz's *bypassing reason* account. An ideal response seems to presuppose that one would reasonably desire this response and this seems to imply that Noggle's account is another example of a *bypassing reason* account. However, it may have some teeth against Gorin's Christian versus atheist case, as arguably the response of the Christian to the atheist argument was ultimately non-ideal[10], meaning it could be identified as an instance of manipulation.

As we have seen, there is no definitive concept of manipulation out there. Combining the variety of manipulative acts, distinguished in the *trickery* and *pressure* accounts, under the heading of *bypassing reason*, still seems to fall short of covering all instances of manipulation[11] while manipulation surely cannot be reduced to either *trickery* or *pressure*. Noggle's attempt to define manipulation by its consequences rather than the means employed blurs the line between coercion and manipulation which should be avoided for any clear account of manipulation. However, this is no reason to give up on these insights. By salvaging what can be used from the above attempts, I think it possible to come up with a satisfying concept for manipulation that is suitably abstract to capture all instances of manipulation we are concerned with. We then only need to add the aim of altering or bolstering one's political preferences to make the concept specific to influencing acts that target *voters*.

---

[9] Such as urinating on the picture of a fly to avoid the chances for a messy urinal. Note, that the ideal consequences may also be reached through employing rational persuasion rather than nudging. In this sense the account of nudging in Noggle's work suffers from a similar deficiency as his account of manipulation, namely it being too broad.

[10] For a paper, in which it is argued that Noggle's account can fend off Gorin's case, see: Bělohrad, Radim. (2019). The Nature and Moral Status of Manipulation. *Acta Analytica*. 34. 1-16.

[11] Namely those that do not tend to bypass reason.

The conception of manipulation I argue for is:

> An act of manipulation is any intentional attempt by A to *unstraigthforwardly* cause B to will/prefer/intend/act in accordance with A's aims.

The conception of voter manipulation in accordance with the above conception is as follows:

> An act of voter manipulation is any intentional attempt by A to *unstraightforwardly* cause B to alter or bolster B's political preference in accordance with A's aim.

The conception of (voter) manipulation as *unstraightforwardness* I will argue, captures *all* relevant instances of manipulation, albeit too broad to capture *only* relevant instances of manipulation. With this I mean this account captures instances of nudging as well that some may hold are qualitatively different from manipulation. I will argue against the qualitative different between the two, in the first section of chapter II.

In this section, I will explain how to interpret *unstraightforwardness* and why it suffices for our conception. Thereafter, I will comment on other elements of the concept such as i*ntention*, the *status* of A and B[12], and the inclusion of the term *bolster* in addition to *altering* political preferences, which differs from most accounts of manipulation that take manipulation to be about altering the decision-making processes, not maintaining it. However, I will begin by showing how I arrived at the above conception of voter manipulation.

Manipulation as we have seen, is an influencing act that is often described in terms of its differences with rational persuasion and coercion[13]. I take the reason for this to be that both rational persuasion and coercion are comparably clear terms. Granted, there are different ways in which one can go about to rationally persuade someone. However, it always entails targeting one's rational faculties with reasons that allow someone to weight the costs and benefits of a certain decision. Coercion can take multiple forms as well, yet always targets one's *"external and objective features"* (Mill, 1995, p.97). In other words, if one has to choose between two options, rational persuasion influences someone by giving epistemically good reasons for one decision over the other, while coercion influences someone by making one decision outweigh the other, ultimately leaving someone with only one rational decision to make.

The distinction between rational persuasion and coercion relates respectively to the *manipulation as trickery* account, that takes manipulation only as a perversion of rational persuasion, and the *manipulation as pressure* account, that only treats manipulation as a mild(er) form of coercion. For a complete account of voter manipulation, we will need to combine insights from both accounts as both methods of manipulation can be employed to influence voters[14].

Manipulation as *bypassing reason* seems promising in this regard, however, assuming Gorin (2014) is right for the most part, this would leave out those instances of manipulation from our account where the reasonable faculties of the manipulated are employed for deceitful purposes. However, I think his argument for manipulation that does not bypass reason, while applying to manipulation proper, does not apply to the specific instance of voter manipulation, as I will argue shortly.

What all treated accounts of manipulation have in common, is that the manipulative act never shows the manipulators intent in a direct manner; A does not make it clear to B what aims underpin

---

[12] We will see this is particularly important for a concept of manipulation that works for cybernetic technology.
[13] See e.g. Rudinow (1978), Sunstein (2016), Whitfield (2020),
[14] Voters can, for example, be subjected to peer pressure from family to vote a certain way or may be the victim of a misinformation campaign that gives them epistemically bad reasons to base their vote upon, as we have seen in the promises of the Brexit campaign or the ploy of the Trump campaign to criminalize Hillary Clinton.

the influencing act towards B. Where with rational persuasion the aims of the persuader are clear as he or she straightforwardly ventures to alter the decision making process of the persuaded, and with coercion too, as there is no doubt that the coercive act aims at making the coerced act in accordance with the coercer's aims, there is no such straightforward relation to be had for the manipulated on the one side, and the manipulator and his or her aims at the other.

Manipulation as unstraightforwardness easily distinguishes those manipulative acts that fall under *manipulation as trickery* from *rational persuasion*, as these deceptive instances, such as dressing bad reasons as good reasons scream *unstraighforwardness*. However, how can it be the case that a manipulative act can take the form of mild(er) coercion but lose the straightforwardness just ascribed to coercion? I will use an example to argue why this is the case.

Imagine a situation where A *threatens* B: "*You either vote for candidate D or I'll make sure your friends hate you*", and this threat works as under no circumstances B wants her friends to hate her. It is obvious here that the relation of B with her friends is only of concern to A as a lever to necessitate B to vote for candidate D. A's intent, to make B vote on D, is clear to B as A *straightforwardly* addresses her own aims to B. A's threat leaves no possibility for speculation about the aims she has for threatening B.

Now take a situation where A *pressures* B that "*if you don't vote for candidate D your friends will hate you*". It is not clear in this case whether A tells this to B out of concern for her relations with her friends, to neutrally inform A about the consequences of her action, or to get B to vote candidate D. As it stands, B may interpret A's utterance as *pressuring,* and it surely applies pressure, but B may interpret it in other ways as well. The intent of A's *pressure* remains obscure to B as A does not straightforwardly tell B that she wants her to vote for candidate D.

This is akin to Whitfield's (2020, p.4) account of manipulation, who conceptualizes it as a method that "*obscures and renders deniable*" the manipulator's intent *vis a vis* the manipulated. However, on my account the manipulating act does not necessarily entail a deceitful intent, it suffices that the intent is not apparent to the manipulated[15].

I think the above example works for showing that there is one important distinction between coercion proper and *pressuring* manipulation. The latter does not violate the conditions of free choice. This means that *pressuring* manipulation can come close to *coercion* proper but is never strong enough to *force* someone to act one way or the other. As soon as an influencing act pressures in a way where the influencer actively participates in making the content of the pressure happen, in the above case A's promise that *B's friends will hate B*, the influencing act becomes coercive rather than manipulative.

This being the case, the manipulator can never straightforwardly apply pressure while conceptually remaining outside the boundaries of coercion. In other words, coercion is straightforward, it does not obscure the intent of the coercer and the coerced knows to what end she is coerced.

With applying pressure, the intent of the one who pressures remains somewhat obscure. We can apply pressure by simply naming a possibility or directing one's thoughts but never by actively intervening in a way that ensures that the content of that with which one is pressured, becomes reality. *Pressuring* does not guarantee a certain outcome, while coercion relies on the coercive party to be able to guarantee a certain outcome, as this makes for a valid threat. Granted, it remains vague where the line between coercion and *pressuring manipulation* is crossed as it is empirically impossible to know for certain if someone's free choice is sufficiently violated to conclusively speak of coercion rather than pressuring. Even the same sentence uttered in a different context or with a different intonation can make the difference between interpreting *pressuring manipulation* as a *threat* or *vice*

---

[15] This is an important distinction to make the account of manipulation applicable to cybernetic technologies, as we will see in Chapter 4.

versa[16]. This being said, I take it that *unstraightforwardness*, as a concept, sufficiently distinguishes between coercion and *pressuring manipulation*, and will leave the complicated problem of empirically distinguishing between the two to others[17].

One may ask, how does the *manipulation* as *unstraightforwardness* account relate to or differ from earlier accounts of manipulation? To this my answer is that manipulation as *unstraightforwardness* combines the *trickery* and *pressure* accounts in one neat concept, thus avoiding the duality of the hybrid concepts. This is no novelty as it is also the aim of the *bypassing reason* account. However, as we have seen, the *bypassing reason* account, suffers from one weakness, namely that, although most instances of manipulation seem to bypass reason, some, think of Gorin's example of the atheist and the Christian, clearly do not. The account of manipulation as *unstraightforwardness* withstands Gorin's critique, as it does not identify the manipulative act by weighing the content of the manipulative act, which does not *bypass reason*, but by looking at the intent of the manipulator. It focusses not on the content of what the atheist says to the Christian, which is reasonable and grounded in science, but on the hidden intent the atheist has for convincing the Christian of the non-existence of God. This inherent hiddenness of the intent of the manipulator, the fact that the manipulator can always lie, deceive and cover-up his or her intent, is the specific quality of manipulation that is identified with the manipulation as *unstraightforwardness* account, and which I think to be common to all accounts of manipulation, even those that do not bypass reason, and this ultimately sets manipulation apart from rational persuasion and coercion.

As promised I will briefly comment on the term *intentional* and the status of A and B. Intention has an important moral connotation, and as I will argue later on in this dissertation[18], it is by intent that we can distinguish between *causal responsibility* and *moral responsibility,* where the latter form of responsibility requires intent. Intent is somewhat connected to the status of A, while I take B to always be a voter as we will treat *voter manipulation*, the status of A is less clear. All above authors on manipulation hold implicitly or explicitly that, the manipulator, in our case A, is an agent, thus possessing agency, which presupposes a certain capability towards autonomous behavior. In this dissertation, the agency of the manipulator is not taken for granted as we will ultimately be concerned with the influence of cybernetic technologies, such as algorithms, on voting behavior. This means that we have to be open to the idea that an entity without agency might be able to manipulate B. For this reason I have left the agency status of A out of the concept.

Although agency is left out of the concept, *intent* has a place in it for two reasons. The first (1) is that intentionality and agency can possibly be decoupled with cybernetic technology, where the intent of the creator of the technology on the one hand, and the self-learning properties of the technologies on the other, make it so that the influencing act lacks agen*cy* but has intent[19]. Whether I think this can amount to manipulation I will touch upon later, for now it suffices to note that I think

---

[16] Take the sentence: "If you go there, an accident might happen", it can be *pressuring* in the sense that a worried friend says it to scare you from going to a place she considers dangerous. However, if uttered by an angry mobster one can take it as *threat* and it may be interpreted as *coercive*. This shows we cannot take manipulation on the value of words only but need a complete context to judge whether an act is manipulative or not.

[17] For those not convinced by this argument, it suffices to adjust the concept of manipulation in the following way to make it sensitive to the difference between coercion and manipulation: An act of manipulation is any intentional attempt by A to *unstraightforwardly* cause B to will/prefer/intend/act in accordance with A's aims without violating B's freedom of choice. I do not deem this necessary as I think no conceptual clarity is lost with the original, shorter version where this addition is taken to be included in the term *unstraightforwardly*:.

[18] See: Chapter II Section 2a.

[19] For example, an algorithm may have some goal to work towards, some programmed *intent,* while being unable to act autonomously, outside the pre-programmed boundaries.

this to depend on how these technologies will evolve in the future[20]. In addition, (2) I want this concept to be able to make sense of the moral status of *voter manipulation* and I will argue in the next section that the intentionality of the act plays a decisive role.

The reason for adding 'to alter or bolster B's political preferences' to specify *voter manipulation* from the concept of manipulation is because I think it is possible to not only manipulate someone to alter his or her political preferences, but also to make someone not prone to doubting his or her political preferences. If, for example, a fictional news broadcasting agency, call it 'WOLF news' targets its viewers with misinformation in favor of a political preference the viewer already has, I take this to possibly be an instance of *voter manipulation* that bolsters ones political preferences.

In the first section of this chapter we have turned to the manipulation literature to look for a clear concept for (voter) manipulation. I have discussed important difficulties that arise when trying to come up with a clear concept of manipulation. By discussing these difficulties and enquiring into the shortcomings of some conceptions of manipulation, I have tried, in the second section, to come up with a conception for manipulation that suffers none of the weaknesses of the treated accounts[21]. Furthermore, the aim of this dissertation is not only to identify (voter) manipulation but to gain clarity as to how to deal with increasingly complex voter manipulation, through cybernetic technologies, in democratic systems. To know how we *ought* to deal with manipulation, we first need to evaluate (voter) manipulation. This will be the aim of the next chapter in which I will treat the moral status of (voter) manipulation in democratic systems.

Chapter II: Evaluating (Voter) Manipulation

In the previous chapter, I have argued that we may identify an influencing act as rational persuasive, coercive or manipulative by looking at the *straightforwardness* of the act, where the former instances of influencing are *straightforward* and manipulation is characterized by its inherent *unstraightforwardness*. With this, I mean that the manipulator does not show the intent of his or her influencing act, leaving the influenced ignorant of the true motives of the manipulator.

In the first section of this chapter, I will argue why not being able to distinguish between manipulation and nudging is inherent to my account of manipulation and why this is unproblematic for our current purposes. I will do this by arguing that from a democratic perspective, there is no qualitative distinction between nudging and manipulation.

Hereafter, in the second section, I will argue for the specific *wrongness* of *voter manipulation* in democratic societies. The negative moral status of *voter manipulation* in this context stems from the diminishing effect *voter manipulation* has on the autonomy of the manipulated. The *unstraightforwardness* inherent in *voter manipulation* infringes upon the *procedural independence* necessary for citizens to make *autonomous* decisions on the vote. *Procedural independence* means that one makes a decision with which, after critical reflection and under conditions of full information, one would agree. Seemingly paradoxical, one's procedural independence is compatible with being influenced under certain conditions, namely that, if, after fully knowing the extend and intent of the influencing act, the influenced person would still wholeheartedly make the same decision. In other words, if one can identify with a decision knowing how the decision came to be, what the genesis of this decision is, one is taken to be procedurally independent. The genesis of a certain preference is

---

[20] Will there, for example, be autonomous artificial intelligence in the future, and if so, lies the intent, if any, with the AI or with its creator or even with some other institution that decided upon the (ethical) boundaries of the AI?

[21] Keep in mind that I will argue why the inability to distinguish between manipulation and nudging will be unproblematic for our purposes in the first section of chapter II.

clear when one is being coerced or rationally persuaded. However, manipulation, through its quality of being *unstraightforward*, blocks the manipulated from knowing the information needed for *procedural independence*, thus diminishing the autonomy of the victim for which *procedural independence* is a necessary condition. Thereafter I will argue that a voting act that results from manipulation fails, through a lack of procedural independence, to properly justify the coercive power of the state. To give a justification for the laws, all citizens have to autonomously support the laws and have to be able to identify with the laws. This means that voter manipulation violates the conditions for the public justification of coercive state power.

Chapter II Section 1: Voter Manipulation & Nudging

The reason why my conception of manipulation does not *only* capture instances of manipulation, stems from the moral neutrality with which I have treated manipulation thus far. This has resulted in a concept that also captures influencing acts that may be better defined as instances of *nudging*. This is unproblematic as I will bite the bullet and claim there is no moral difference between manipulation and nudging[22]; in the case of *voter manipulation*.

The concept of *manipulation* may be made sensitive to the difference between *manipulation* and *nudging* by adding a moral status to it. Cas Sunstein and Richard Thaler (2009, p.151) conceptualize *nudging* as a *"deliberate introduction of subtle, non-coercive influences into people's decision-making to get them to make more optimal choices"*. Remember, that Noggle (1996) argued that manipulation results in a *non-ideal response.* If A manipulates B to get the non-ideal response, this manipulation seems to be *pro tanto* wrong as B would have fared better with the outcome where A had not interfered and thus A ought not to have interfered in B's dealings. A nudge on the other hand, as a stimulant towards an ideal response, seems to be *pro tanto* right as it steers B to behavior (s)he ought to desire. Nudging, according to the above conception, is a good thing, while our intuition about manipulation runs in the opposite direction[23].

The view that manipulation includes an element of moral disapprobation is echoed by the authors we have treated in the first section. Raz (1988, p.378), holds that manipulation "*perverts*" our decision-making process, Mills (1995, p.99-100), claims manipulation gives us "*bad reasons*", going as far to compare it to the "*morally unacceptable act"* of drugging someone to make this person change his or her mind, and Faden, et al. (1986), argue it violates the conditions for informed consent which constitutes a wrong in the doctor-patient relationship. I have sympathy for the view that manipulation is *pro tanto wrong*. It fits our intuition, the common usage of the word, and it distinguishes manipulation from *nudging* in a simple and effective manner. Despite the intuitive appeal of this view and it's practical implications, however, the *pro tanto wrongness* of manipulation is a matter of debate[24], as is the qualitative distinction between *nudging* and *manipulating*[25] and no consensus has

---

[22] Those who are convinced there is no qualitative difference may ignore the claim that the moral neutral concept is too broad.

[23] This is why I claimed in the first section that Noggle succeeds in distinguishing *nudging* from *manipulation.*

[24] Some authors argue that at least some forms of manipulation are praiseworthy, for their argument see: Dowding, K., & Van Hees, M. (2008). In Praise of Manipulation. *British Journal of Political Science, 38*(1), 1-15.

[25] See for an overview of the debate amongst others: Bovens, L. (2009). The Ethics of Nudge. In T. Grüne-Yanoff & S.O. Hansson (eds), *Preference Change: Approaches from philosophy, Economics and Psychology*. 207-219 Dordrecht: Springer Netherlands, Hausman, D.M., & Welch, B. (2010). Debate: To Nudge or Not to Nudge. *Journal of Political Philosophy, 18*(1), 123-126, Noggle, R. (2018) Manipulation, Salience and Nudges. *Bioethics, 32*(3), 164-170, Nys, T. R.,& Engelen, B. (2017), Judging Nudging: Answering the Manipulation Objection. *Political Studies, 65*(1), 199-214 and Reach, G. (2016) Patient education, nudge and manipulation: Defining the ethical conditions of the person-centered model of care. *Patient Preference and Adherence, 10*, 459-468.

been reached as of yet.

In this dissertation, I will contribute to this debate by looking at the specific practice of voter manipulation in democracies. Quite different from earlier approaches, I will argue that the *wrongness* of *voter manipulation* depends on it violating the *procedural independence* of the victim. This means that I will identify the *wrongness* of manipulation independent of any (ideal or non-ideal) outcomes of the influencing act. Because both *nudging* and *manipulation* depend on *unstraightforwardness*, and both diminish the *individual autonomy* of citizens when deciding on their vote, I hold there is no qualitative difference between *nudging* and *manipulation* as both instances of influencing violate the core democratic value of *autonomy*. My approach focusses on the procedure of voting, the act through which citizens in democracies make their political preferences know. This focus on process means that the outcome of the manipulative act is of lesser concern to us in this dissertation. This is the result of treating the subject of voter manipulation as a problem of democratic societies through it violating democratic values, as I will argue in section 2 of this chapter.

Obviously, the consequences of a manipulative or nudging act do matter, as, even when we conclude that manipulation is *wrong*, it may still be desirable to manipulate voters if the consequences of non-manipulation are worse[26]. Furthermore, one *wrong* consequence of manipulation is always attached in the case of voter manipulation, namely that it results in the inability of citizens to justify coercive state power. This consequence follows necessarily from a voting process in which the autonomy of voters is undermined, as will be argued in the following section.

To duly weigh the act of manipulation against the consequences, we ought to evaluate manipulation on itself first and lay the groundwork for further research into the relation between the *pro tanto wrongness* of manipulation and the consequences. This means that for this dissertation, I take it that procedurally, both *nudging* and *manipulation* employ *unstraightforwardness* and thus are not qualitatively different instances of influencing. I note again, that the emphasis on procedure stems from the specific democratic perspective I argue from in this thesis. Would one not take the value of *individual autonomy* to be of fundamental importance, as it is for democratic theory and the voting process, differentiating between *nudging* and *manipulation* in the voter context, thus focusing on the outcomes of influencing acts rather than the procedure involved, might be relevant. This being the case as the importance of procedural independence, and the resulting democratic legitimacy of the procedure, would be smaller or even non-existent[27][28]. That being said, in those systems where this is the case, the democratic value of voting is already subordinated to non-democratic rule of power elites, rendering a treatment of voter manipulation for those systems less relevant as the individual vote does not translate into political influence for individual citizens but symbolically to legitimize autocratic rule.

---

[26] If for example voters are manipulated to stop a totalitarian criminal from taking control of a country, the *wrongness* of the manipulative act may well be outweighed by the consequence of keeping democratic rule.

[27] Take as an example contemporary China, where *individual autonomy* is hardly valued at all and manipulation of political preferences to maintain order is an accepted government practice. Here, obviously the outcome rather than the procedure of the manipulative act is taken to be the justifying feature of the manipulative act.

[28] For a discussion of the view I am in agreement with, namely that manipulation accounts based on outcomes are incompatible with democratic legitimacy (which stems from procedure) see: Paulo, N., & Bublitz, C. (2016). Pow(d)er to the People? Voter Manipulation, Legitimacy, and the Relevance of Moral Psychology for Democratic Theory. *Neuroethics*, *12*(1), 55–71

In the introduction to this chapter, I claimed that the *unstraightfowardness* inherent in *voter manipulation* diminishes the *autonomy* of the manipulated. Before arguing how this comes about and why *voter manipulation* is always wrong in democracies, it will be crucial to give an account of moral responsibility, autonomy and public justification. *Moral responsibility*, to make clear how my evaluation of manipulation is grounded, *autonomy* as it is the fundamental democratic value, thus making our evaluation of manipulation specific to instances of *voter manipulation,* and *public justification* as the failure to arrive at a *public justification* of coercive state power is the negative moral consequence of *voter manipulation.* This will be treated in part a of this section. In part b, I will tie moral responsibility autonomy and public justification to the practice of *voter manipulation.*

a: Moral Responsibility, Autonomy & Public Justification

We can distinguish between two relevant types of responsibility, namely, *causal* and *moral* responsibility. Only the latter form of responsibility has an effect on the *moral status* of one's act. However, it is a necessary condition for bearing *moral* responsibility that one bears *causal* responsibility too. Simply put, one cannot be held *morally responsible* for an act one did not bring about or let happen[29].To be *morally* responsible then, is to have some added responsibility to one's *causal responsibility*. This added responsibility is generated by one's understanding of the *moral status* of one's *act* or the *effects* the act would generate. If one lacks this moral understanding for non-culpable reasons[30] , one cannot be held *morally responsible* for an action while one may bear obvious *causal responsibility* for acting thus*.* For someone to be properly held responsible for a *wrong* then, one has acted in spite of knowing the *wrongness* or while having had to know the *wrongness* of the act. In other words, there was either intent or a unwarranted deficiency in one's *moral knowledge*.

On the above account, *moral responsibility* amounts to *causal responsibility + (culpable lack of) moral knowledge.* This account allows us to make sense of someone's *moral responsibility*. In accordance with Frankfurt's (1969) views[31], it sees moral responsibility as resulting from our second-order volition, namely our capacity to critically reflect on our first-order desires and (try to) act accordingly. If we would only be able to follow our first-order desires, we would merely be *causally responsible* for our actions, according to Frankfurt, being not different from ants in lacking the capacity for *moral reasoning*. It is exactly our capacity for critical reflection on our first-order preferences that makes us persons in the sense that we are moral agents bearing moral responsibility for our actions.

This account of moral responsibility, to be useful, must be substantiated with values. Adding values will allow us to decide whether someone ought to have *known (or not)* the *moral status* of his or her action and ought to have critically reflected on it. Ultimately, our practices ought to be compatible with these values and *moral reasoning* is aimed at making our actions accord to those values. For our current purposes, these values are readily available in the form of the liberal democratic values that underpin voter practice. We may evaluate the *moral status* of *voter manipulation* for democracies by invoking these associated values.

---

[29] Being passive can be culpable in this view.

[30] A doctor who prescribes a certain drug to a diabetic without knowing the side-effects, resulting in the patient dying because one of the side effects triggered a lethal response specific to the diabetic condition is an example of culpable ignorance.

[31] As we are concerned with responsibility in this thesis, I assume at least a compatibilist view on moral responsibility as any further move towards determinism makes a discussion on responsibility irrelevant. In addition, we assume the existence of free will and corresponding responsibility in the daily practices of liberal democracies such as the legal system.

Two fundamental and interrelated values of democratic practice are *autonomy* and *public justification*. Liberal democratic practice is based on the idea that all humans, being *free* and *equal*, have a right to *self-government*. This right to *self-government* in liberal democracies is based on the capacity of citizens to make *autonomous* decisions based on one's own interests and/ or preferences. The *autonomy* of citizens plays a key role in liberal democratic theory not in the last part for justifying and legitimizing coercive state power. As authors of our own lives, liberal thinkers agree that, as Berlin (1969, p.131)[32] aptly puts we "*wish to be a subject not an object… deciding, not being decided for, self-directed and not acted upon by external nature or by other men as if I were a thing, or an animal, or a slave incapable of playing a human role, that is, of conceiving goals and polices of my own and realizing them*".

From the early modern period onward, liberal political philosophers[33] have argued for the compatibility of *individual autonomy* and state power. Contract theorist such as Thomas Hobbes (2005, p.87) and Benedictus de Spinoza (2007, p.227) argue that citizens *voluntarily* lay down their "*right to everything*" for the better life in the state, and the voluntary nature of this bargain in fact legitimizes state coercion. Hobbes going so far as to say that when the state punishes someone, the punishment actually derives from the will of the punished. Jean-Jacques Rousseau (2014, p. 55) holds that anyone who does not comply with the outcome of the democratic procedure should be "*forced to be free*", essentially buying into the Hobbesian view that state power derives from one's own will. Immanuel Kant (1991), argues that obedience to the state is a moral maxim, therewith claiming that obedience to the laws is necessarily an *autonomous* act, by obeying the state, we are essentially obeying ourselves. John Rawls (1971, p.576), holds that we may act autonomously when we "*act from principles that we would consent to as free and equal rational beings"* arguing for the compatibility of *autonomy* and *state coercion* on the basis of collective consent to his principles of justice. Democratic theorist over centuries have tied *individual autonomy* to *public justification*, arguing that preserving the former may result in the latter. When the *autonomy* of individuals is infringed upon, *public justification* becomes shaky and therewith the justification of state power so valued by liberal democratic thinkers.

For this reason it is held by liberal democratic thinkers that democratic systems are best able to justify state power as the *autonomy* of citizens is *procedurally imbedded* in the right of citizens to vote, which in fact is the democratic way of practicing self-government. The word *autonomous* stems from the ancient Greek *autós*, "self", and *nómos*, "law". To be autonomous then, is to give oneself laws to life by, to govern oneself. To be autonomous in a political body is to be able to actively participate in the process of making laws as one would if such a political body was lacking. The practice of voting in democracies then, is nothing other than the practice of self-government in a collective. The right to vote, to act out one's right to self-government, is only granted to citizens who are taken to be *autonomous*. Those humans who are thought to be able to make *procedurally independent* decisions by virtue of not being (too) dependent on some other person[34]. In addition, to not violate the innate equality of democratic citizens which is derived from the idea that everyone, being author over their own life, ought to have an equal say in collective issues, every citizen receives one vote in elections. Call this the *one person one vote principle* wherewith every citizens *individual autonomy* is accorded equal weight in political matters.

If all citizens that are able to act out their autonomy can do so by participating in the body

---

[32] See: p. 131: I, Berlin. (1969) *Four Essays on Liberty*. Oxford University Press: Oxford.
[33] Note that the scale of the democratic process has increase over time has changed. The democratic element in Hobbesian political philosophy rests only with the genesis of the social contract whereas later thinkers extend the reach of the democratic procedure to cover more elements of the state . However, the democratic procedure, and thus autonomy, is also the basis for the legitimization of state power in Hobbes.
[34] This is why voting is restricted to certain age groups and why in most democracies severly mentally handicapped are not eligible to vote.

politic through the act of voting, and thereby govern themselves, the resulting laws this collective decides on can be taken to be collectively justified through the democratic process, meaning that every individual has governed him or herself through participation in the collective action of voting. Because hereby the laws are justified for each autonomous individual, the coercive power of the state is justified for the public at large.

Now that we have seen the broader connection between autonomy and public justification, and why these values are of utmost importance to democrats, I will put forward a conception of autonomy by hand of Gerald Dworkin (1988, p.20) that I will use throughout this thesis:

"Autonomy is conceived of as a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values. By exercising such a capacity, persons define their nature, give meaning and coherence to their lives and take responsibility for the kind of person they are".

On the above account, the autonomy of a certain action is not expressed by someone's liberty to act in accordance with first order preferences[35] but rather with the agreeableness this person feels with a certain act after critically reflecting upon it. By invoking higher-order preferences as the defining preferences for autonomy, Dworkin succeeds in conceptually differentiating between *autonomy* and *liberty*. Where the latter expresses that one act without external impediments, the former only means one may identify with a certain act whether one was at liberty to act or not. This means that on Dworkin's account, if I grant someone the power to vote in my stead, trusting this person to make the right, perhaps even a better decision than I would, while I am not at liberty to influence the outcome after some point, I have still acted *autonomously* if, after critical reflection, I am in agreement with the procedure I have taken. Dworkin's account makes clear the connection of *autonomy* with both our *capacity* to decide and reflect on our preferences and *authenticity,* namely our identification with those preferences.

This further allows decisions based on considerations of loyalty, membership of a group or other special relations to be taken as *autonomous* decisions if these result from a procedure in which the independence of the actor is not violated. We may thus be influenced by others, even giving decisional power away, while keeping *procedural independence*. However, not all influencing acts are compatible with procedural independence. To know whether my procedural independence has been maintained, according to Dworkin (1998, p.18), I need to be able to distinguish between "*those ways of influencing people's reflective and critical faculties which subvert them from those which promote and improve them*". In other words, I need to know if the information I use to critically reflect on my first order preferences, namely "*the genesis of a certain need*" (Ibid, p.29) is not tainted, with which Dworkin means that the discovery of this genesis does "*not tend to undermine my acceptance and identification with it*" (Ibid, p.29). To know if *voter manipulation* has effectively diminished the autonomy of the manipulated, is to ask the difficult and hypothetical question whether discovery of *voter manipulation* by the manipulated would lead to a suspension or acceptance and/or identification with the voting act and or the procedure that resulted in it.

This entails that the manipulated should know if he or she is being manipulated, which is paradoxical, as it is exactly the unstraightforwardness, and thus the hiddenness that makes manipulation what it is. This being the case, I will explain in chapter III, that citizens ought to have a suspicious attitude towards information that may be used to alter or bolster their political preferences to maintain their procedural independence and therewith their autonomy.

---

[35] For an account of first and second-order preferences please see: Frankfurt, H. (1971) Freedom of the Will and the Concept of a Person, *Journal of Philosophy*, 68: 5–20.

<p style="text-align: center;"><u>b: Manipulation Cases</u></p>

To see how this works, and to make up on my promise to show the relation between *moral responsibility* and *autonomy*, I will treat a case of *global manipulation*[36] and extrapolate insights to a *voter manipulation* case. Derk Pereboom (2014) has come up with a thought experiment suited for our purpose. This thought experiment is meant to trigger an intuitive response about moral responsibility. In Pereboooms' *manipulation case* victim A, a gentle, non-violent fellow, is manipulated by neuroscientist B, resulting in him killing C. In one of the scenario's, A's neural states can be directly and instantly changed by B with radio-like technology. A has no knowledge of B's ability to use this technology nor would have consented to B using this technology under any circumstances. In spite of this, B proceeds to alter A's neural states, making him highly aggressive which results in A killing C.

We can take away some important insights from Pereboooms' thought experiment. First, it shows that B's manipulation results in the diminishment of A's autonomy as A loses the capacity to critically reflect on the first-order preference to kill C. The lost autonomy is transferred to B who unilaterally and effectively decides over A's actions in this regard and is in a position to critically reflect on his first order preference to have C killed. B is in a position to use his capacity for moral reasoning and thus, we have a strong intuition that B ought to have considered the moral status of killing C. If we were to ask who is *morally* responsible for killing C, a suitable response would be to hold B accountable, as A, although bearing obvious *causal* responsibility, could not act out his capacity for *moral reasoning* and not for any culpable reasons either, as A did not voluntarily led B alter his neural states. Would A after the act, critically reflect upon him killing C, A would be unable to identify with the act as the genesis of the preference to kill C stems from B's manipulation of his neural states alone, and the second-order capacity of reflection needed for identification with this preference never came into play, meaning that A is in no position to take responsibility over his action. It is obvious that the procedural independence, needed for A to act *autonomously* was appropriated by B. I claim that, and I think this uncontroversial, B has *moral responsibility* over A's action and he has *moral responsibility* to the extent that he infringed on the procedural independence of A, which in this case, amounts to complete moral responsibility over the act of killing C[37]*.*

Now we need to bridge the gap between this rather fantastical but clear case of manipulation and a case of manipulation that could actually take place and is not dependent on science fiction technology that instantly alters one's neural states. How does Pereboom's manipulation case translate to an ordinary case of *voter manipulation*?

Imagine a scenario where B purposely misinforms A to alter A's political preference to X, resulting in A voting X? A is clearly not deprived of the second-order capacity to critically reflect on the decision to vote X. In addition, A may well identify with choosing candidate X for the reasons supplied by B after reflecting on the procedure. However, this only works to the extent that A believes the information given by B to be true, which it is not. Certainly, the misinformation given by B resulted in A using his capacity to critically reflect on his voting preference, but the critical reflection fails to generate any true identification with the preference, as when A would find out that the genesis of the preference to vote X, which stems from B, was made his preference through misinformation, he would most likely suspend identification with this preference[38].

---

[36] Global manipulation is complete manipulation. An example of this is brainwashing.

[37] Note that we can swap the act of A killing C with A saving C's life without any change to B bearing moral responsibility over this act. This is because our account of moral responsibility is inconsequential as it relies purely on procedure.

[38] And if not, and A is happy to have decided in favor of candidate X, A may still feel diminished by the means employed by B, as Scanlon (1988, p. 68) writes about such a case: *"I want to choose the furniture for my own apartment, pick out the pictures for the walls, and even write my own lectures despite the fact that these things might be done better by a decorator, art expert, or talented graduate student. For better or worse, I want these*

The procedural independence of A was not appropriated as in the first manipulation case. Surely it was infringed upon by B as the misinformation limited A's capacity for critical reflection by corrupting A's available information but A is not a victim to the same extend as in the prior case. Perhaps A ought to have collected more information about X. It might have been naïve of A to trust B to be honest. Perhaps B was a campaigner and A ought to have been suspicious as campaigners cannot be trusted to give objective or unbiased information. Maybe A ought to have questioned B's intentions or the misinformation could have easily been discarded but and A was just too lazy to look into it. In other words, there a viable reasons for holding that A's lack of critical reflection, despite B making it so, is nonetheless culpable[39].

We thus end up in a situation where both A and B have a certain *moral responsibility* over A's preference to vote candidate X. B's moral responsibility extends as far as A's autonomy was diminished, namely as far as giving misinformation to A. A is morally responsible to the extent that he relied on B's misinformation, thus allowing to have his autonomy diminished, for culpable reasons, if there are any[40]. The extent to which both are morally responsible than, is depended on the specific conditions under which A came to accept B's misinformation.

We have seen that, the autonomy of A is diminished because the act of *voter manipulation* renders A unable to properly use his capacity for critical reflection. The *unstraightforwardness* employed by B makes it so that A lacks the necessary information to act out this capacity. We have seen that A may bear moral responsibility to the extent that him failing to critically reflect stems from culpable reasons. B purposefully infringes on the *procedural independence* of A therewith bearing moral responsibility over his infringement, which in this case amounts to his use of misinformation. This suffices to argue for the *wrongness* of *voter manipulation* from a liberal perspective. As A and B share moral responsibility over A's voting decision. The *one person one vote principal* and therewith the equal weight accorded to each vote, is violated in spirit. A's individual autonomy is weighed less than B's individual autonomy as A is only in part responsible for his own vote, where B bears moral responsibility over his own vote and a part of A's. In addition, as the conditions for A are not met to identify with the procedure of voting and its outcome, the voting procedure is tainted in a way that fails to make it generate *public justification* for the resulting outcome as the outcome is not the direct result of A expressing his individual autonomy. A is not able to identify with the outcome, he may feel played, as if his vote is somehow stolen from him, and although in part he might have himself to blame, he has good reasons for holding B accountable for generating this regrettable outcome.

The *unstraightforwardness* resulted in a lack of information on A's part to critically reflect on the decision to vote X and therewith liberal democratic values were violated. It is both in the interest

---

things to be produced by and reflect my own taste, imagination, and powers of discrimination and analysis. I feel the same way, even more strongly, about important decisions affecting my life in larger terms: what career to follow, where to work, how to live"

[39] Those who hold *manipulation* is always wrong, may find it preposterous of me to claim that the 'victim' of this manipulation case also bears responsibility. How can it be that one who trusts easily and is deceived ends up bearing moral responsibility over the resulting action. For one, I think that these people have too narrow an understanding of what critically reflecting entails. We ought to be realistic about humans and we therefore, how regrettable it may be, ought to be suspicious. To truly critically reflect is not only weighing the plausibility of information but extends to reflecting on the reliability and interests of the source. One who fails to do this, is not critical to a sufficient extend and has oneself to blame for it. However, the probability of success for critical reflection rests both on the availability of information to use to reflect upon, as a proper, perhaps somewhat suspicious, mind-set. This will be the topic of the next chapter.

[40] We can think of a case where B has successfully argued against all grounds on which A might find B suspicious, and tried to find evidence contrary to B's information without success, therewith leaving A in a situation where he trusts B's misinformation for no culpable reasons at all.

of citizens of liberal democratic states as the state itself to guarantee the availability of certain minimum conditions needed for successful critical reflection. In the next section, I will treat some of the duties of both citizens and the state consequent on my assessment of the wrongness of voter manipulation.

Chapter III: Duties of Citizens and the State

In the previous section, I showed how the *unstraightforward* element of *voter manipulation* leads to a violation of the liberal democratic value of *autonomy* and undermines the *public justificatory* role of voting. In what will follow, I will argue that citizens have a civic duty to be suspicious to safeguard them and the state from the wrongs of *voter manipulation*. In addition, I will argue that the state has a corresponding duty to provide citizens with the means necessary to effectively carry out the duty of suspiciousness, call this, the duty to alleviate suspicion.

Chapter III Section 1: Civic Duty of Suspiciousness

Before I explain the duty of suspiciousness, I want to stress that I take this duty to be a strictly civic duty. It is as Raz (2006, p.1004) puts it "*a positive duty to take steps to secure the safety and advance the interests of her country*". The duty of suspiciousness is due to all citizens of liberal democratic states qua their being citizens. Although it may be prudent for all people to have a suspicious disposition towards information that targets their political preferences, it is a necessary, albeit insufficient, duty for citizens of liberal democracies to remain as such by maintaining a healthy liberal democratic state. Only when the duty of suspiciousness is fulfilled do citizens safeguard the state from the lack of legitimacy that results from a failure to justify the coercive power of the state through voter manipulation.

Citizens of autocratic states or monarchies, more generally those citizens of states that do not share their political power with the people in accordance with liberal democratic values, do not necessarily have to fulfill a duty of suspiciousness to counteract voter manipulation, as voter manipulation does not undermine the legitimacy or public justification of states where there is no voting process[41].

With this, the duty to be suspicious does not impart on any dealings of citizens that do not relate to the civic act of voting. As I have only argued for the specific wrongness of *voter manipulation*, I do not want to claim here that manipulation in general is wrong, nor do I intent to argue that people ought to be suspicious at all times. This would be too stringent a duty as most of the times being suspicious is not warranted nor necessary. In addition, I do not claim that all forms of manipulation constitute a *wrong* towards the manipulated, as possibly with *nudges* outside the voter context, and it would go too far to hold that citizens have to be suspicious in cases where this may be against their own interest.

I do hold that, in the capacity of citizens of liberal democracies however, it is a duty to be suspicious, as without a proper suspicious disposition, as we have seen in the previous section, citizens may end up in a position where they share moral responsibility over the violation of liberal democratic values that are fundamental for citizenship. By letting manipulators manipulate, citizens of liberal democracies undermine their own role as citizens and with the corresponding diminishment of autonomy that functions to legitimize the voting process, this ultimately undermines the state.

As we have seen that *voter manipulation* alters the political preferences of citizens through

---

[41] Whether such states can be legitimate at all is a question I will leave to others.

unstraightforward means, and the *unstraightforwardness* results in the violation of liberal democratic rights, I take *being suspicious* to be a disposition a citizen has qua citizen *vis a vis* influencing acts that may alter or bolster political preferences when convincing grounds to trust an influencer are lacking[42]. The disposition being one where the citizen has to *seriously consider*[43] the possibility that an influencer is being unstraightforward when engaged in an influencing act that *seems to be aimed* at altering or bolstering one's political preferences.

I will explain the terms in *italics* first, thereafter I will give an account of suspiciousness as an *intellectual virtue* for citizens of liberal democracies.

First, what does it mean to *seriously consider* this possibility? To seriously consider the possibility is to take the possibility that one is being manipulated as the truth of the matter, and try figuring out why and how one is being manipulated. Serious consideration of this kind is the active part of the latent disposition of suspicion that ought to be triggered when one encounters such information.

To be autonomous, we have seen[44], one has to critically reflect on the basis of higher order preferences how one's first order preferences relate to the higher-order ones, and how the latter preferences came into existence. *To take serious* the possibility that one is being or has been manipulated, is to critically reflect on the possibility that the genesis of one's preference to vote a certain way has come about through a manipulative act. This means that one has to attempt to trace back in memory how the voting preference came about, and when being engaged in an influencing act relating to the act of voting, on has to assume that persons one interacts with are *unstraightforward* bout their intentions if there are no convincing grounds for being trustful.

Only if one has taken the step to *seriously consider* the possibility of manipulation, I take it that one has succeeded in the critical reflection needed for *autonomous decision-making*. If not, the neglect to have done so, the *"benighting act"*, constitutes a wrong on the part of the citizen as she failed to act out the active part of *being suspicious* (Smith, 1983 p.547). As the citizen failed to properly be suspicious, this may well result in the *"unwitting wrongful act"* of failing to be completely morally responsible over one's vote thus violating the *one person one vote* principal (Ibid.). If this is the case, the manipulated may be held responsible for failing to fulfill the duty of suspiciousness, therewith making oneself susceptible to have ones *procedural independence* infringed by a manipulator resulting in the violation of one's autonomy.

Furthermore, the active part of *being suspicious*, the *serious consideration*, ought not to only trigger when one's political preferences are obviously targeted, as in a discussion about politics or politicians. Because of the *unstraightforward* nature of manipulation, *serious consideration* must come into play when information *seems to be aimed* at one's political preferences. Manipulators may try to feed nonpolitical (mis)information to manipulate citizens such as misrepresenting climate data, economic forecasts or the situation and attitudes of the working poor, to alter or bolster political preferences. In those cases where these topics may relate to one's political preferences, one ought to be suspicious too.

However, if manipulation is effectively applied, a citizen will most likely not know if he or she has been manipulated, even after *serious consideration* of the possibility. *Being suspicious* will never rule out the practice of manipulation. It is simply impossible to reveal the intent of every influencer through *seriously considering* the possibility of being manipulated. In addition, manipulation may work in subtle, almost undetectable ways, as when one politician is portrayed vividly by a website while the opponent looks tired, and triggering *serious consideration* in all the right instances is demanding and difficult.

---

[42] More on convincing grounds in the next section as this relates to the state's duty to alleviate suspicion.

[43] I will use the term 'active suspicion' for *serious consideration*. Note too that *serious consideration* always applies to considering it a possibility that one is being manipulated.

[44] Chapter II Section 2.

This being said, a suspicious disposition is all we can ask of citizens for guarding themselves against *voter manipulation*, while not infringing upon freedoms that allow influencing acts to take place, such as the freedom of speech. As it would be illiberal to infringe on basic rights of citizens that allow them to influence and even manipulate we are left with a situation in which we must allow citizens the possibility to manipulate each other on the basis of their rights, while combatting manipulation without touching these. This means educating citizens to have the proper attitudes to guard themselves, and by extension the state, from manipulation and discourage the use of manipulation while unveiling instances of it where possible[45].

I claim, the proper attitude of citizens to protect themselves from manipulation, is *suspiciousness.* The attitude of s*uspiciousness* is an intellectual virtue for citizens of liberal democratic states, to arrive at knowledge about political information. To act in accordance with this virtue is to act *epistemically responsible* in this area of knowledge. Here, I will take a hybrid approach, using both insights from reliabilist and responisiblist to treat the attitude of suspiciousness and argue where its weakness lies. Overcoming this weakness will be an important theme in the next section.

According to virtue reliabilists, an intellectual virtue is "*any stable and reliable or truth-conducive property of a person*", including among other things *"memory, introspection, sense perception and reason"* (Baehr, 2006, p.479). Virtue responsibilist take intellectual virtues to be *"the traits of a successful knower",* where these traits include *"fair-mindedness, open-mindedness, intellectual attentiveness, thoroughness, tenacity and courage"* (Ibid.). Where the reliabilist traits are methodical, in the sense that these are traits related to a stable and reliable interplay of cognition and reflection, the traits of the responsibilist are character traits, in the sense that these are dispositions of the thinker, rather than methods the thinker employs to arrive at knowledge.

The intellectual virtue under attention, namely *suspiciousness* consists of an active and a passive part. The passive part is made up by a latent disposition *vis a vis* information. When information seems to be aimed at one's political preferences, the passive part of *suspiciousness* ought to generate a sense of suspicion in the listener which allows the active part of *suspiciousness* to be triggered. *Passive suspiciousness*, is akin those traits described by virtue responsibilist, as one need attentiveness, thoroughness, open-mindedness and perhaps even courage to second guess an influencer and consider the possibility of being manipulated.

The active part of suspicion, *serious consideration*, is related to the traits put forward by reliabilists. Active suspiciousness stands in need of a fair bit of *introspection* and *memorizing* as to how one's political preferences came about to ultimately decide whether one's autonomy is diminished[46]. Sense perception may play a role as one tries to identify suspicious behavior, in for example body language, from an influencer. In addition, the critical reflection needed for this is generally taken to be closely related to *reason*, as critical reflection allows us to act on the basis of reason rather than impulse.

However, it seems implausible that *seriously considering* the possibility of being manipulated results in a stable and reliable method for discovering the truth, as reliabilists would have it. There is simply not a reliable method for detecting manipulation, as a single citizen can only guess, sometimes more, other times less accurate, the intent of a manipulator.

Considering the function of *suspiciousness*, we can, however, discern a successful knower from an unsuccessful one, in the sense that the former has applied suspicion through using those character traits described by responsibilist. With this I mean that the latent disposition of suspicion has been made active in the right instances[47] through tenacity, intellectual attentiveness, thoroughness and the

[45] More on discouraging manipulation in Chapter 4.
[46] See: Chapter II Section 2b on procedural independence.
[47] When information is seemingly aimed at one's political preferences.

like. This means that for passive suspiciousness, the virtue responsibilist way of understanding intellectual virtues fits as to when one's suspicion ought to be made active. However, when suspiciousness is made active, *serious consideration* by individual citizens fails to arrive at a method that is stable and reliable enough to be deemed an intellectual virtue according to reliabilists.

The disposition of suspiciousness, its passive form, is an important intellectual virtue for citizens in liberal democracies and citizens ought to strive to possess the right character traits to cultivate this virtue and thus act out their duty of suspiciousness. In addition, this entails that active suspicion, *serious consideration*, must come into play when citizens are confronted with information that seems to be aimed at their political preferences. When suspicion becomes active at the right moments, this shows that citizens have the proper disposition of suspicion, and that they act epistemically responsible in this regard.

We have seen that active suspicion suffers from the weakness that it has no reliable method for discerning truth, which makes it unfit to properly detect manipulation. Therefore, it would be absurd to claim that citizens have a duty to act out the active part of suspicion successfully. Citizens who act on a proper disposition of suspicion, cannot be held responsible for committing unwitting wrongs that result from mistakes made when *seriously considering* the possibility of being manipulated, as these citizens have not omitted the critical reflection that would be necessary for them to commit a wrong, the benighting act, in this regard.

As such, the duty of suspiciousness for citizens only relates to successfully acting out the passive disposition of suspiciousness and not successfully carrying out the active part. In other words, citizens ought to act epistemically responsible but not epistemically reliable when a reliable method for discerning the truth is unavailable. If such a method were available, citizens ought to act both epistemically responsible and epistemically reliable.

As it is unlikely that an individual citizen can come up with a reliable and successful method for detecting manipulation and thus to successfully act out the active part of suspicion, citizens ought to be supported. The state has a responsibility in this regard as the state is dependent for the justification of its coercive power on citizens acting out successfully both the passive and active part of suspicion. As such I take it to be the duty of the state to support citizens in this regard and help them carry out the active part successfully.

Chapter III Section 2: Duty to Alleviate Suspicion

Corresponding with the duty of citizens to be suspicious, the state has a duty to alleviate the suspicions of citizens by making available the means necessary for citizens towards this end. As citizens ought to be suspicious when lacking convincing grounds to trust someone who influences their political preferences, the state must strive to give convincing grounds for trust to its citizens.

I take this to be a duty of the liberal democratic state as the state would undermine itself when it cannot grant citizens the means necessary to distinguish between manipulative and non-manipulative influencing acts in the voting context. Without the ability to distinguish between the two, citizens are vulnerable to having their autonomy diminished in the voting process. As we have seen[48], individual autonomy is a necessary condition for the public justification of state power in liberal democracies. If the individual autonomy of citizens is not properly protected by the state, the state therewith invites the causes of its own demise, namely the dissolution of the justification of its own coercive power.

The duty to alleviate suspicion for the liberal democratic state is twofold. Considering that the state also has a role to play as prime educator of the people, the first part of the state's duty is to

---

[48] Chapter II Section 2a

cultivate in its citizens the latent disposition of suspicion. As one cannot alleviate someone from a pain one does not feel, the state cannot alleviate suspicions where there are none. The state must educate its citizens to have the proper disposition towards those who influence their political preferences. This disposition has to be taught to citizens as the state therewith prepares those citizens who are not yet eligible to vote for the time they are expected to vote as autonomous individuals. Citizens who have no disposition of suspiciousness are prone to being manipulated and therewith to have their autonomy diminished in the voting process. As such, the state cannot hold that merely age suffices as a benchmark for autonomy. Proper education that teaches citizens the disposition of *suspiciousness* is necessary as a bulwark against those influences that diminish autonomy and therewith erode the public justification of the state's coercive power.

As citizens ought to have a passive disposition of suspiciousness, which is an attitude of disbelief towards those who influence political preferences, the state has to give citizens the means to belief the right influences, the right influences being those that make use of rational persuasion rather than manipulation, and detect those influences that employ methods that diminish autonomy.

If citizens are to be properly prepared to act out their duty of suspiciousness, this means in practice that the state needs to guarantee citizens access to information they can trust and which citizens can use to assess whether they are being manipulated. This may include among other information, accurate data on political information but also allegiances of certain important influencers. This information must be straightforward, meaning that there is no hidden intention other than the non-hidden intention of informing citizens as objectively as possible[49]. Citizens ought to be able to consult this information to see whether it corresponds with the information given to them by influencers. In this way the state can support citizens in acting out the active part of suspiciousness by giving citizens the means needed for a somewhat reliable method to separate manipulation form rational persuasion.

This only tackles one side of the problem though, as emotional manipulation cannot be easily offset by giving citizens access to information but relies on the most part in the capacity of citizens to detect emotional manipulation in the first place, resist it in their voting decision and rely on reason in its place. For this part of its duty, the state can only try to make its citizens resilient to emotional manipulation through proper education, allowing citizens to identify emotional rather than reasonable arguments and supporting citizens through access to professionals who may be able to properly identify between the two. Again, it would be easier if the state could ban the use of emotional manipulation in for example political messaging, however, it would restrict the freedom of speech and would therefore be illiberal.

Those who hold we should not trust the state too much are justified in holding the view that while the above duties might seem prudent in theory, it is naive to trust the state to alleviate the suspicion of its citizens in practice. Why would the state, if for example lead by politicians who depend on manipulative methods, cultivate its own citizens to detect manipulation? It seems as if the duty to alleviate suspicion places a lot of trust in the liberal democratic state and assumes it to be benevolent towards its citizens. Any philosopher in his or her right mind ought to think of the possibility that the state is somewhat malevolent and theorize accordingly. How do we protect the citizens and the liberal democratic state from manipulators if they are in power, and how do we ensure that the state cultivates in its citizens the disposition of suspiciousness and fulfills its duty to alleviate suspicion? This will be the topic of the next section.

---

[49] As information always is presented from a certain perspective, there ought to be openness about the perspective, the source of information, the writer, etc.

Before responding to the charge that too much trust is placed in the benevolence of the state, let me note that, although the state has a duty to alleviate suspicion, it does not have to be the state itself that actively makes it so that this duty is being fulfilled. A feature of liberal democracies is their relatively open public sphere[50]. Rather than creating a ministry of information that regulates the flow of information towards all citizens, the state ought to allow the public sphere to flourish and discourage manipulative information while encouraging straightforward information through subsidies and other instruments that are available to it. Keeping the sources of information and the state separated is a liberal thing to do, however, the state has a duty to maintain a certain level of information in the public sphere that allows citizens to act out their active duty of *suspiciousness* meaning the state should intervene in the public sphere if this level of information is not reached. Taking control of the information flow in its own hands should be a last resort for the state if their public sphere consistently and utterly fails to generate reliable information that citizens can use to safely inform their political preferences.

Indeed, the responsibility of the state is great as citizens are dependent on it for acting out their duty of suspiciousness if the public sphere fails to sufficiently generate reliable information. How can we ensure the state acts out the corresponding duty to alleviate suspicion? Even more when the state itself is ruled by politicians who might gain from not acting out this duty at all. There are certain legal tools. Codifying laws that force politicians to act out this duty is a logical place to start. However, it is not enough and the harsh truth is, there may be no way of ensuring the state fulfills its duty by making the executive itself responsible for acting on it. There may simply be too much incentives for politicians to manipulate as they are an important subject in the forming of political preferences and have a direct interest in it. As this is the case, we must force to state to act out its duty by creating checks and balances outside the control of politicians. Independent bodies ought to be raised that can scrutinize the state and take legal measures when the state does not comply with its duty. In addition, citizens ought to hold the state responsible for acting out its duty of alleviating suspicion, electorally punishing those politicians who do not comply.

As the active part of the duty of suspiciousness relies on a collective effort of a society to generate reliable information from and for citizens and considering this is needed to safeguard the liberal state in which these citizens live, further considering the citizens choose their own rulers in liberal democracies, it are ultimately citizens themselves that have a duty to bring about a situation in which the state alleviates suspicion. Much akin Rawls's (1971, p.259) duty of justice that holds citizens ought to strive to bring about just institutions when these do not exist, so too, citizens have a duty to bring about a situation where their individual suspicions are properly addressed through means to successfully act on their active suspicion. The citizens being a collective, have a forward looking collective responsibility, meaning citizens ought to look if they can remedy any harms they have caused by failing their duty of suspiciousness, rather than putting the blame on the collective (Young, 2021). The upshot of forward looking collective responsibility over backward looking collective responsibility is its alignment with the ultimate goal to create a society in which the duty of suspiciousness can be made active through the fulfillment of the duty to alleviate suspicion. Ultimately, we want to better society. Blaming citizens for failing their duty of suspiciousness, and thus invoking backward looking collective responsibility would unnecessarily shift the burden of acting out the duty of suspiciousness successfully to citizens rather than the state, while the state has a pivotal role to play in making available the means to act out the duty of suspiciousness to citizens. Rather than blaming citizens for a state of affairs in which the duty of suspiciousness cannot be successfully carried out, citizens ought

---

[50] When compared to authoritarian, dictatorial or theocratic, thus illiberal, states.

to strive to remedy the incapability of acting out their duty successfully by acting on their duty to bring about a situation where the state alleviates suspicions.

As the citizens rule over themselves in liberal democracies via politicians, it are fundamentally the citizens who ought to force their chosen rulers to alleviate their suspicion, to be able to live in a society in which one can to a reasonable level detect manipulation and thus legitimize the coercive state power the citizens as whole, through voting, have decided upon.

Citizens can do this in multiple ways. By contributing to the public sphere for example, through deciding which outlets of information to choose, through political activism against rulers who do not fulfill their duty to alleviate suspicion and by voting for those who do.

Ultimately, the question whether the state is malevolent or benevolent is not as important as the quality of the citizenry of the liberal democratic state as it is ultimately the citizenry that can decide how to shape the state. This answer may feel unsatisfying as it leaves a lot to chance, and the quality of the citizenry is for an important part affected by the state, for instance through education. However, this is an intrinsic quality of democracies and not so much a specific problem of the duty of suspiciousness and the duty to alleviate suspicion.

To see how to apply both duties to voter manipulation in democracies, we will in the next chapter treat contemporary technologies that can be utilized to manipulate voters. These so-called cybernetic technologies put forth new challenges for a democratic process in which citizens make autonomous decision and therefore invoke the duties treated in this chapter.

Chapter IV The Cybernetic Age

We have identified and evaluated voter manipulation and treated the duties citizens and the state have in relation to it. Thus far, I have kept our treatment of voter manipulation divorced from technology and specific manipulative actions. In this chapter, I will treat how our conception of manipulation and the duties of the state and it citizens relate to contemporary, cybernetic forms of manipulation. In the first section, I will give a brief overview of new cybernetic technologies that exacerbate the potential and reach of voter manipulation. In section two, I will discuss the implications of these technologies for liberal democracies and how these technologies affect the duty of suspiciousness and the duty to alleviate suspicion.

Chapter IV Section 1: Cybernetic Technologies

President Biden in the 2021 State of the Union predicted that *"we'll see more technological change in the next 10 years than we saw in the last 50, that's how rapidly artificial intelligence and so much more is changing. "*[51].Technologies progress at an unprecedented rate and with this the means to manipulate the vote. So called, cybernetic technologies, technologies that are developed by making use of feedback loops, such as algorithms and AI are at the forefront of this technological progression. Most prominent for the means of manipulation is the unprecedented capability of technologies to utilize highly personalized data to create detailed profiles that can be analyzed to gain insights into psychological activity and characteristics of individuals (Kosinski, Stillwell, & Graepel, 2013). These technologies have been developed for marketing purposes where predictability of costumers makes advertisement more effective and therewith lucrative, driving capitalist to "*acquire ever-more predictive sources of behavioral surplus: our voices, personalities and emotions* (Zuboff, 2019, p.11).

---

[51]Retrieved from: https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/04/29/remarks-by-president-biden-in-address-to-a-joint-session-of-congress/, on 10-05-2021

However, lucrative, the consequences of these technologies for consumers *"tend to be rather small and unimportant for the way in which people's behavior is affected in general"* (Calo, 2014, p.1002). Although the effects on humans in the marketplace is rather small, we must not underestimate the impact these technologies have in the political realm, where "*affecting the political preferences, decisions or actions of even a small proportion of voters in a competitive election could be critical to the outcome* (Burkell & Regan, 2019, p.3). Traditionally political messages target certain voter groups on the basis of demographic information, tailoring the message according to the dominant group in certain neighborhoods. *Psychographic profiling*[52] allows for surgical precision. The combination of voter profiles with knowledge gained from psychology, can be used to develop persuasive messages specifically tailored to the interests, desires and vulnerabilities of certain voters (Issenberg, 2016). As we have seen with the Cambridge Analytica scandal, where *"individuals were rated on a five-factor personality model and messages tailored on the basis of the respective ratings"*, using this method in political campaigns can be impactful (Chester & Montgomery, 2018, p.23-4).

Digital platforms are ideally suited for gaining and exploiting psychographic information. As Susser (2019, p.7) points out: "*digital interfaces (…) are configured in real time*" using psychographic information while continuing *"to learn about us as we interact with them (…) they do not wait passively (…), rather they send text messages and push notifications, demanding our attention, and appear in social media feeds at the precise moment they are most likely to tempt us*". Because of algorithms, this is an automated process, shaping a highly fragmentized and personalized digital environment for every user. An environment in which, "*the vulnerabilities catalogued through pervasive digital surveillance are put to work to influence our choices"* (Yueng, 2017, p.122).

The techniques for tailoring messages are being develop too, to the point where these can have considerable subconscious influence on their target. One of these techniques, morphing, where characteristics of two faces are built into a single image has recently shown to have considerable applications in the political arena. By morphing the face of a candidate to slightly look more like that of a targeted voter, the targeted voters expressed greater support for these candidates than for those who weren't morphed to look more alike them (Bailenson, Ivengar, Yee & Collins, 2008). Other developments, such as deep fake and advances in three dimensional animation coupled with AI, will exacerbate the problem of not being able to distinguish between reality and the slightly altered version of it, provided by influencing parties.

At the heart of psychographic profiling lies the data we as internet users provide to tech companies by making use of their products. By combining the data of individuals and analyzing behavior, tech companies' algorithms create models that predict and steer human behavior. While the tech companies *"know everything about us"*, they design their operations "*to be unknowable to us"* (Zuboff, 2019, p.11) , the individual knows next to nothing about how their technologies function and utilize our data. Moreover, when users grow accustomed to technologies, the medium fades from conscious experience. This so called, *"technological transparency"* allows users to focus on their tasks while utilizing certain technologies making the technology itself hidden and influences it mediates harder to recognize (Van Den Eede, 2011). This asymmetry in information is significant as it bars the individual user from knowing how, why and if (s)he is being manipulated. It is the source of the *unstraightforwardness* that is characteristic of manipulation.

In summary there are three main characteristics of cybernetic technologies in relation to voter manipulation. The first is the scale of data that can be used, in combination with automated processes, to generate a psychographic profile and therewith enables an unprecedented level of personalization. Second, are the digital platforms that enable new and persisting means to influence users. Last is the

---

[52] The combination of personalized information and psychological studies.

inherent hiddenness of the technologies and the way these recede from consciousness when users are adapted to them, making them prone to being manipulated.

Chapter IV Section 2 Technology, Duty and Policy

In this section, I will explore how the duties described in chapter three may be achieved by the state in relation to cybernetic technologies by suggesting certain policy options. I will treat the technologies with related duties and suggested policies in accordance with the three main characteristics I have identified in the first section of this chapter. First, I will treat personalization, hereafter the digital platforms to finish this chapter with a treatment of the inherent hiddenness of these technologies.

a: Personalization

The main implications of personalization in relation with voter manipulation are twofold and pose significant problems for acting out the duty of suspiciousness.

First, the fragmentation of the public sphere makes it increasingly difficult to compare and weigh information and thus to have the means to critically reflect on information received. As the public does not receive the same information collectively but every person receives information individually, it becomes harder to come together with other citizens and compare information and deliberate about the viability of certain arguments or the validity of information.

Second, personalization and the tailoring of messages that accompanies it makes it increasingly difficult to make suspicion active at all, as the receiver will most likely identify with the information received. The tailoring of messages to fit personal preferences makes manipulation hard to detect as our capacity for detecting manipulation is somewhat negated when a message neatly fits our preferences. In other words, we want it to be true, meaning that the attitude of suspiciousness needed to begin critical reflection is somewhat negated.

Bodo (2019, p.2) calls the first type "*personal isolation*" and the latter "*technological isolation*". These can be mutually reinforcing as "*People are diversity averse, and algorithms reduce diversity. Together, users and algorithms create a spiral, in which users are one-dimensional and prefer their information diet to be filtered so that it reflects their interests, and in which this filtering reinforces the individual's one dimensionality*".

Both types of isolation are undesirable as they undermine the capacity for critical reflection by isolating the individual from the collective with whom to reflect and undermining the tendency for reflection by generating appealing information for the target. To counter this, Burkell & Regan (2019, p.16) suggest we should broaden the right to access information as "*a basis for legitimate restrictions on the use of sophisticated analytics in targeting political messages*". This may be a viable avenue as according to the OECD (2018) the right to access is important to "*establish and support the democratic and participatory practice in government*". If democratic practices are undermined by personalization as people receive one-dimensional information this may be a reason to limit personalization to protect the public sphere from fragmenting. Another avenue may be to see personalization as a violation of the right to receive information. As Mart (2003, p.178) argues "*it is the right of the viewers and listeners. Not the right of broadcasters that is paramount… It is the right of the public to receive suitable access to social, political aesthetic, moral and other experiences which is crucial*" adding that (Ibid p.181) to receive these ideas "*is a necessary predicate to the recipients meaningful exercise of his own rights of speech, press and political freedom".* A third way of restricting personalization in political

messages is suggested by Walker Wilson (2010, p.742), who argues that *"the definition of corruption ought to be expanded to include the potential for distortion in voting-behaviour as a result of heavy handed psychological tactics"*, as democracies cannot function without a "*free and willing voting public*" and a process that is not tainted by *"systematic, wide-scale manipulation by any segment of the public, individual candidate or political party*". If Walker's argument were to be judicially agreed upon this would have significant consequences as it would prohibit spending campaign money on personalized digital political advertising.

If personalization leads to a one-dimensional information diet, citizens may be undermined in practicing their rights and this may be a valid reason to prohibit or restrict personalization. As information is fundamental for making autonomous decisions and for the capacity of critical reflection, the state has a duty to provide citizens with the means necessary to practice their suspicion and to prohibit or restrict practices that undermine this capacity, broadening the right to access information or expanding the definition of corruption to include the relationship between money and potentially manipulative communication strategies may be a good place to start.

However, this does not tackle the issue completely and a more radical approach is needed to ensure that liberal democracies are protected against the threat of personalization. As the market brings personalization through the internet, the state can counter this with supporting or creating powerful unpersonalized media outlets. A place where the information for all citizens is the same and which citizens can use to deliberate amongst each other on equal terms. This does not have to be a new technology. Either regulated private or state-owned television or radio can fulfill this role and citizens who carry out their duty of suspiciousness ought to commit themselves to those sources of information that generate no personal isolation and do not use agreeably tailored messages to capture the attention of viewers. It is the responsibility of the state to make available such mediums and the responsibility of citizens to restrict their information diet to those sources that do not personalize political information.

b: Digital Platforms

In the first section of this chapter, we saw that digital platforms are ideally suited for continuously barraging us with tailored messages in the most effective and intrusive manners. Being structurally exposed to this type of messaging makes influencing us a continuous effort on the part of the automated systems that rule these platforms. This continuous exposure to information, if political, means that citizens have to continuously strive to remain suspicious. However, the nature of the messaging on these platforms, being systems that run on user feedback, makes this hard as the information and its form adapts continuously *"in order to most effectively influence us"* (Kaptein & Eckles, 2010). Here again the psychographic data employed by these platforms play an important role in undermining the critical faculties needed for autonomous decision making. How can the state help the citizens to successfully employ their suspicious disposition?

The European Union may have an answer to combat the potential manipulation that arises from use of these platforms. The freedom of expression entails a right to refuse information to protect citizens from *"a level of information exposure diversity that could infringe upon individual freedom"* (Burkell & Regan, 2019, p.15). Obviously too much information can be as bad as too little information and if the way in which information is provided to the public reaches the point where individuals cannot duly reflect on the information received, there may be a positive obligation of the state to intervene and protect its citizens from the potential manipulation that they are exposed too.

In addition, voters ought to be taught to be digital literates, meaning that the state provides its citizens with the education needed to restrict one's information diet, recognize tailoring and make

them resilient against the psychological tricks employed on platforms. Such an initiative has already been argued for, funded through a special tax on social media (Cooke, 2018). However, recognizing the subtle forms of manipulation may be too hard and there is reason to suspect that platforms will adapt to the education.

The duty of suspiciousness, if the platforms adapt successfully to this education steer to a more extreme approach, namely exiting the platforms and postpone using them until these platforms stop employing manipulative methods to capture the attention of audiences. If digital platforms keep adapting to create more subtle forms of manipulation ,retreating from these platforms altogether may be the only way to avert manipulation.

The state has a role to play as well. Flagging platforms that make use of manipulation and for example forcing these platforms by law to warn users for their manipulative methods, would help alleviating suspicion as it makes clear that these platforms cannot be trusted, therewith triggering the disposition of suspiciousness in citizens who use these platforms. Again the state ought to support or create platforms that are as free of subtle manipulation as possible. Restricting for example images that invoke emotional responses or phrases that are emotionally laden. Although this sounds as a difficult task, it can actually be done by making use of the exact research that forms the basis of psychographic profiling as it is on the basis of this research that conscious subtle manipulation is currently utilized. Instead of using this research to make subtle manipulation more effective, it can be used to rollback these methods and rid platforms of the subtle manipulative tools they employ to influence users.

Furthermore, guidelines can be made as how to create healthy platforms. Codifying these guidelines into law and enforcing it would most likely be too big an infringement on the freedom of speech. The state is only allowed to go as far as to discourage use of platforms that employ manipulative tools and encourage the use of platforms that do not.

<div align="center">c: Hiddenness</div>

This leaves us with the inherent hiddenness of new technologies. The asymmetry of information between the owners of digital platforms on the one hand, and citizens on the other, is staggering. This asymmetry entails that citizens do not get to know why, how, by whom and if they are being influenced or manipulated. As I argued that *unstraightforwardness* is the fundamental characteristic of manipulation, the inherent opacity of digital platforms maintains and reproduces the manipulative tendencies into the future. The main problem with hiddenness is not that the technology becomes unrecognizable to the user, of course this may exacerbate manipulative practices as it makes users more vulnerable, but that digital platforms operate like black boxes. The platforms are not transparent as they operate as influencers, mostly through (political) advertisement and being open about why someone is being targeted and by whom reduces the likelihood of someone being influenced. From the perspective of the state's duty to alleviate suspicion, to break open the black boxes for the public would go a far way in fulfilling its duty as then the state ought only to support people to do relevant research and maintain oversight on the functioning of digital platforms.

Current European developments can be used as a model to further reduce the hiddenness of digital platforms. The 2018 Code of Practice on Disinformation makes digital platforms responsible for ensuring transparency about sponsored content specifically for political advertisements and restriction the targeting and personalization options of platforms. In addition, explanations of the platforms have to be provided on the functioning of their algorithms, and fact-checkers, researchers and authorities are enabled to continuously monitor digital platforms on misinformation, a means of manipulation.

However, the self-regulatory nature of the code has led some digital platforms, such as Facebook to resist research into their advertisement database. The EU could go further, as it has already threatened and formalize the code into law to be able to penalize platforms that are non-cooperative.

Holding platforms legally accountable for abiding by the rules set in the code of disinformation goes a long way in breaking the information asymmetry between citizen and platform that stands at the basis of contemporary widespread voter manipulation. In tandem with the code, the EU commission tries to ensure a *"pluralistic, diverse and sustainable media environment"* to give its citizens viable and transparent options to turn to for their information besides the platforms (Leerssen, 2019, p.6). If successful, the EU might just have a viable policy option to fulfill its duty to alleviate suspicion towards its citizens.

The state then, to alleviate suspicion, must force any organization or individual that aims at influencing the political preferences of citizens to be completely transparent about their allegiances, thus making public where revenues are made and to what aims information is shared with citizens.

The duty of citizens to be suspicious entails that citizens ought to retreat from digital platforms that are not willing to cooperate with rules such as the Code of Practice on Disinformation, choosing other platforms or information sources that hold up to standards that make voter manipulation at least researchable and thus discoverable. In addition, citizens must inform themselves on the allegiances of the platforms they use so they can guard themselves from manipulative information beforehand.

Conclusion

We will ultimately never be free of manipulation and it would not be desirable to be. To prohibit manipulation would be to prohibit too many valuable freedoms and not all forms of manipulation will ultimately turn out for the worse.

What I have tried to show in this dissertation is how to identify manipulation of voters. I have identified the *unstraightforwardness* employed in acts of manipulation to be the fundamental characteristic. In addition, we have seen that this *unstraightforwardness* undermines individual autonomy. As voter manipulation violates the individual autonomy of voters, and this autonomy is taken to be the fundamental democratic value, it also undermines democracy itself. Hereafter, I have argued that both democratic citizens and the state have duties to resist voter manipulation. The duty of the citizen is to have an suspicious attitude towards political information, while the state has a duty to alleviate the suspicion of its citizens by granting them access to the means to detect manipulation and create or support alternative sources of information if they cannot yet provide these means. Furthermore, the state has a duty to educate citizens to have a proper attitude of suspiciousness when they are eligible to vote and citizens ought to encourage the state to alleviate suspicion by actively striving for a situation in which the state successfully does so. This can be done though the democratic voting process, but also through political activism.

To bridge the gap between abstract philosophical argumentation and current developments relating to voter manipulation, I have dedicated the last chapter to see how these duties, in a non-exhaustive list of contemporary problems, translate into policy options.

I hope the treatment of *voter manipulation* has both been illuminating conceptually and useful in thinking towards solutions to the pressing issues for democracies trying to cope with voter manipulation in contemporary times. As I have noted in Chapter I Section II, the advent of cybernetic technologies carries with it difficulties for conceptualizing (voter) manipulation. My suggestion for further conceptual work on manipulation would be to delve in the interrelated concepts of *agency, intent* and *responsibility*.

The use of cybernetic technologies will probably grow significantly and intellectual work has to be done to understand what this means for our societies. Cybernetic technologies such as algorithms fundamentally blur the analysis as to whom is responsible for being influenced or manipulated through use of algorithmic technologies. Is it ultimately the company owning the algorithm, the designer of it, the consumer using it or the self-learning element in the technology itself we have to hold responsible? Can we speak of intent when an algorithm has developed itself in a certain way to influence political preferences, or must we place the intent with the designer or original idea of the algorithm?

Answers to questions like this are necessary to create legal frameworks that can address voter manipulation, to build working policies to address it as well as to advance our understanding of the morality of influence in a day and age where this is becoming increasingly urgent and important.

Literature:

Ackerman, F. (1995). "The Concept of Manipulativeness", *Philosophical Perspectives*, 9: 335–340.

Bailenson, J.N., Iyengar, S., Yee, N., & Collins, N.A. (2008). Facial similarity between voters and candidates causes influence. *Public Opinion Quarterly, 72(5)*

Jason S. Baehr. (2006). Character in Epistemology. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 128*(3), 479-514.

Berlin I. (1969*) Four Essays on Liberty*. Oxford University Press: Oxford.

Bodó, B., Helberger, N., & de Vreese, C.H. (2017). Interested in Diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalisation. *Digital Journalism (7)2*

Burkall, J., & Regan, P.M. (2019). Voter preferences, voter manipulation, voter analytics: policy options for less surveillance and more autonomy. *Internet Policy Review 8(4)*

Calo, R. (2014). Digital market manipulation. *George Washington Law Review, 82(4)*

Cooke, N.A. (2018). *Fake News and alternative facts: Information literacy in a post-truth era.* Chicago: ALA editions.

Chester, J., & Montgomery, K.C. (2019). The digital commercialization of US politics – 2020 and beyond. *Internet Policy Review, 8(4)*

Dworkin, G. (1988). *The Theory and Practice of Autonomy* . Cambridge: Cambridge University

Faden, R. R., Beauchamp T.L., & King N.P.M. , (1986). *A History and Theory of Informed Consent*, first edition, New York: Oxford University Press.

Feinberg, J. (1989). *Harm to Self*, *The Moral Limits of the Criminal Law Volume 3*, New York: Oxford University Press

Frankfurt, H. G. (1969), "Alternate Possibilities and Moral Responsibility", *The Journal of Philosophy*, 66(23)

Gorin, M. (2014), "Do Manipulators Always Threaten Rationality?" *American Philosophical Quarterly*, 51(1).

Hobbes, T., Schuhmann, K., & Rogers, G.A.J. (2005). *Leviathan: Vol. A critical edition by G.A.J. Rogers and Karl Schuhmann*. Thoemmes Continuum.

Issenberg, S. (2016). *The victory lab: The secret science of winning campaigns.* New York: Broadway Books.

Kant, I., & Wood, A. (1996). On the common saying: That may be correct in theory, but it is of no use in practice (1793). In M. Gregor (Ed.), *Practical Philosophy* (The Cambridge Edition of the Works of Immanuel Kant, pp. 273-310). Cambridge: Cambridge University Press

Kaptein, M., & Eckles, D. (2010). Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. In T. Plouug, P. Hasle, & H. Oinas-Kukkonen (Eds.), *Persuasive Technology*

Kligman, M., and Charles M. C. (1992). "An Analysis of Interpersonal Manipulation", *Journal of Medicine and Philosophy*, 17(2): 173–197

Kosinksi, M., Stillwell, D., & Graepel;, T. (2013). Private traits and attributes are predictable from digital records of human behaviour. *Proceedings of the National Academy of Sciences, 110(15)*

Leerssen, P., Ausloos, J., Zarouali, B., Helberger, N., & de Vreese, C.H. (2019). Platform ad archives: Promises and pitfalls. *Internet Policy Review*. *8(4)*

Mart, S.N. (2003). The right to receive information. *Law Library Journal, 95(2)*

Mills, C. (1995). "Politics and Manipulation", *Social Theory and Practice*, 21(1): 97–112.

Noggle, R. (1996). Manipulative actions: A conceptual and moral analysis. *American Philosophical Quarterly* 33

OECD (2018). Right to Access Information. Retrieved 11-5-2021 from:
    https://www.oecd.org/mena/governance/right-to-access-information-2018.pdf

Pereboom D. (2014) *Free will, agency, and meaning in life*. New York, NY: Oxford University Press

Rawls, J. (1971). *A Theory of Justice*. Cambridge: Harvard University Press.

Raz, J. (1988). *The Morality of Freedom*, Oxford, New York: Oxford University Press

Raz, J. (2006). The Problem of Authority: Revisiting the Service Conception, *Minnesota Law Review*, 90:
    1003–1044.

Rousseau, J. J., & Williams, D. (2014). *Rousseau's Social Contract: An Introduction* (Cambridge
    Introductions to Key Philosophical Texts). Cambridge: Cambridge University Press.

Rudinow, J. (1978). "Manipulation", *Ethics*, 88(4): 338–347

Scanlon, T. M. (1988). *The Significance of Choice*. In Sterling M. McMurrin (ed.), The Tanner Lectures
    on Human Values (Vol. 8). University of Utah Press.

Smith, H. (1983). Culpable Ignorance. *The Philosophical Review, 92*(4), 543-571.

Spinoza B., Israel, J. I., & Silverthorne, M. J. (2007) *Theological-political treatise*. Cambridge: Cambridge
    University Press;.

Sunstein, C. R. (2016). *The Ethics of Influence*. Cambridge: Cambridge University Press

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet
    Policy Review, 8(2)*

Thaler, R. H., & Sunstein, C.R. (2009). *Nudge: Improving Decisions About Health, Wealth, and
    Happiness*, Revised & Expanded edition. New York: Penguin Books.

Van Den Eede, Y. (2011). In Between Us: On the Transparency and Opacity of Technological Mediation. *Foundations of Science, 16(2/3)*

Walker Wilson, M.J. (2010). Behavioral Decision Theory and Implications for the Supreme Court's Campaign Finance Jurisprudence. *Cardozo Law Review*, 31(3)

Whitfield, G. (2020). On the concept of political manipulation. *European Journal of Political Theory*.

Young, Iris, 2011, *Responsibility for Justice* Oxford: Oxford University Press.

Yueng, K. (2017). Hypernudge: Big Data as a Mode of Regulation by Design. *Information, Communication & society, 20(1)*

Zuboff, S. (2019). *The Age of Surveillance Capitalism.* London: Profile Books ltd.