

ARTICLE

## Algorithms, Manipulation, and Democracy

Thomas Christiano

Department of Philosophy, University of Arizona, Tucson, Arizona, USA  
Email: [thomasc@arizona.edu](mailto:thomasc@arizona.edu)

### Abstract

Algorithmic communications pose several challenges to democracy. The three phenomena of filtering, hypernudging, and microtargeting can have the effect of polarizing an electorate and thus undermine the deliberative potential of a democratic society. Algorithms can spread fake news throughout the society, undermining the epistemic potential that broad participation in democracy is meant to offer. They can pose a threat to political equality in that some people may have the means to make use of algorithmic communications and the sophistication to be immune from attempts at manipulation, while other people are vulnerable to manipulation by those who use these means. My concern here is with the danger that algorithmic communications can pose to political equality, which arises because most citizens must make decisions about what and who to support in democratic politics with only a sparse budget of time, money, and energy. Algorithmic communications such as hypernudging and microtargeting can be a threat to democratic participation when persons are operating in environments that do not conduce to political sophistication. This constitutes a deepening of political inequality. The political sophistication necessary to counter this vulnerability is rooted for many in economic life and it can and ought to be enhanced by changing the terms of economic life.

**Keywords:** Algorithms; nudge; manipulation; political equality; low-information rationality; democracy; informational power

Algorithmic communications pose a number of challenges to democracy. The three phenomena of filtering, hypernudging, and microtargeting can have the effect of polarizing an electorate and thus undermine the deliberative potential of a democratic society. Algorithms can spread fake news throughout society, undermining the epistemic potential that broad participation in democracy is meant to offer. They can pose a threat to political equality in that some people may have the means to make use of algorithmic communications and the sophistication to be immune from attempts at manipulation while other people are vulnerable to manipulation by those who use these means. My main concern in this paper is with the danger that algorithmic communications can pose to political equality. This arises because democratic politics are often low-information rationality politics. Most citizens must make decisions about what and who to support in democratic politics with only a sparse budget of time, money, and energy. Algorithmic communications such as *hypernudging* (communications that nudge a person toward some choice while constantly updating the information on which the communications rely) can be a threat to democratic participation when persons are operating in environments that do not conduce to political sophistication. *Microtargeting* (communications that are targeted at highly selectively defined groups) can also be problematic for the democratic participation or persons who are not sophisticated. Low-information citizens can be vulnerable to a kind of manipulation when they do not participate in groups that enable them to act with good information, and this constitutes a deepening of political inequality.

© The Author(s), 2021. Published by Cambridge University Press on behalf of Canadian Journal of Philosophy. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

The basic problem can be seen when we contrast the effects of hypernudging on individuals who are well informed and know what they want to the effects on individuals who are not well informed and do not know what they want, though they may have some aims. There is an inequality here in the differences in information and informational abilities. This inequality may not always be unjust; for instance, it is not unjust when we can hold the person responsible for the difference. But I will argue that inequalities of this sort are pervasive in democratic politics. They are unjust inequalities because individuals have deep interests in being able to understand what is at issue in politics and how to advance their interests and moral aims, and they cannot be held entirely responsible for their lack of information under certain conditions. Or so I shall argue.

One difficulty here is, as far as I can see, that studies of the effects of algorithmic communications on political participation are still in their infancy. The effects of hypernudging and microtargeting have been primarily studied in the context of commercial activity. There is still some significant debate on how powerful algorithmic communications are on social media and within other contexts.<sup>1</sup>

I will start by examining some different conceptions of manipulation. I will settle in [section 1](#) on a sufficient condition for manipulation that is morally salient for my purposes. In [section 2](#), I will lay out how algorithms in political communications can contribute to manipulation and when they don't. In [section 3](#), I will discuss some elements of democratic theory that are relevant to thinking about how manipulation can happen and what its moral relevance is. In [section 4](#), I will lay out a simple model illustrating how objective manipulation can occur. In [section 5](#), I will conclude with some implications of the model for thinking about democracy and institutions.

## 1. Notes on manipulation

Here I will start with the idea of manipulation of opinion and discuss the core problem of manipulation in the context of democracy. Since the idea of manipulation itself is primarily an ethical concept having to do with the ethical motivations of the putative manipulator, it is not primarily suited to the democratic context. I will try to develop a concept of objective manipulation that can serve us for the democratic purpose, but we need to explore the notion of manipulation to get there.

One prominent notion of manipulation is that *manipulation* involves influencing a person in a way that does not sufficiently engage that person's rational abilities (Sunstein 2016, 82). This is an interesting idea, but it does not strike me as compelling because it seems clear that people often engage in all kinds of bad forms of reasoning while thinking that the reasoning is not bad. One person may do this and persuade another and thus, in some sense, not sufficiently engage the other person's cognitive faculties, but the person may be persuaded without knowing this or having any inkling of it. It does not strike me as manipulation in this case.

Another notion may be,

*Manipulation as undercutting reason:* A manipulates B if A influences B toward changing B's mind by doing something that undercuts or subverts the operation of B's rational abilities. (Wood 2014, 289)

This thesis is problematic because it is not obvious that manipulation generally subverts or damages someone's rational abilities. Sometimes it does; sometimes it doesn't. What manipulation often does is to in some way make use of someone's weakness in thinking. It need not damage it or subvert it.

---

<sup>1</sup>See Benkler, Faris, and Roberts (2018) for the argument that algorithmic communications do not yet have an outsized influence on politics. They argue that polarization seems to be occurring much more on the right-wing part of the spectrum of views and that this is among people who do not use social media very much.

A third conception of manipulation based on Coons and Weber (2014) is:

*Manipulation as indifference to reason:* A manipulates B into believing  $p$  or doing  $d$  when A influences B to believe  $p$  or do  $d$  by certain means and A is indifferent about whether B comes to the belief or action in a justified way.

This definition cannot be right because presumably one can regretfully manipulate someone. In that case, one is not indifferent about whether the other person comes to appreciate the reasons for  $p$  or  $d$ , one simply wants B to change his mind and does not expect that giving the reasons will work. One would prefer to operate through reason, but one values the change of mind more.

Another tough case for this conception is the case in which A gets B to believe  $p$  or do  $d$  by giving B an excellent justification (by both A's and B's standards) for believing  $p$  or doing  $d$  but A is willing to influence B without a good justification, at least for the purpose of changing B's mind. A might have done something quite different if A could have changed B's mind or action in some entirely nonrational way. A is a manipulative person in this example, but it is not clear that A is manipulating B. We might say in this latter case that the manipulation is benign, and we might also say that this is not a case of wrongful manipulation.

I am not sure we are going to be able to get necessary and sufficient conditions for manipulation or wrongful manipulation. I want to try to determine a sufficient condition for manipulation that could help us here. This might be helpful because it might suggest that though there are some wayward cases, the central cases are ones in which reason is in some way importantly compromised. This might be especially important if the central cases are the ones that are connected with certain fundamental interests. Hence the sufficient conditions will be morally salient conditions.

Here is the idea:

*Wrongful manipulation as knowing use of flawed reasoning:* A manipulates B when A influences B in the direction of believing  $p$  or doing  $d$  by influencing B to believe  $p$  or do  $d$  by means of a flawed process of reasoning, which A knows works because B is unaware of the flaw in the process or is unable to rectify it. There must be some conflict of purpose here between A and B such that without the flaw, B would go in a different direction that goes against A's purpose.

The flawed process could be that one of the premises is false and A knows it, but B does not. It could involve emotional manipulation where emotions of fear or anger or pride interfere with B's reasoning. It could involve a hasty inference that B does not notice, such as affirming guilt by association. Emotional manipulation does involve subversion of the rational abilities of B, but merely relying on B's not knowing the falsity of a premise or the hastiness of an inference does not seem to me to be subverting or undercutting B's reasoning, but rather using a weakness in it. A takes advantage of the flaw in B's capacities. Taking advantage of the flaws in B's rational capacities means that if B were to approach the matter with a great deal of time and without flaws, B would arrive at a different conclusion or intention that A does not want.

Objective wrongful manipulation takes place when one person or group sets in motion processes of influencing people's minds that take advantage of flaws in the recipients' rational capacities. Minds are changed in such a way that the appreciation of reasons is set back either intentionally or at least is welcomed. Furthermore, the manipulation is directed impersonally at many, though only some are successfully manipulated. The manipulative activity is probabilistic and may manipulate only with a low probability. As long as the harvest is sufficient, it is worth it even if only a fraction of all targeted people is affected. Another condition of objective manipulation is that the manipulated person is not responsible for being manipulated. If a person has all the resources in the circumstances to avoid manipulation but is reckless or foolish, we are not talking about objective manipulation. This notion is important for political equality,

which is primarily concerned with establishing the conditions which enable people not to be manipulated.

From a democratic standpoint, there are three problems with manipulation so understood. First, manipulation appears to deprive a person of what is important or undermines a person's ability to think on their own to a result that can be justified to them. Second, manipulation seems to substitute one's person's justified aims or beliefs with another's. To the extent that each person has an interest in coming to their own understanding of what they ought to believe or do, wrongful manipulation looks like it sets back the interest of the manipulated person by subverting their reasoning or depriving them of important tools of reasoning, and, in the second case, sacrificing that interest for the sake of someone else's interests. Third, manipulation can undermine the important democratic aim of getting people to have an appreciation of other people's interests. The threat to political equality arises here when there is a systematic sorting of persons into groups of persons engaged in manipulation and those effectively subjected to it. Such a systematic sorting implies the setback of the interests of members of the subject group in favor of the interests of the perpetrators.

A kind of case of possibly wrongful manipulation I will not pursue here is the kind in which a person willfully leaves himself open to manipulation; he has all the opportunity in the world to avoid it and does not. This strikes me as an interesting case for democratic theory. A person may have all the cognitive ability to avoid getting carried away by some rhetorically highly charged appeal but may simply allow themselves to be carried away. This may involve wrongdoing on the part of the manipulator and the manipulated. It is a problem if it occurs systematically in a society in that it sets back the deliberative potential of a democratic society and damages the epistemic value of democratic processes, but I do not think that it is a threat to political equality.<sup>2</sup> And it is this latter threat that I want to focus on in this paper.

## 2. Algorithmic manipulation

The processes of hypernudging, microtargeting, and filtering are not necessarily cases of objective manipulation, but they are the main instances of manipulation that scholars discuss when they think of algorithmic manipulation.

Let us discuss the idea of nudging before we approach how it works in algorithmic communications. Discussions of nudging have occurred primarily in three different contexts: government, commercial, and political. The initial impetus of discussion derives from Richard Thaler and Cass Sunstein's work defending nudges for paternalistic reasons (2009). One classic case consists in setting a default for a choice on whether or not to acquire a pension plan. The traditional default is that if one makes no choice, one gets no pension plan. Sunstein and Thaler argue that this greatly increases the number of people who have no pension plan. They argue that one ought to set the default to having a pension plan and allow people to opt out if they wish. This has the effect that many more people acquire pension plans. The basic character of a nudge here is to design the choice architecture (or the way the choice is presented and structured) in a way that does not change the options or the costs of the options but nevertheless influences how the choice is made. The important point here is that not everyone's choice is influenced. Those, for instance, who clearly desire not to have a pension plan will choose not to have the plan whatever the default is. But the nudge makes a difference because there are many who will simply not make a choice out of inertia or laziness or some other nonrational process. We can say that they select the default. How the choice is structured will influence what they end up having. Intentionally designing the choice to ensure that many people will select a certain outcome because they are known to act irrationally in this context does involve a kind of manipulation of those persons who are so influenced (Bovens 2008).

---

<sup>2</sup>I thank Marc Fleurbaey for getting me to appreciate this distinctive kind of case.

Still, this is not a case of wrongful manipulation in most cases, since the choice made concurs with what the person would choose in more favorable circumstances (such as, having plenty of time and unmarred processes). There is clearly some residual problem here, but it does not rise to the level of wrongful action. Moreover, the very same structure does not manipulate others who make deliberate choices no matter what the choice architecture. Many choose the default because they like the default, but they are not the ones who are manipulated. Only those who are expected to select the default out of inertia are being manipulated. Hence, in this context, we see that while nudging may usually have manipulative intent, it does not manipulate everyone. Whether someone is manipulated depends on the person.

Sunstein and Thaler discuss paternalistic nudges. These nudges are meant to be in the interests of the persons being nudged, though a person may not recognize it. But much nudging occurs in commercial contexts in which it is in the interests of the nudgers to nudge and not generally in the interests of the nudged. Here, the aims of the person being nudged and those of the person or group doing the nudging are in conflict and the manipulative character of the nudges is therefore increased.

Politics is an interesting intermediate case. It is always an indistinct mix of conflict of interest and disagreements about legitimate interests and the common good. Activists are trying to persuade you that your legitimate interests will be advanced by a certain candidate or party within a framework of a just set of aims. But, of course, the activists may be mistaken about your interests, and they also act from their own set of interests when they are persuading you. Political persuasion usually presupposes a mix of potential for similarity of interest and aim and conflict of interest or aim. It is fully legitimate when both of these are acknowledged, and it becomes manipulative when the potential for conflict is not acknowledged. It is wrongful when it involves taking advantage of the flaws in rational capacities to achieve an end that the recipient would not endorse were they more fully rational.

What is distinctive about algorithmic manipulation? *Hypernudging* is a form of nudging that continuously reconfigures nudging based on data that an algorithm is getting from you and others and is changing the nudge depending on information and the algorithm in play (Yeung 2017; Lanzing 2019; Danaher 2017). An easy example of a hypernudge is the Google map directions given when a person is already driving toward some goal. The system tells one the fastest way to a destination, but it can recalculate the directions if the person makes a wrong turn or goes too far on a particular route. A recommender system also has this characteristic. Every time someone buys a book on Amazon, the system recommends other books they might like given what they have bought in the past and given what others who have purchased the same book have additionally bought. The system continually adjusts as it is given new information about the person being nudged.

*Microtargeting* involves using algorithms to target messages to certain groups of persons who would be particularly likely to be moved by those messages given all the data collected about what they and others like them would be moved by (Benkler, Faris, and Roberts 2018). Microtargeting can be highly personalized in the sense that the algorithms applied to the data tells one what kind of message to send a very small group of persons who have the relevant traits. It can be used in get-out-the-vote campaigns. It can be used to shape the development of opinion in a particular direction.

Microtargeting can be manipulative in many ways. It can be highly emotional, triggering anger or fear; it may involve half-truths, outright misinformation, or hasty inferences.

*Digital gerrymandering* is a kind of targeted hypernudging. Here one can try to nudge people with a certain set of characteristics to go out and vote or engage in other political activities (Zittrain 2014).

We should note that these activities are not always manipulative. For example, hypernudging, in the case of the expert-recommendation algorithms, does not seem manipulative when it is being used on reasonably well-informed people. When I buy books on Amazon, I receive a list of recommendations for further books. They are described as books that other people, who have bought the book I am buying, have bought. Sometimes I see an author I might have not have

thought of and that might make me think about purchasing the book. But usually, the recommendations are too far from my interests or I have already read them. Sometimes a book comes up that looks interesting but is not close enough to the work I am doing. So, normally, this system has no effect on me, though sometimes it makes a small difference to me.

It seems to me that the distinctive character of algorithmic communications is the sheer scale of the data on which the algorithms can operate, the precision with which they can target people, and the speed with which this can be calculated and constantly updated. Some authors have claimed that hypernudging is especially opaque—that it makes the recommendation from a large body of data acquired by surveillance and operates on a weakness in people's cognitive systems to get them to do things (Yeung 2017; Danaher 2017 Lanzing 2019). But I am not sure how distinctive these features are. After all, I am not sure how people come up with the advertisements I am confronted with on television or the newspapers or how a bookstore owner features certain books for everyone to see. Of course, I have hypotheses, but the same is true regarding the expert-recommender systems. To be sure, the way algorithms work is quite obscure to me and the way complex algorithms do their jobs is obscure to everyone, including the programmers. But is this opacity distinct in quality from the opacity of how advertisers or campaign professionals design election campaigns? Nor am I convinced that algorithmic communications must rely on weaknesses in the cognitive system. This depends on the algorithms and the people using them. To be sure, expert-recommender systems rely on the fact that I can be moved by an *availability heuristic* (that is, that I choose based on the most recent evidence made available to me). And other messages rely on confirmation bias to ensure that people are better persuaded. But one could microtarget people in a way that enhances their ability to deal with new information. For instance, one could have algorithms that ensure that people are confronted with contrary points of view. The fact that they often are not used for this purpose is a function of a polarized environment, not the algorithmic method of communications. And that polarized environment is having its effect on television news broadcasts and newspapers as well. It is probably also partly a function of the fact that the main use of algorithms has been for advertising, which is not particularly interested in exposing people to contrary ideas. Finally, I am not convinced that surveillance is distinctive here. Observing people's behavior is an essential part of political activism and of advertising. For example, in politics, door-to-door or telephone campaigners use information about the people they have visited or spoken with to determine who should be recontacted to remind them to go to the polls. Voters for opponents are not recontacted. The surveillance is far less intrusive, but it is there.

The major differences in all these areas have to do with the size of the data on which communications are based, the precision with which the targeting of messages is achieved, and the speed with which all of this is done. This may exacerbate the problem of manipulation, but it is at least conceivable that the problem could be partly alleviated by big data, precision, and speed. After all, big data, precision, and speed may help find and target people who could potentially be persuaded to change sides.

Let us return, however, to the potential for manipulation in the commercial case. The intention behind the expert recommendation is to get me to buy more books that the people who set up the recommendation system want me to buy. So, it is designed to appeal to me in a way that is meant to get me to act. The algorithms that are used to establish these recommendations are quite opaque to me, and the net effect is to provide a very modest assistance in finding books that might be helpful to my projects.

It is not obvious to me that this is manipulating me in any way, as it stands. At best, it is providing me with hints for what I might want to look at; at worst, it is wasting my time. In this example, it seems to me to realize a kind of mutual advantage: the recommender is giving me hints for further things I might be able to use, while increasing the probability that I might buy them. It must be carefully calibrated to my interests, otherwise it has no impact and is wasted effort. The availability heuristic has extremely limited significance when I am informed about the literature. But since I am experienced in looking for books, it only has an impact when I am not informed about the literature.

Furthermore, the recommendations can have a very modest benefit, so it is hard to think of this as a case of manipulation. I am putting privacy concerns to the side here; I am simply concerned with manipulation.

There are two remaining concerns here that are of interest. One, the recommender system may have a manipulative impact on people who are not well informed and who must make decisions quickly. Two, there may be an aggregate effect of importance even in the case of relatively benign influences. If the effect of the recommender system is to increase the probability by a very small amount that I purchase something, that probability will translate into significant numbers when millions of similarly situated people are affected. If the probability is .001 percent that I purchase something the recommender wants me to buy, then when 10 million people are approached, ten thousand purchases will occur. This is an interesting effect that we need to keep track of in democratic theory.

Take the first concern. Change the case. Suppose I believe that I am in urgent need of medical treatment necessary to keep me alive. Suppose further that I know very little about medicine. And suppose I need to decide very quickly. I am thinking here of a privately run system, though it might be subsidized. An expert-recommender system might be quite manipulative in this context. The combination of ignorance, necessity, limited time, and anxiety might make the availability heuristic and emotional framing very powerful in this circumstance. The expert recommender may nearly determine my choice. That will depend on several things about me and the circumstance but, in this case, the fact that we are dealing with an interested party who very much wants me to adopt a form of expensive treatment and I am very vulnerable to their suggestion makes me highly vulnerable to manipulation. And the expert-recommender system may be very good at figuring out what I am likely to respond positively to.

But we need to note here that even the circumstance described above is not sufficient for manipulation. People do sometimes face these kinds of medical decisions. We usually think that a doctor in this circumstance can present the options in a way that is not designed simply to advance the interests of the doctor or others and that is well calibrated to my needs. A highly conscientious doctor or medical team can present alternatives and predictions in ways that are not manipulative, even though the decision may not end up being fully rational. And an expert recommender could conceivably be set up in a similar way.

But we are in a situation where someone with a desire to sell the most expensive product to me could take advantage of my ignorance and the vulnerability of my cognitive system. In this case, the interests and wills of the influencer and influenced are not aligned and the influencer may well be willing to do quite a bit to get the influenced to make the purchase.

Acts do not in themselves manipulate people. This is the case even when the person performing the action is acting in a manipulative way. The reason for this is that whether an action manipulates a person depends on the condition of the person who is supposed to be manipulated. Well-informed people are not easily manipulated. Cautious people are not easily manipulated. In some cases, an action that manipulates one person can be helpful to another. System 1 (the fast and frugal system of decision making in Kahneman [2011]) may be activated in people who are not paying much attention. But even emotional appeals can be illuminating for persons if they activate an empathetic response. We see this kind of distinction working in the legal apparatus of advertising. Advertising to children is subject to more restrictions than advertising to adults. Presumably this is because it is thought that children are more susceptible to manipulative advertising than adults. Adults, in significant part, seem to realize that advertising is not to be taken at face value. So, they are not manipulated by it, at least not to the extent that children are.

A further point worth observing here is that manipulative actions are not always wrongful even when they succeed in manipulating someone. At least some forms of paternalistic manipulation can be acceptable when the manipulator is highly attuned to the interests of the manipulated and when the manipulated can be expected to act irrationally or imprudently (but not for moral reasons) and there is a particularly simple and noninvasive form of manipulation. Some of the nudges that Thaler

and Sunstein (2009) describe, such as setting defaults, can fit into this category. But it is not my concern in this paper to pursue the issue of justified paternalistic manipulation.

### 3. Democracy and manipulation

From the standpoint of democratic theory, the possibility of manipulation presents several important challenges. First, it can undermine deliberative processes and it can set back the understanding people have of each other's interests and the ways in which one can advance them. Second, and this will be my focus in what follows, it opens an important potential source of inequality among citizens. The reason for this is that in democratic politics, citizens must operate based on low-information rationality. In modern democracy societies with complex divisions of labor, most people have only a small budget of time, money, and energy to expend on politics and so must conserve on information gathering costs (Downs 1957). Low levels of information are part of what make persons susceptible to manipulation through microtargeting or hypernudging. In the medical case, people like me are easily manipulated into thinking that a kind of medical treatment is necessary because we don't know much about medicine and because we are afraid of the consequences of making a bad decision. What saves most of us is not our understanding but the network of people within which we find ourselves and the knowledge that we are in such a network. We can get opinions from other, better-informed people. And, of course, the medical profession itself, to a degree that no doubt varies among doctors, practices restraint when it deals with uninformed and frightened people.

We live in a society in which some people are massively less well informed than others about politics. And we live in a society in which the social conditions under which people can become informed about politics are very unequal in value. In other words, we live in a society in which the social conditions that sustain cognitive ability in politics are quite unequal.

Let me say a few things about the democratic theory I am employing as I think this through. The first major idea is that the fundamental principle behind democracy is the principle of political equality (Christiano 2008). This principle asserts that persons are to have equal shares of effective political power over the society they live in and that inequality is defensible only when it ensures that the less powerful are more powerful than they otherwise would be (Christiano and Braynen 2008). This notion of political power can be broken up into two components: one, resources for participating in collective decision-making, including resources for voting power and the ability to participate in democratic deliberation; two, informational power.

I want to focus on informational power here, on the assumption that equal voting rights and rights of expression and association are secured as well as rights to run for office in a representative system, which is what I call *minimally egalitarian democracy*. It is the lack of informational power that makes people vulnerable to manipulation. And they are vulnerable to people who have a lot of such power.

The significance of informational power can be seen when we contrast the world we are in with a world in which everyone is completely informed about the alternatives. There we are informed about who is promoting which alternatives, and everyone has highly reflective conceptions of their interests and the common good as well as how politicians and policy alternatives relate to them. In such a world, equal voting power, opportunity to run for office, and rights of expression and association would be sufficient for complete political equality. In the world we live in, people are initially completely uninformed and can hope to be at best modestly well informed. Information dissemination and reception are costly and difficult. And this is the source of the problem of achieving political equality in democracies that assign the political liberties to persons.

There are two dimensions of informational power. The first dimension has to do with the ability to understand information and seek out intelligible information about the way the political system can advance one's interests along with the common good. This requires resources for developing one's conceptions of one's interests and the common good as well as basic resources for assessing

whether the society is pursuing these aims—that is, which politicians and parties are committed to pursuing them and whether they are pursuing them through effective policy. Both of these abilities are necessary to an adequate conception of political power. To have voting power and power to participate in deliberation without the cognitive conditions is like having a car without knowing how to use it or any conception of where to go. You have a resource that gives you a kind of impact, but you do not have power to pursue your aims. You may be able to harm other people with this power, but you are not reliably able to advance your interests or moral aims.

Cognitive conditions are meant to enable people to do two things: fulfill their natural duty of justice by attempting to make society as just as possible by their lights, and make sure that their interests are properly accommodated within the society. Hence, people need to have conditions under which they can advance justice by understanding the interests of people in different parts of the society, properly situate their interests within this framework, and do these things in part by ascertaining that the political system is pursuing these concerns.

The other dimension of informational power is the power to disseminate information. Disseminating information involves election campaigns, media broadcasts, and—the thing we are most interested in here—algorithmically determined communications. Each of these activities requires a great deal of funding. One normally cannot run an election campaign without serious financial backing. Large media services are also well-financed operations. And algorithmic services require a great deal of expertise and funding as we saw with the cost of employing the services of Cambridge Analytica. Hence, all these activities will tend to be most well financed by the top income strata of society. The usual view is that the main source of funding for election campaigns comes from the top ten percent of the income distribution (Sorauf 1992) and the top one percent have an outsized influence even here. To the extent that all these activities are expensive, they will be financed by those better off financially. And though the better off are a diverse group of people, their opinions are on average distinct from the rest of the society in a variety of ways. One main example is that the higher-income strata tend not to like much redistribution while the lower-income strata tend to be more in favor of redistribution, at least when they are informed (Erikson 2015). So, if we expect the higher-income strata to be financing algorithmic communications and we expect them to do so with some significant (not necessarily intended) bias toward their own interests, one condition of significant and wrongful manipulation is satisfied: there is serious conflict of interests and aims between those who are financing the algorithmic communications and many of those who are receiving them.

While there is conflict of interest between those who are financing the expertise devoted to algorithmically determined communications and many other people, one essential condition of manipulation is that the receivers of the communications are vulnerable to manipulation. Even highly interested parties are not usually capable of manipulating well-informed persons who are able to tap into networks of persons with a lot of information. The problem in modern society is that the personal and social cognitive conditions are highly uneven. Some have quite good conditions in which they can develop knowledge and understanding while others have lesser conditions.

These social conditions are of the essence. Each person must develop effective understanding while essentially depending on a cognitive division of labor. In order to do our jobs, take care of family and friends, make sure our professions are properly maintained, we can only devote a small amount of time and energy to most of the things we depend on. In our private lives, we depend on doctors, mechanics, financial consultants, plumbers, and many others to know what is necessary to make our lives go well. We depend on a kind of distributed cognition, which is such that much of the knowledge we rely on even in our own decision-making is held by others (Hutchins 1995). For example, when I take medicine prescribed by the doctor, I am usually acting on the basis of good information even though I do not possess most of the information myself. I rely on the doctor's understanding. I do not get a medical degree or even read all the scientific papers connected with the medications. And this is just one of many ways in which we act on the basis of information that is possessed by others. This is the import of distributed information for practical deliberation. It is

essential that I be able to act on the basis of such information I do not possess. At the same time, the complex division of labor in society requires that we depend on others being informed about matters that are important to us and even to our decision-making.

This cognitive dependence is one of the central facts of our private lives and of our lives as citizens. This is the context of low-information rationality that is central to my concern with manipulation here. This cognitive dependence is seen throughout the social sciences in the form of shortcuts, heuristics, and distributed cognition (Lupia 2017). Cognitive dependence makes it the case that our abilities to navigate through life, whether in politics or in private life, are heavily dependent on the quality of the networks of people in which this distributed cognition is occurring. If I live in a society in which doctors, banks, mechanics, electricians, and many, many others cannot be trusted, my interests will be set back and probably the whole society will fail to thrive. This would be a consequence of a kind of massive adverse selection in which people make suboptimal decisions because they cannot depend on others to know about a topic, give good advice, or engage in sound practices.

Cognitive dependence is what makes us potentially vulnerable to manipulation. This is because our cognitive abilities are not very developed to take up the slack when those we might depend on are no longer performing their roles as they ought. And cognitive abilities are quite weak when the social environment on which we depend is not, epistemically speaking, reliable.

#### 4. Simple model

Let me illustrate both the cognitive dependence and differential conditions with a very simple two-group model of the economics of information gathering (ultimately derived from Downs [1957]).

The basic principle of information gathering is that each person collects information up to the point where the marginal benefits of collecting information are equal to the marginal costs. Benefits and costs can be determined by many different concerns. In economic transactions, we might measure the costs and benefits in terms of self-interested concerns. In political life, the costs may be measured in wider terms, including the common good and justice. In the case of an apartment search, for example, you look for an apartment up to the point where the chance that looking at a new apartment will give you something better is sufficiently small that it is not worth expending time and energy to look further. This marginal principle is an expression of the fact that seeking and understanding information is costly and must be evaluated in terms of costs and benefits. Low-information rationality, or rational ignorance, which is quite common in most of our activities, is merely the fact that we must often stop pursuing information even when we are short of being fully informed. We have to do this because the costs to other activities is too great.

This raises the question about how one knows what the marginal benefits and marginal costs of collecting information are when one does not have the information to start with. One does not know what one does not know. This is a key feature of information. The idea is that one starts with free information, or information that one has acquired in some other context and that is available for use here. One has not sought it out with the idea that the marginal benefit from it is greater than the marginal cost. For example, when I talk to my friends, I might, in passing, acquire information about the neighborhood they live in, and that will come in very handy when I look for an apartment myself. The information gathered in casual conversation may prove very useful for my apartment search. Standard sources of free information include education, casual discussions with family and friends, information one acquires from doing one's job, information one has because one wants to know about some subject, and information that may grab one's attention for a variety of reasons.

Education is a clear example of free information. Many children learn things not because they have a sense of how important it is to learn things but because they want to get good grades in order to please their elders. In the absence of this incentive, they would not learn. It is partly from this base of free information that they can go on and pursue more information when they need it. But a solid base of free information is necessary. It is hardly worth repeating the fact that education in many

countries is highly unevenly distributed. This already implies very different social conditions for the development of cognition both in private and political life.

Let's consider another source of free information. Imagine two groups: Employers and Employees. The members of each group interact mostly among themselves except the members of some are the employers of others. Employers are businesspersons, lawyers, professionals, and others who need to know a fair amount about law and politics to know how to do their jobs because law and politics determine the environment in which their work succeeds. And they receive a benefit from success. Employee's work is fairly simple work, which involves little or no engagement with law or politics. Assuming that we are talking about workplaces that are not unionized and in which there is no workplace democracy or rights, employees simply follow orders and ignore the consequences. Their sole concern is the steady flow of wages.

Employers will acquire a great deal of free information from work because they need to know the consequences of what their firms do, and they receive bonuses when things go well. Hence, they acquire sophistication about politics and law. In addition, they also get a great deal of free information from casual discussions with other Employers since they all get high-quality free information from work. They may even develop an informal division of labor in developing ideas. Some learn more about certain aspects of law and politics; others learn about other aspects. And they may learn also from family members, who, let us assume in this example, are also in the Employer class. Employees do not get much quality free information from work. They are not, in this example, members of unions. And let us assume the other people they interact with are in the same position. They don't get much high-quality free information from these sources. They rely on what they hear on television and social media, what grabs their attention, and so on. And they do not have sophisticated tools with which to evaluate these sources. In addition, in the US, Employees are not given the same kind of high-quality education that Employers are given because they live in low property tax neighborhoods.

Notice that free information is not merely new pieces of information, it also includes skills and tools with which to analyze information. In a complex work environment in which one has to deal with many divergent sources of information and disagreements about its interpretation, one acquires the skills to navigate an environment in which there is a lot of disagreement and sources of conflicting information.

Four remarks are necessary here before we proceed. First, the differences in free information and the consequent differences in political sophistication are entirely explained in this model by differences in positions in the division of labor. There is no suggestion here that some are innately smarter than others. Second, I do not mean to suggest that Employees are not able to recognize what their interests are or that they are being set back. Their sophistication in thinking about their interests are diminished somewhat by lower-quality education. But the set back to their information capacities in the workplace is primarily due to the fact that they have less of a sense how the political system might connect with their interests. This latter type of knowledge is essential for politics. Third, this model is highly simplified for illustrative purposes but even here we would not be able to predict that unskilled workers are all unsophisticated. The model only suggests broad tendencies. Finally, Employees may still be able to advance their interests to some extent in the political system if they can identify with a political party that advances those interests; they needn't have much sophistication to do that. But it is well known that party identification is a weaker basis for choosing candidates than more extensive knowledge because politicians are significantly more responsive to those with some policy knowledge (Erikson 2015).

Free information, on this model, is the base from which one then figures out how to acquire more information, what sources are reliable, and how to think about that information. It is often not the end of the process of collecting information, especially for those who have high-quality free information. They can see the need to acquire more information and skills. They can see what they do not know. So, high-quality free information tends to induce a person to become even better informed. Low-quality free information often leaves a person unconcerned with acquiring more

information. They don't have much experience of being rudely disillusioned or finding that their information is not adequate. And they have not acquired the skills to do that.

In relation to Employers, Employees in the model are going to be at an important disadvantage in their abilities to collect further information about politics. They have less high-quality free information and thus less of a basis on which to pursue further information in a discriminating and fruitful way. They are less able to acquire political sophistication and they are less able to learn from each other.

One important consequence of this will involve a greater vulnerability to being swayed by messages that are merely attention grabbing but have little epistemic merit. Furthermore, one is more susceptible to being influenced by microtargeting that sends messages that always confirm one's initial beliefs. They have less ability to discriminate between messages of very different quality than the Employers. Employers will, as a result of the high-quality information they receive as part of their work and as a result of the need to carefully discriminate between what is reliable and what isn't, will have a lesser vulnerability to messages that have little epistemic merit or merely confirm one's biases. No one is invulnerable, but the levels of vulnerability are quite different.

The hypothesis that I am going by here is that low-quality free information and the consequent low-quality ability to discriminate between good messages and not very good messages make those persons highly susceptible to the kind of manipulation that we are concerned with. And the concern is worrisome precisely because of the capacity of groups to microtarget messages to particular groups that have highly emotional framing or that merely confirm initial beliefs. The first way in which manipulative messages can be sent is by sending highly emotionally charged messages to select groups, who are expected to be responsive to them. It works by capturing the attention of persons by highly charged appeals that work on the basic emotions of fear or anger. A second way in which manipulative messages can work is by sending low-quality information to persons that may be misleading or not very relevant to the issues at stake but that give the appearance of such relevance. A diminished ability to discriminate may make it possible for those messages to seem persuasive.

Algorithms enhance the possibility of microtargeting by putting together large amounts of data on persons and then selectively sending messages to people with very particular profiles. They may eventually outperform human beings in selecting targets. But algorithms can involve manipulation because, in the end, the process of selecting targets depends on the set of ends that the programmer set for the algorithm and for the machine-learning process if we are talking about a deep-learning process. This means that the messages are sent in order to promote a particular end, which may or may not be clear in the messages themselves. The manipulation, in this case, can consist of targeting select individuals who are not particularly well informed and sending them messages that may persuade them either by emotional means or by half truths.

## 5. Democracy and epistemic inequality

This model suggests a context in which objective manipulation may take place, and it is an idealized version of circumstances in which individuals in modern societies actually find themselves. Objective manipulation takes place here when agents take advantage of the fact that some people are likely to suffer from serious epistemic weaknesses in the social environment to advance purposes that might not be advanced were the epistemic weaknesses not present.

To the extent that the epistemic weaknesses generated by a particular environment contribute to many people being deprived of important tools for thinking about how to advance their interests or aims, and vulnerable to having their rational abilities subverted by others with different purposes and interests, we have a serious problem for democracy. We have a situation in which the cognitive conditions of citizenship of some people are not well supported by the environments they are in, while the cognitive conditions of others are well supported by the environments they are in. The difference in epistemic conditions makes for a difference in capacities of various groups to advance

their legitimate interests within the framework of a conception of justice and the common good. This creates an important inequality in power over the collective decision-making that violates the fundamental principle of political equality. Objective manipulation is an exercise of that superior power.

We have, on the one hand, inequality in the ability to finance algorithmic communications, which would seem to suggest that these will be primarily dominated by better-off people with distinct interests and aims. And we have, on the other hand, very unequal social conditions under which people can develop the kinds of information and informational skills to help them critically evaluate these communications. Some, in particular the worst-off groups, live in the weakest social conditions for developing political sophistication. So, we have two basic conditions for manipulation in communications: conflict of interests and aims and systematic inequality of political sophistication.

This completes the picture of the possibility of manipulation by means of algorithmic communications in a democracy. What can be done? I have described the ability to transmit information and the capacity to receive and critically assess information as two dimensions of informational power. Those who have more of any one of these have more political power since differences in political power essentially consist of differences in informational power in societies with minimally egalitarian systems of collective decision-making processes. But if one has more on both dimensions, one has a massive advantage and is in a position to manipulate the others.

There are three basic strategies one can envision to try to achieve less manipulation and greater political equality. The first is to require that content providers or the platforms, such as Facebook, use algorithms that ensure that some diversity of opinion gets to each participant. This should not be extremely difficult. This would be some kind of equivalent to the Fairness Doctrine in the United States that was imposed on the networks in the 1950s, '60s, and '70s and was repealed in the late '80s. A lesser remedy might be for the platforms to attempt to give notice when information circulated on the platforms is thought to be highly manipulative and dubious (Cohen and Fung 2021). The second strategy, and the one most pursued in my recent work, is to strengthen the social bases of the informational power of middle-class and lower-class persons (Christiano 2019). The point of this strategy is to enhance the capacities of persons to respond to manipulative efforts. The third strategy is to strengthen the power of a diverse set of intermediate institutions to finance participation in deliberative communications. This is what election campaign finance law is concerned with and could be extended.

I proceed from the basic principle that each person has a right to participate as an equal in collective decision-making and that this right is set back by the conditions I have described. We need to confront this problem on two distinct fronts. We must understand how to make the proper conditions under which persons can acquire political sophistication so they can enhance their political power, and we must make the use of algorithmic communications more democratic.

There are two main ideas for enhancing the epistemic conditions of worse-off persons. An obvious one, though surprisingly not very well pursued in the US, is to improve education in neighborhoods of worse-off persons. The second, more surprising, one is to make the workplace more democratic for all. If Downs's conjecture (1957) that the division of labor poses an important challenge to the ability of those in less powerful parts of it to realize their aims in politics and thereby to achieve political equality, then one important means to improve the cognitive conditions of less well-off people and, thus, to realize political equality, is to change the division of labor. If one of the causes of vulnerability to manipulation is being part of a group with little economic power (which in turn creates low-quality free information), a clear remedy would be to enhance the economic power of this group. This could be done either by empowering unions in the economy or by introducing a significant amount of worker participation in firms. Under these circumstances, ordinary workers would acquire free information in their work because they would need knowledge of the legal and political environment in which work takes place to exercise power in the workplace. This could improve the quality of free information. To be sure, in order for this effect to take place, the unions

or the workplace would have to be genuinely democratic. Democracy in the workplace would also enhance the capacities of persons to engage in discussion and debate over the terms of their work and may enhance their abilities to engage in debate more broadly (Pateman 1970). This would have a positive impact not only on people in the workplace but also on the informal discussions about politics among family and friends, assuming that friends and family are in similar positions in the division of labor. Their discussions can benefit from the kind of sophistication that is needed to participate in union or democratic organized workplaces and, in so doing, they acquire some degree of immunity from efforts to manipulate them.

In this way, there is a kind of complementarity between effective participation in the workplace and effective participation as a citizen in political society.

This approach has substantial empirical backing. It has been observed by a number of scholars that members of unions are significantly better informed on economic and workplace issues than are others who are in the same jobs (Kim and Margalit 2017; MacDonald 2019), and the usual tendency of politicians to be significantly more responsive to wealthier individuals than to others (Bartels 2008) is bucked in the case of representative districts in which there is high union density (Becher and Stegmüller 2020).

With these ideas in mind, I want to argue that each person has a right to the cognitive conditions of effective participation. They have this right because they have deep interests in the kind of political sophistication that enables them to participate more effectively, indeed the very same interests they have in the right to participate in the political system generally (Christiano 2008).

There is a further basis for this positive right. Often, we assign negative rights to persons when they can be held responsible for securing the benefits of the right. For instance, we may leave it to a person to secure the benefits of a proper job once we have made sure that they have the education that prepares them for the workplace. We hold them responsible for the efforts they put out and thus for the exact kind of job they end up with. Their negative right to take a job is protected but, at a certain point, there is no further positive right.

Information has the peculiar property that at some point, we do not know that we do not have information or that we need it and, thus, we are not held responsible for not seeking it out. We do not expect children to get educations for themselves and would not hold them responsible for the failure to be educated. I think something like this structure of reasoning holds for workplace participation and the cognitive conditions of workers. Work is the most important commitment a person has after their families and takes up a great deal of time and energy. If one does not work under conditions that enhance one's information and informational abilities regarding politics, one may well not know what one is missing, especially if all the other persons one knows work under similar conditions. Hence, one cannot be held responsible for failing to have these abilities. But this implies that one may be unable to secure something in which one has a profound interest without workplace participation. Hence, I think we have reason to think that persons have democratically grounded rights to participation in the running of the affairs of their workplaces either as members of democratic unions or in a democratically run workplace.

With regard to the financing of communication and in particular the financing of algorithmically determined communication, we need to enhance the presence and power of working people's associations such as unions (O'Neill and White 2018). These can counteract the tremendous advantage that wealthy groups have in sponsoring communications that attempt to persuade people in society. They can then also undercut the possibilities of manipulation that arise when there are great asymmetries of informational power. A number of proposals have been circulated for increasing unions. One might first provide much more protection for union organizing drives, enabling persons to become members of unions. We know that the great majority of workers in the US do want to become members of unions (Freeman 2007). The fact that they aren't is in significant part a function of a very hostile political environment. Furthermore, some have proposed a system for popular finance of a broad array of associations that can represent the interests of persons by distributing vouchers that can only be used to finance such associations (Ackerman and Ayres 2002).

## 6. Conclusion

I have tried to lay out a conception of objective manipulation in this paper and its relevance to democratic thought and I have tried to show how the increasingly important methods of algorithmically determined communications are playing a role in it. I have tried to show that the problem of manipulation becomes especially important in democracy and what some underlying causes might be. I have also attempted to draw some conclusions about institutional reform in democracy.

All of this is extremely speculative and relies on a still fragile grasp of some of the underlying mechanisms. In particular, algorithmically determined communications are still in need of a great deal of study. This paper is intended as setting a possible framework for reflection on some main issues of political equality that are posed by algorithmic communications.

**Acknowledgments.** I want to thank Marc Fleurbaey, Alex Verhoeve, Kate Vredenburg, Annette Zimmerman, and an anonymous reviewer for this journal for helpful comments on previous drafts of this paper.

**Thomas Christiano** is professor of philosophy and law at the University of Arizona. He authored *The Constitution of Equality: Democratic Authority and Its Limits* (Oxford: Oxford University Press, 2008) and *The Rule of the Many* (Boulder, CO: Westview Press, 1996). He is co-editor in chief of *Politics Philosophy and Economics*.

## References

- Ackerman, Bruce, and Ian Ayres. 2002. *Voting with Dollars: A New Paradigm for Campaign Finance*. New Haven, CT: Yale University Press.
- Bartels, Larry. 2008. *Unequal Democracy*. Princeton, NJ: Princeton University Press.
- Becher, Michael, and David Stegmüller. 2020. "Reducing Unequal Representation: The Impact of Labor Unions on Legislative Responsiveness in the US Congress." *Perspectives on Politics* 19 (1): 92–109.
- Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda: Manipulation, Disinformation and Radicalization in American Politics*. Oxford: Oxford University Press.
- Bovens, Luc. 2008. "The Ethics of Nudge." In *Preference Change: Approaches from Philosophy, Economics and Psychology*, edited by Till Grüne-Yanoff and Sven Ove Hansson. Berlin: Springer.
- Christiano, Thomas. 2008. *The Constitution of Equality: Democratic Authority and Its Limits*. Oxford: Oxford University Press.
- Christiano, Thomas, and Will Braynen. 2008. "Inequality, Injustice and the Leveling Down Objection." *Ratio* 21 (4): 392–420.
- Christiano, Thomas. 2019. "Democracy, Participation and Information: Complementarity between Political and Economic Institutions." *San Diego Law Review* 56 (4): 935.
- Cohen, Joshua, and Archon Fung. 2021. "Democracy and the Digital Public Sphere." In *Digital Technology and Democratic Theory*, edited by Lucy Bernholz, Hélène Landemore, and Rob Reich. Chicago: University of Chicago Press.
- Coons, Christian, and Michael Weber. 2014. "Introduction: Investigating the Core Concept and Its Moral Status." In *Manipulation: Theory and Practice*, edited by Christian Coons and Michael Weber. Oxford: Oxford University Press.
- Danaher, John. 2017. "Algoocracy as Hypernudging: A New Way to Understand the Threat of Algoocracy." *Philosophical Disquisitions*, <https://philosophicaldisquisitions.blogspot.com/2017/01/algocracy-as-hypernudging-new-way-to.html>.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Erikson, Robert S. 2015. "Income Inequality and Policy Responsiveness." *Annual Review of Political Science* 18: 11–29.
- Freeman, Richard. 2007. *America Works: Thoughts on an Exceptional US Labor Market*. New York: Russell Sage Foundation.
- Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kim, Sung Eun, and Yotam Margalit. 2017. "Informed Preferences? The Impact of Unions on Worker Policy Views." *American Journal of Political Science* 61 (3): 728–43.
- Lanzing, Marjorlin. 2019. "'Strongly Recommended' Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies." *Philosophy and Technology* 32: 549–68.
- Lupia, Arthur. 2017. *Uniformed: Why People Seem to Know So Little about Politics and What We Can Do about It*. Oxford: Oxford University Press.
- MacDonald, David. 2019. "How Labor Unions Increase Political Knowledge: Evidence from the United States." *Political Behavior* 43: 1–24.
- O'Neill, Martin, and Stuart White. 2018. "Trade Unions and Political Equality." In *Philosophical Foundations of Labour Law*, edited by Hugh Collins, Gillian Lester, and Virginia Mantouvalou. Oxford: Oxford University Press.
- Pateman, Carole. 1970. *Participation and Democratic Theory*. Cambridge: Cambridge University Press.
- Sorauf, Frank J. 1992. *Inside Campaign Finance: Myths and Realities*. New Haven, CT: Yale University Press.

- Sunstein, Cass. 2016. *The Ethics of Influence: Government in the Age of Behavioral Science*. Cambridge: Cambridge University Press.
- Thaler, Richard, and Cass Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth and Happiness*. Rev/exp. ed. New York: Penguin Books.
- Wood, Allen. 2014. *The Free Development of Each: Studies on Freedom, Right, and Ethics in Classical German Philosophy*. Oxford: Oxford University Press.
- Yeung, Karen. 2017. "Hypernudge: Big Data as a Mode of Regulation by Design." *Information, Communication & Society* 20 (1): 118–36.
- Zittrain, Jonathan. 2014. "Engineering an Election." *Harvard Law Review Forum* 127: 335–41.