PUBLIC REASON

Journal of Political and Moral Philosophy

PUBLIC REASON

Journal of Political and Moral Philosophy

Public Reason is a peer-reviewed journal of political and moral philosophy. Public Reason publishes articles, book reviews, as well as discussion notes from all the fields of political philosophy and ethics, including political theory, applied ethics, and legal philosophy. Texts from other areas of philosophy may be published, if they are relevant for issues in moral and political philosophy. The journal encourages the debate around rationality in politics and ethics in the larger context of the discussion concerning rationality as a philosophical problem. Public Reason is committed to a pluralistic approach, promoting interdisciplinary and original perspectives as long as the ideals of critical thinking and clarity are respected. The journal is intended for the international philosophical community, as well as for a broader public interested in political and moral philosophy. Public Reason publishes two issues per year, in June and December. Public Reason is an open-access e-journal, but it is also available in print. Public Reason is an open-access journal, which does not charge authors a processing fee – neither directly nor indirectly (where applicable) through the author's institution of affiliation. Hence, publication-related decisions are not made on financial considerations

PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 15, No. 2, 2023 & Vol. 16, No. 1, 2024

ARTICLES

Introduction

Jakub Szczepanski
Jagiellonian University

This volume contains texts presented at the Summer School of Political Epistemology, organized by the Kantian Political Thought Standing Group of the European Consortium for Political Resarch ECPR), held on July 27-31, 2020. The school was co-convened by Keele University and Jagiellonian University, jointly with the ECPR. The purpose of the meeting was to give scholars of varying degrees of academic careers the opportunity to meet and discuss key methodological issues in normative political theory, in particular in political epistemology. During the school, lectures were given by accomplished scholars, and papers were also presented by doctoral students and scholars with doctoral degrees.

Presentation topics included the role of knowledge and justification in politics; the problem of deep disagreement; epistemic injustice; democracy and its problems; the role of theory in politically relevant epistemic processes; constructivist and contractualist accounts of justice; the role of sincerity and trust in politics; the epistemic value of electoral processes; the use of ignorance in political processes, including populism, propaganda and manipulation; uncertainty and freedom of speech.

Among the many papers presented, this volume includes articles by authors conducting research in many different countries (Estonia, Germany, the United States, Canada, the United Kingdom) and who are at different stages of their research careers.

What exactly are the topics covered by the texts presented here? First of all, if we are talking about political epistemology, then the question of truth undoubtedly becomes one of the primary issues. With regard to this question Wes Siscoe, in an article entitled: "Epistemic Democracy and the Truth Connection," asks about so-called epistemic democracy and its relationship to truth. In the theory of democracy of this type, it is assumed that the role of truth is very significant. Underlying such a claim is the belief that democratic institutions do truth-tracking better than other political institutions. This belief, called Truth-Tracking, in its turn, has to do with the thesis that in a democracy, a citizen is expected to justify his or her own beliefs, even if this justification is purely internal. This kind of reasoning, in its turn, is what Truth-Tracking requires. In the author's view, however, there is a problem related to the question of the equivalence of a justified belief and one that strives for truth. One does not necessarily correspond to the other. At the level of deliberation, Justification does not necessarily bring us closer to the truth. And here, an important question arises: Should we prefer Justification or Truth-Tracking in the face of deliberative democracy? If the truth in a community begins to be hidden and manipulated, the need for whistleblowing arises.

The question of the role and meaning of this phenomenon is answered by Linyu Jing in a text entitled: "Speaking Truth to Power? A Foucauldian Theory of

Whistleblowing in A Nihilistic World." According to the theory described by the author, a whistleblower is focused on "speaking truth to power." As is well known, power has an indelible tendency to manipulate the truth and treat it in an instrumental way. Whistleblowing can thus be seen as a "practice of freedom", in the sense of normative engagement in political discourse, aimed at "recreating" truth-telling subjectivity.

But the problem arises if the truth we seek does not exist. The pluralist approach and the related conflicts are described by Manuel Knoll in the text "The Significance of Deep Disagreements on Justice, Values, and Morals for Political Epistemology." The subject of this article is the extremely important question of the reference point of political epistemology, namely, "knowledge about values, the good, and about what is just and morally right." According to the author, political struggles are part of the fundamental dispute about values, about what is "God" and what is "Devil." Following Max Weber, the author states that there is no knowledge that can guide "a politician's decisions and actions in the struggle of values and ideals." Isaiah Berlin's value pluralism thesis is also important to the author here. This is how the conclusion of "deep disagreement" is arrived at, as the claim that "there is no possibility to rationally arbitrate between ideals or to rationally resolve value conflicts." Since consensus is not possible, the primary task of political epistemology becomes the search for knowledge, in order to reduce conflicts and enable peaceful coexistence.

Another epistemic problem in modern politics is the difficulty of defining basic concepts with pragmatic meaning. An example of such an issue is the problem described by Jaanika Erne of defining the meaning of the term "democracy." This author, in her text: "Defining Democracy for the European Union", focuses on the limitations of defining the principle of democracy in EU law, where the reference point is the Kantian ontological categorization. The problem, of course, is the lack of a clear definition of democracy, which can lead to a methodological void. What we need, then, are meta-categories, through which it will be possible to correctly determine what the nature of the concept under analysis is. How to realize this postulate? Here the author proposes to capture the EU principle of democracy through a constructed system of seven categories, which allowed various classifications "diachronically and synchronously", while difficulties will be generated by the political nature of the concept. This is because the questions "Which?", "Where?", "When?", "Who?", "What?" and "How?" lend themselves to empirical investigation. Only the question „Why?" will unveil the explanatory dimension.

One significant problem in political epistemology is raised by the question of the possession of knowledge in the context of the choices made. Kristoffer Ahlstrom-Vij, in his text: "Political Knowledge: Measurement, Elitism, and Dogmatism," shows the importance of the question of measurability of knowledge so understood and then its significance in political elections. This is because political knowledge can be seen as a certain resource: having a large amount of such a resource means that we are able to do well in the political world and, consequently, that we are more likely to combine

our political goals with effective means. To demonstrate the difference of having such a stock of knowledge, so-called counterfactual modelling is used. Faced with the objection that scales of knowledge determine certain assumptions, it can be answered that even if political knowledge understood in this way gives rise to dogmatism, this dogmatism can be expected to perform a protective function in the hostile epistemic environment that the political domain represents.

In addition to the primary epistemic issues related to the role and importance of truth in politics, we also have other issues, such as the question of social justice. Questions in this area are answered by John E. Roemer in the article: "Epistemological Issues in Equal Opportunity Theory." The article's author poses the question: what is the optimal policy of equal opportunities? The general answer to this question is that it is a policy that will compensate disadvantaged types with resources that will improve their income distributions. Romer's Equal Opportunity Theory is supposed to answer this question precisely. Its main application is the fair treatment of individuals competing for desirable positions in society.

As can be seen from the preceding brief summary of the articles included in the volumes of this special issue, political epistemology touches on a number of extremely important issues. By taking up epistemic considerations, we situate ourselves at the same time on a meta-level. The range of topics that can arise in this field may therefore also be discussed when considering the relation of this field with others. Since politics and political philosophy in general open up a wide range of considerations, an investigation of the epistemological dimension of these considerations will also have wide scope, as shown by the wide range of topics presented here.

I thank Sorin Baiasu for the co-convening of the Summer School and the ECPR for looking after the administrative aspects of the event and for helping us to organise the event.

j.szczepanski@iphils.uj.edu.pl

Epistemic Democracy and the Truth Connection

Robert Weston Siscoe
University of Notre Dame

Abstract. If political decision-making aims at getting a particular result, like identifying just laws or policies that promote the common good, then political institutions can also be evaluated in terms of how often they achieve these results. Epistemic defenses of democracy argue that democracies have the upper hand when it comes to truth, identifying the laws and policies that are truly just or conducive to the common good. A number of epistemic democrats claim that democracies have this beneficial connection to truth because of the type of deliberative environment created by democratic political institutions. Democratic political cultures encourage giving and exchanging reasons, ultimately improving the justification that citizens have for their political beliefs. With this improved justification comes a better chance at truth, or so the story goes. In this paper, I show that attempts to forge a connection between justification and truth in epistemology have encountered numerous difficulties, making the case that this causes trouble for deliberative epistemic defenses of democracy as well. If there is no well-defined connection between truth and justification, then increasing the justification that citizens have for their beliefs may not also increase the likelihood that those beliefs are true, revealing a serious flaw in charting a connection between political justification and political truth.[1]

Keywords: epistemic democracy, justification, truth.

Epistemic theories of democracy hold that democratic political institutions have a valuable connection to truth. While democracies do not ensure that the optimal policy will always be chosen, democratic decision-making increases the likelihood of selecting such outcomes when compared with other political arrangements. This is the first component of epistemic defenses of democracy, that democracies have a truth-tracking feature that makes them preferable to other political institutions:

Truth-Tracking – Democracies track the truth better than other political institutions

This truth-tracking feature of democracies plays out in terms of the laws and policies selected by democratic institutions – these institutions are better, the thought goes, at selecting the best legislation. The shape of such legislation depends on the task at hand. In some contexts, the goal is to select just laws and policies. In others, the aim is to choose legislation that will be most effective in securing the common good. Of course, in many instances, it could be that making laws that promote justice and laws that secure the common good coincide. Truth-Tracking can, in principle, be applied to all of these scenarios – it is possible to ask whether a law truly reflects the principles of justice or truly provides for the common good. The common thread running through

---

Truth-Tracking arguments in favor of democracy is that democratic decision-making is preferable because of its ability to track the truth about these important matters.[2]

It is not by chance that democracies are thought to satisfy Truth-Tracking. Instead, many advocates of truth-tracking theories, like David Estlund and Hélène Landemore, hold that democracies are effective at tracking the truth because a democratic political culture improves the epistemic justification that citizens have for their political beliefs.[3] For this reason, there is taken to be a tight connection between Truth-Tracking and Justification:

Justification – Democratic political cultures create deliberative environments that increase the justification that citizens have for their political beliefs

The thought with Justification is that there are features of democracies that make citizens better justified in their political beliefs, features including cognitive diversity and egalitarian sharing of reasons. A number of truth-tracking accounts of democracy thus have two features – that democratic procedures are more likely to select correct political policies and that this is the case because voters are better justified in their political beliefs. After addressing some preliminary issues in Section 1, in Section 2 I examine Estlund's and Landemore's accounts of epistemic democracy, noting that they emphasize not only the truth of the policies selected by democratic institutions, but also that this occurs due to the justification citizens have for their political beliefs. These accounts take it that, because democracies fulfill Justification, they will also satisfy Truth-Tracking.

In Section 3, I point out that a key assumption of this two-pronged strategy, that there is a well-defined connection between truth and justification, faces serious difficulties. A necessary assumption for such epistemic defenses of democracy to go through is that epistemic justification has a connection to truth. A constituency with better justified beliefs can only more effectively identify the truth if having better justified beliefs in fact makes one better at locating the truth. This is, of course, a very natural thought, that justification marches lockstep with likelihood of truth, but it is a thought that has failed to be substantiated within contemporary epistemology, undermining democratic theories that appeal to the connection between truth and justification. There is thus a substantial difficulty for Truth-Tracking defenses of democracy that also depend on Justification, a challenge which I detail in Section 4.

I conclude the paper in Section 5 by showing that the combination of Justification and Truth-Tracking in many theories of epistemic democracy obscures an important

---

2] For more on how to make sense of the connection between truth and various political policies, see the discussion in Section 1.

3] Discussions of political institutions often debate whether or not such institutions are justified, but that is not the sort of justification I have in mind. By 'epistemic justification', I refer to the notion of justification central to epistemology, the justification of belief.

choice point in epistemic defenses of democracy more generally: Would we rather have laws that we have more reason to endorse or laws that hew closer to the truth? If the arguments of this paper are correct, these two questions can come apart, requiring that political theorists get clearer about what they are pursuing when they advocate an epistemic defense of democracy. Views that focus on the ways in which democracies improve the reasons citizens have for their beliefs might instead have reason to opt for accounts on which democracies select, not policies that are more likely to be true, but policies that are better justified than alternatives.

I. TRUTH-TRACKING CONCEPTIONS OF DEMOCRACY

One issue that is relevant before we begin is that there has already been a fair amount of work criticizing Truth-Tracking. The general thrust of these objections has been that, as a matter of fact, actual democracies do not do a good job of tracking the truth. The reasons for this are myriad: The pressure for consensus drives groupthink rather than critical evaluation, voters have little incentive to be well-informed about political issues, and politicians have more reason to cultivate rhetorical flair rather than political expertise. What these worries all have in common is that they attempt to deny Truth-Tracking by undermining Justification. It is thought that, because of voter ignorance, groupthink, or political incompetence, it is not plausible that citizens have better reasons for their political beliefs.[4] My critique, however, goes further. I argue that, even if the issues that lead to problems with Justification can be resolved, this still will not mean that democratic institutions satisfy Truth-Tracking. My argument instead drives a wedge between Justification and Truth-Tracking, pointing out that even if advocates of epistemic democracy can rebut the criticisms of Justification, this is not yet to argue that democracies satisfy Truth-Tracking, showing that resolving those issues is not the only challenge facing deliberative epistemic defenses of democracy.

The second issue that needs clarification is the place of truth in political decision-making. There are a number of different positions on offer that fall under the banner of epistemic defenses of democracy – what they all have in common is a commitment to an independent standard of correctness for political decision-making. In many cases, that independent standard of correctness is truth, as whether propositions like raising the legal minimum wage increases unemployment or nationalized healthcare reduces the cost of major medical procedures are true is obviously of political concern. But citizens often vote for particular politicians or political policies, things that are not true or false in the same way. It doesn't make sense, for instance, to ask whether a particular political candidate is true or false, nor does it make to ask whether a particular nationalized health care plan is true or false. To speak of these things as true or false is simply a category error.

4] For views along these lines, see Brennan 2016, Caplan 2007, Solomon 2006, and Somin 2016.

In order to deal with this issue, epistemic defenders of democracy have instead emphasized independent standards of correctness. In his seminal work on epistemic concepts of democracy, Joshua Cohen says that, in order to make sense of epistemic arguments for democracy, we need "an independent standard of correct decisions."[5] That is, we need some, procedure-independent standard for judging whether voters are choosing correctly or incorrectly. Truth is obviously a procedure-independent standard, but it is not the only one. Selected politicians and political policies can also, for example, best satisfy collective voter preferences or avoid major harms.[6] Whatever the independent standard we choose, then, allows us to reintroduce questions about truth. We can always ask, for instance, whether it is true that a chosen politician or political policy avoids major harms, re-establishing the connection between truth and political decision-making. In this paper, I will focus on views that are Truth-Tracking in the sense that they hold that democracy makes decision that track with some independent standard of correctness.[7]

A final issue that is worth clarifying is that we will not be concerned with all epistemic defenses of democracy. Amongst Truth-Tracking accounts of democracy that rely on some independent standard of correctness, not all views endorse both Truth-Tracking and Justification. One variety of epistemic defense that does not explicitly endorse Justification, for example, are accounts that rest on preference aggregation. According to this sort of epistemic argument for democracy, democratic procedures are valuable for selecting the best political policy because they reveal what policy best satisfies collective preferences. Even though these accounts may be committed to discovering the truth about the preferences of its citizens, preference aggregation defenses of democracy are not committed to Justification. On this model, it does not necessarily matter how citizens arrived at the preferences they have – what counts is whether democratic procedures are effective for aggregating those individual preferences that already exist. Because aggregation views of democracy do not take up Justification, the arguments of this paper are not directed towards such views.

Another sort of view that lies beyond the purview of this paper are the results touted by defenders of the Condorcet Jury theorem. On political interpretations of the theorem, if independent voters are each more likely than not to get the right answer about a political question, then a large numbers of voters virtually ensures that the best political policy will be chosen.[8] One assumption of Condorcet's theorem, however, is

---

5] See Cohen 1986, 34. For those who echo this position while making the epistemic case for democracy, see Estlund and Landemore 2018, 113, and Goodin and Spiekermann 2018, 17-18.

6] Both Estlund (2008, 160-66) and Landemore (2013, ch. 8) appeal to this second standard of correctness.

7] For the purposes of this paper, I will assume that Cohen's ideas about an independent standard of correctness can come to the aid of Truth-Tracking defenses of democracy, though if it is not possible to bridge this divide, this would pose another serious problem for Truth-Tracking cases for democracy.

8] For a discussion of the original jury theorem and various ways to strengthen its results, see List

that citizens form their views independently of other voters, practically the antithesis of the deliberative environment encouraged by Justification. According to Justification, it is precisely because voters do not form their political beliefs independently that they better track the truth. Thus, even though defenses of democracy that rely on Condorcet do advocate Truth-Tracking, they do not do so because of Justification. If all voters form their beliefs independently, then there is no collective deliberation about political decisions. For this reason, epistemic defenses of democracy that rely on the Condorcet result are also not the target of this paper.[9]

Instead of focusing on preference aggregation or Condorcet-motivated views of epistemic democracy, this paper instead hopes to make a criticism of epistemic defenses of democracy that advocate for both Justification and Truth-Tracking. On such views, it is because democracies create particular kinds of epistemic environments that they satisfy Truth-Tracking, not because democracy is effective for preference aggregation or independent belief formation. The target of this paper then are views that take the middle route to Truth-Tracking in Figure 1, accounts that argue that democracies track the truth because of their unique deliberative environments. For the sake of clarity, I will refer to such defenses of democracy as deliberative epistemic defenses of democracy.

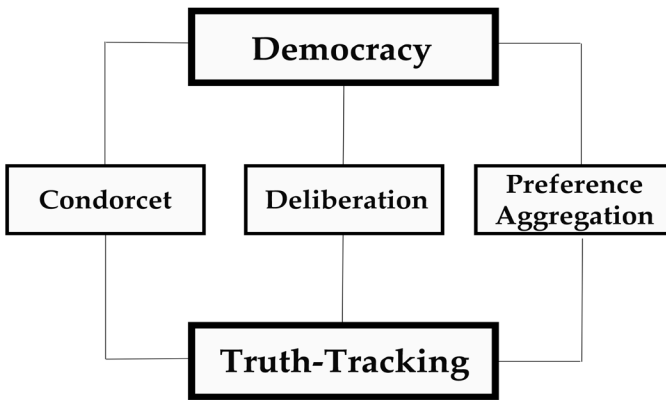*Figure 1:* Deliberative Epistemic Democracy.

DELIBERATIVE EPISTEMIC DEMOCRACY

One recent account of deliberative epistemic democracy that traces a connection between Justification and Truth-Tracking is David Estlund's *Democratic Authority.*

and Goodin 2001.

9] Though for work that modifies the independence assumption in order to account for real world deliberation, see Goodin and Spiekermann 2018, 67-82.

Estlund holds that democracies are preferable because of their effectiveness in identifying the true requirements of justice. This epistemic power is quite modest – on Estlund's view, democracies are (1) better than random when it comes to selecting the true principles of justice and (2) better than other political arrangements.[10] In this way, Estlund's account has a truth-tracking component. Why think that democracy does better than other political arrangements at selecting the true principles of justice? Estlund holds that democracies are effective at approximating model epistemic deliberation in that they promote everyone having equal opportunity for sharing their reasons for their political beliefs.[11] With access to more evidence for and against what they believe, citizens are able to take more into account when forming their political beliefs. Because citizens are able to appreciate more relevant reasons and have better justified beliefs than they would have otherwise, the likelihood that their judgments will lead to the best outcome increases. The reason then that Truth-Tracking occurs is because democracy supports the kind of epistemic environment that also satisfies Justification. Estlund's account of democracy thus ties together these two strands – both truth and justification play a role in his deliberative epistemic defense of democracy.

Another recent example that draws a link between Truth-Tracking and Justification is Hélène Landemore's deliberative epistemic argument for democracy in *Democratic Reason*. According to Landemore, democracies are better than other political arrangements at selecting the best political policies. Like Estlund, Landemore argues that democratic procedures are both better than random at identifying the best political policies and more effective than other political arrangements, including dictatorships, oligarchies, and epistocracies. This, of course, is the Truth-Tracking plank of epistemic defenses of democracy – that democracies have an inside-track when it comes to making the right political decisions. What is it about democracies that give them this epistemic power? Landemore makes the case that inclusive deliberation and the cognitive diversity that comes with it plays an important role in effective group decision-making, the kind of diversity that democracy is well-positioned to capture. Inclusive deliberation has this beneficial effect on group deliberation because a diverse group of citizens bring more to the table, both in terms of potential solutions to political issues as well as reasons for and against adopting different solutions.[12] This exposure to a diversity of reasons and solutions then improves the ability of the group to identify the best solution to political problems. Landemore thus argues that, to the extent that democracies are the most cognitively diverse political arrangement, they also satisfy Truth-Tracking. Here, then, we can see the connection again between Justification

---

10] See Estlund 2008, 98. Estlund also holds that the policy selection procedure must be acceptable to all qualified points of view, a criterion he uses to rule out forms of government that are not justifiable via public reason. The epistemic element of Estlund's view is thus only one component of his defense of democracy. For how the qualified acceptability requirement bears on the arguments of this paper, see Section 5.

11] See Estlund 2008, Ch. 9.

12] See Landemore 2013, ch. 4.

and Truth-Tracking. It is in virtue of the deliberative process produced by democratic political cultures that citizens improve the reasons for their political beliefs, leading to better political decisions.

Even though Estlund and Landemore give different reasons for the epistemic powers of democracy, they both forward deliberative epistemic defenses of democracy. On the one hand, Estlund argues for the Truth-Tracking conclusion because democracy provides all citizens with equal opportunity for sharing their reasons and viewpoints, coming closest to instantiating model epistemic deliberation. Landemore, on the other hand, contends that inclusive, diverse deliberation is what enables democracy to satisfy Truth-Tracking. What they have in common is that they both defend Truth-Tracking because of Justification, that it is the deliberative environment created by democracies that enable them to make the best political decisions. These, then, are the sorts of views that are issue in this paper, views which require a connection between justification and truth in order to make a case for Truth-Tracking.

II. THE TRUTH CONNECTION IN EPISTEMOLOGY

Theories of epistemic democracy accept Truth-Tracking, that there is a close connection between democratic procedures and truth. A number of these accounts, views like Estlund's and Landemore's, hold that the connection with truth is forged due to Justification. On these accounts, democratic procedures give way to better judgments because collective democratic decision-making makes citizens better justified in their political beliefs. There is thus a hope running throughout deliberative epistemic democracy that there is a connection between better justified political beliefs and better policies, that Justification leads to Truth-Tracking. Close attention to recent work in epistemology, however, undermines such optimism. As we will see, the thought that justified beliefs have a clear connection to truth has yet to be substantiated. In this section, I will survey possible routes to securing a connection between justification and truth, ultimately concluding that none of them provide a firm foundation for those who offer an epistemic defense of democracy.

Before we examine these views on the connection between truth and justification, I should note that I do not take the objections listed in this section to be decisive. There are a number of possible avenues for response, many of which I note, and I will not have the space to consider all such possibilities. My intention, rather, is to illustrate the difficulties that have prevented each of these proposals from enjoying wide endorsement within epistemology. Likewise, I do not take the options presented in this section to be exhaustive of the possible links between truth and justification – there are other possibilities, and variations on those possibilities, that will not be surveyed here. Summarizing the difficulties, however, of some of the most prominent proposals attempting to connect truth with justification makes clear why there is no consensus in epistemology that justification is coupled with truth. This fact alone is enough to

make trouble for the deliberative epistemic democrat – until we have a better idea of how justification is connected to truth, we also will not know if Justification actually leads to Truth-Tracking.

Entailment

An early view of epistemic justification that attempted to make the connection between justification and truth explicit was Descartes's view that having a justified belief that *p* entails that *p* is true. The Cartesian thought is that no belief should be taken on unless it is beyond doubt, and a belief is beyond doubt only when it follows deductively from one's evidence. The *cogito* then gains purchase because, regardless of how deceived I am, it is entailed by the fact that I am thinking that I exist. The failure of the Cartesian project, however, came because this standard of justification is far too strong. There are many everyday propositions that we are justified in believing that are nevertheless not entailed by our experiences. It is conceptually possible that many of our beliefs are due to the machinations of a Cartesian demon, but we are nevertheless justified in our ordinary beliefs about the external world. The majority of philosophers have thus taken the lesson from DesCartes that the relationship between justification and truth is something less than entailment.[13]

Process Reliabilism

Because the Cartesian project was ultimately untenable, epistemologists have explored views of epistemic justification on which the connection with truth is weaker than entailment. An influential view in this spirit is process reliabilism. Reliabilism attempts to characterize justification in terms of how well certain processes of belief formation result in true beliefs. For simplicity's sake, let's consider a process reliable if it produces more true than false beliefs.[14] Reliabilism is one strategy for attempting to make the relationship between justification and truth explicit without a Cartesian entailment requirement. On reliabilism, what it is to have a justified belief is to have a belief that was produced by a truth-related process, in this case, a process that produces true beliefs better than half of the time.

Process reliabilism runs into issues, however, because the link it proposed between truth and justification does not always seem to obtain. In some situations, it seems that believers can be justified regardless of the truth-efficacy of their belief forming processes. Consider the following counterexample to reliabilism:

13] It should be noted that Descartes does not explicitly discuss epistemic justification. Instead, he discusses when one ought to assent to beliefs. One promising interpretation of this view, of course, is that Descartes is taking a stand on when one is justified in holding a belief, see Bonjour 2009, 39-40. Even if this is not the best way to characterize Descartes, however, there are nevertheless others who have defended a connection of entailment between justification and truth, see McGrew 1995 and 1998.

14] This is a toy theory of reliability in comparison to Goldman's 1979 original thoughts on reliability, but it should be adequate to point out the difficulties for process reliabilism.

New Evil Demon

Suppose that I have a counterpart in a possible world that is controlled by a Cartesian demon. Everything appears to my counterpart precisely as it appears to me and they have every reason to believe what I do in the actual world. Now, as it turns out, their belief-forming processes are completely unreliable due to the malevolence of the Cartesian demon.[15]

The counterpart is clearly justified in their beliefs, just as justified as I am in the actual world, but none of their belief-forming processes are reliable. The majority of their beliefs are false and always have been, but they are nevertheless justified. The way the process reliabilist conceives of the interface between truth and justification thus seems to be misguided – justification does not require that a belief is produced by a reliable process.[16]

Probability-Raising

In the wake of the challenges to Descartes's necessitation account and Goldman's process reliabilism, there have been other proposals for how to connect justification with truth. One influential thought has been that, instead of the reasons that I have for thinking that *p* necessitating *p*, the reasons for *p* can instead just raise *p*'s probability. A number of epistemologists have endorsed such a conception, with some even taking it to be a fundamental assumption about epistemic justification. Richard Fumerton, for example, says that "whatever else epistemic justification for believing some proposition is, it must make probable the truth of the proposition believed."[17] Following closely on the heels of a probabilistic account of justification is a probabilistic account of evidence. Just like with justification, probabilistic accounts of evidence say that all evidence for *p* makes *p* more probable.[18] The probabilistic conception of justification and evidence is

---

15] This example is due to Cohen 1984, 281.

16] Though the New Evil Demon is widely taken to decisively show that reliable processes are not necessary for justification, dissenters include Bach 1985, Brewer 1997, Engel 1992, Goldman 1986, Littlejohn 2009, and Sutton 2005. Lasonen-Aarnio (Forthcoming) holds that the counterpart in the demon scenario, while perhaps not having justified beliefs, nevertheless has reasonable beliefs. While distinct from justified belief, reasonable belief is extensionally very close to justified belief and thus can account for the intuition that the counterpart is doing something right in the New Evil Demon scenario.

17] See Fumerton 2002, 205. For another formulation of justification in terms of probability, see Steup 2005.

18] Those who defend a positive relevance account of evidence include Carnap (1962), Hesse (1974), Kearns and Star (2009), Kronz (1992), Maher (1996), Roush (2004), and Swinburne (1973). For those who are evidentialists, taking the degree of justification to be solely a factor of evidential support, a probabilistic account of justification follows straightforwardly from a probabilistic conception of evidence. This will not be true for those who think that there are other factors involved in justification – Cohen's (1998) view is that to have a justified belief is to be able to rule out all salient error possibilities (p. 292, fn. 11) or Fantl and McGrath's (2002) view that which beliefs are justified is effected by the stakes of one's practical situation. On

thus deeply ingrained in contemporary epistemology as an attempt to connect truth and justification.

The difficult question for a probabilistic interpretation of justification, however, is how to construe an epistemic notion of probability. If the counterpart in the demon world is justified in their beliefs, whatever notion of probability we give will have to rule that their beliefs are probable given their reasons even though the majority of their beliefs end up being false. One potential thought is that they find themselves in a quite unusual world. Perhaps, for the most part, possible worlds in which agents have all of the counterpart's evidence are like the actual world. This promising thought, however, is difficult, if not impossible, to make good on. To begin with, there are an infinite number of possible worlds, and so we will be appealing to equivalence classes of infinite possibilities. On this way of understanding the problem though, there is no way to capitalize on the thought that p is true in a greater proportion of the worlds in which the subject has good reason to believe p. Worlds where the subject has good reason to believe that p are either veridical or deceptive – in the former case, p is true, and, in the latter, the subject is deceived. Because there are an infinite number of veridical worlds, there is a way to sort the worlds such that, for each deceptive world, it is sorted into an equivalence class with only one deceptive world and an infinite number of veridical worlds. But because there are also an infinite number of deceptive worlds, there is also a way to sort the worlds where, for every single veridical world, there are an infinite number of deceptive worlds. This suggests that the thought that a greater proportion of worlds where an agent possesses good evidence for p are veridical is unfounded. Finding an understanding of probability that tracks epistemic justification is thus a significant challenge.

Perhaps there is an interpretation of the sort of probability at play with justification that can respond the above worries – my goal is not to argue that the objection I have offered is decisive. Rather, my objective is to note the difficulties that have prevented epistemologists from widely accepting any particular account of the connection between truth and justification. Just as necessitation and process reliabilist accounts of this link have serious issues, so do epistemic interpretations of probability. The difficulty runs so deep, in fact, that many epistemologists have simply abandoned the thought that epistemic probabilities are connected to the truth, instead arguing that epistemic probabilities are just subjective degrees of confidence. There is thus far from a consensus that the probability-raising view of justification is correct, and if it is, whether these probabilities have anything to do with truth.

III. THE TRUTH CONNECTION AND DEMOCRACY

---

these views, even though evidential support is not the only consideration when determining the strength of one's justification, the force of one's evidence nevertheless also plays a role in the justification of belief.

So what is the relationship between justification and truth? The challenges facing the preceding accounts – entailment, process reliabilism, and probability-raising – undermine the thought that there is a straightforward connection between truth and justification, and some epistemologists now do not think that there is a way to make sense of justification in terms of truth.[19] Given that there is no widely accepted way to secure the connection between justification and truth in epistemology, it will also be difficult to make sense of the idea that Justification promotes Truth-Tracking. The deliberative epistemic arguments for democracy that we have seen, though, depend on the thought that, by improving the reasons that citizens have for their beliefs, democracies will be better able to track the truth. In this section, we will see that the ways to understand the proposals of deliberative epistemic democrats mirror the proposals that attempt to connect epistemic justification and truth, showing that the difficulty in epistemology carries over to epistemic defenses of democracy.

Estlund and Landemore clearly do not think that democratic procedures guarantee Truth-Tracking, so the Cartesian view is off the table. The other views though, process reliabilism and probability-raising, can make sense of some of the comments by Estlund and Landemore. Suppose, for example, that we interpret Estlund's and Landemore's comments about democracy having a better than random chance of making the correct political decision as advocating that democratic procedures are justified because they select the right answer more than half the time. The best way to understand the epistemic defense of democracy, then, is as a form of democratic reliabilism, that democracies are the best political arrangements because they embody a reliable process, positioning epistemic democrats as depending on the reliabilist proposal as the correct account of the connection between justification and truth.

Just as with process reliabilism, however, democratic reliabilism runs into issues given the possibility of the demon world. The New Evil Demon world demonstrates that there is no straightforward connection between justification-conferring processes and truth. Likewise, a Democratic Demon scenario would be one in which, despite using democratic procedures, no democratic political decisions end up getting the right result because of the deception of a Cartesian demon. Due to open deliberation and cognitive diversity, the citizenry of a demon world democracy would be very well justified in their political beliefs, but that does not also mean that they will have a better than random chance at identifying the best policy. Because they suffer from demon-deception, their beliefs will be mostly false, preventing them from making sound political decisions. Thus, even if Estlund and Landemore are right that democracies satisfy Justification, this does not guarantee that they also satisfy Truth-Tracking.

19] See, for example, Berker 2013a and 2013b and Cohen 1984. How widespread are these concerns? On the most recent, 2020 Philpapers Survey, 36% of philosophers said that they accept or lean towards an internalist conception of justification. Internalist accounts are typically crafted to accommodate cases like the New Evil Demon, and some even sever the connection between justification and truth entirely.

Perhaps, instead of taking Estlund and Landemore to be reliabilists, we should instead take them as thinking that democratic procedures raise the probability that a political decision is correct, an alternative way of understanding democracy's "better than random" chance of getting the right result. Estlund at points appears to explicitly endorse such a conception, that policies chosen by democratic procedure are more likely than not to be true.[20] Like with reliabilism, however, because the probability-raising conception of epistemic justification runs into serious difficulties, it is difficult to make sense of a probability-raising conception of democratic deliberation. As with epistemic justification, it is challenging to describe how, across infinite possible worlds, democracies are more likely to choose the best policies than not. For this reason, deliberative epistemic democrats should not want to be interpreted as endorsing a probability-raising connection between Justification and Truth-Tracking.

The lesson from epistemological work on justification and truth should now be clear. It is dubious that there is a well-defined connection between truth and justification – likewise, there is a serious concern that there is no such connection between Justification and Truth-Tracking. Deliberative epistemic democrats may be right about Justification, that open deliberation and cognitive diversity provide citizens with better reasons for their beliefs than they otherwise would have had, but this is not enough to also claim Truth-Tracking for democracy.

IV. THE WAY FORWARD FOR EPISTEMIC DEMOCRATS

Given the difficulty connecting truth and justification, how should deliberative epistemic democrats proceed? A few avenues present themselves. Deliberative epistemic democrats could attempt to save the link between justification and truth, singling out a particular strategy for connecting truth and justification. Making this case would restore the close association between Justification and Truth-Tracking, making it viable to argue for both in a defense of democracy. Failing this, however, deliberative epistemic democrats will have to make a choice – what is the more important, foundational political value, Justification or Truth-Tracking? In this section, I will make the case that epistemic defenses should opt for Justification over Truth-Tracking, putting the reasons that citizens have to accept their political institutions at the center of the deliberative case for democracy.

Restoring the Link?

One way to restore the link between truth and justification would be to adopt a knowledge-first account of the relationship between truth and justification. On Timothy Williamson's view, epistemology went wrong when it tried to analyze knowledge. The task, instead, is to understand epistemic notions like justification *in terms of* knowledge.

20] See Estlund 2008, 114.

In keeping with this approach, Williamson proposes that knowledge is what justifies belief, making known propositions the only candidates for justifying a person's beliefs. Such a view, of course, is a factive account of justification. Only truths can be known, so if one's justification consists only of propositions that one knows, then the justification that a person has for their beliefs is completely made up of truths.

At first glance, this seems advantageous for epistemic democrats. On Williamson's way of thinking, Justification and Truth-Tracking cannot come apart – the more truths that citizens know, the more justification they have for their political beliefs. Using a knowledge-first account to restore the link between Justification and Truth-Tracking, however, comes with a significant cost. Because justification is limited to knowledge, the citizens in the demon world hardly have any justification whatsoever for their beliefs. Most of what they think they know is false, precluding them from having justification for the majority of their beliefs, including what they believe in the political realm. Williamson confirms that this is the case by characterizing victims of skeptical scenarios as excused yet nevertheless unjustified in their beliefs.[21] This position is even worse than what we saw in Section 4. While before, Justification could diverge from Truth-Tracking, democracy in the demon world could still increase the justification that citizens had for their beliefs even when it failed to increase the number of truths that they believed. On the knowledge-first view, however, inhabitants in the demon world are deprived of both Justification and Truth-Tracking, stripping away any case that democracy can guarantee improved epistemic results.

Another possibility for restoring the link between truth and justification would be to ask why we should be concerned with demon worlds or infinite sets of possible worlds at all. What counts – what matters to Estlund and Landemore – is whether democracy is truth-tracking in the actual world. Why, then, do we need to consider the epistemic merits of other possible worlds?

While promising at first glance, this line of response also ends up being unsatisfactory. To begin with, proceeding along these lines makes the deliberative epistemic case for democracy merely contingent – only in worlds like the actual world is democracy the best form of government. It is not clear, however, that deliberative epistemic democrats would be satisfied with a merely contingent defense of democracy. Take, for instance, worlds in which other political institutions come closest to tracking the truth. There have been political states that are founded on the premise that our best access to truth is via divine revelation, truths that cannot be established by human reason. According to these political arrangements, the best epistemic form of government is rule by those who have the best access to God's revelation.

21] See Williamson (Forthcoming). Clayton Littlejohn, another recent advocate of the factivity of justification, stumps for the same view, that the deceived are excused but not justified, see his 2012 and Forthcoming.

In the possible worlds where the best access to truth is in fact through divine revelation, I doubt that epistemic democrats would be open to defending theocracy. This might be because they think that democracy has other benefits that theocracies would not obviously offer, including universal political participation and freedom of thought and conscience. But this presents a dilemma. If democracy is still the best political arrangement even when it is not epistemically beneficial, then it becomes difficult to see what role its supposed epistemic benefits play in its justification. What started out as a concern about the mere contingency of democracy's epistemic benefits thus turns out to be a challenge to the Truth-Tracking strategy writ large. If democracy is the best political arrangement regardless of how well it performs in terms of Truth-Tracking, then Truth-Tracking does not seem to play an essential role in selecting democracy over other political institutions.

Justification or Truth?

I seriously doubt that deliberative epistemic democrats would be satisfied with a merely contingent defense of democracy. In particular, David Estlund already has a ready response to the possibility of theocracy with his qualified acceptability requirement. On Estlund's view, it is not just the form of government that gets closest to Truth-Tracking that is preferable. Such epistemic merits are considered only once political institutions are "justifiable in terms acceptable to all qualified points of view (where 'qualified' will be filled in by 'reasonable' or some such thing)."[22] On Estlund's view, the epistemic success of a form of government is only considered if it is acceptable to all qualified or reasonable points of view. Estlund uses this criterion in order to rule out forms of government that might do better at Truth-Tracking than democracy. It may well be that certain forms of epistocracy will outperform democracy so far as the truth is concerned, but it is also possible for reasonable people to disagree about who the experts should be in such political arrangements, and so these proposals fail to even get off the ground. Estlund could press the same point against the theocracy proposal, as there will presumably be qualified objections to who the religious experts are in a theocracy. Such a move allows Estlund to rule out theocracies before epistemic considerations come into play.

Deliberative epistemic democrats thus have a couple of options. They can restore the link by limiting their claims to the actual world, or they can keep their global claims while adopting something like Estlund's qualified acceptability requirement. The difficulty with going this latter route, however, is that it raises a serious question about whether the priority in choosing a a form of government is Truth-Tracking or Justification. Do we want political institutions that track the truth, or instead those that people have the most reason to endorse? Deliberative epistemic democrats think we

---

22] See Estlund 2008, 41.

can have both, that Justification and Truth-Tracking march in lockstep, but we have seen that work in contemporary epistemology has not borne out this assumption.

Estlund's qualified acceptability requirement avoids this worry by privileging Justification over Truth-Tracking, but this is tantamount to switching strategies for defending democracy. What is most fundamental, on Estlund's view, is that everyone has enough reason to endorse their political institutions. Whether or not those institutions are truth-conducive is a secondary consideration. Arguing for a qualified acceptability requirement thus puts the emphasis on the reasons citizens have for their political beliefs, not on the truth of those beliefs, making the updated view quite distinct from strict Truth-Tracking. If, in response to difficulties connecting Justification to Truth-Tracking, epistemic democrats opt for something like qualified acceptability, then perhaps it is ultimately Justification that matters after all. Because it removes the fundamental emphasis on truth, a move to qualified acceptability represents a shift from a truth-centered focus to a justification-centered focus in defending democracy, admitting the two do not always coincide.

V. CONCLUSION

Amongst epistemic defenses of democracy, deliberative epistemic democrats have a unique strategy. Though all epistemic democrats argue that democracies are effective at tracking the truth, deliberative epistemic democrats emphasize the type of discursive environment created by democratic political cultures, pointing out the epistemic virtues of these settings. Hélène Landemore argues that these environments capture the widest range of cognitive diversity, while David Estlund contends that democratic political cultures best instantiate ideal epistemic deliberation. What these views have in common is that they argue that democratic citizens are better justified in their political beliefs and that this explains why democracy is better at tracking the truth than other political arrangements. Deliberative epistemic democrats have overlooked, however, that there is no consensus in epistemology that there even is a connection between justification and truth, so even if Estlund and Landemore demonstrate Justification, this is not yet to show that Truth-Tracking also obtains. If it is not possible to have both Justification and Truth-Tracking, how deliberative epistemic democrats respond to this difficulty reveals what really matters in political institutions. If we cannot have both, perhaps we should prefer political institutions that we have the most reason to accept, rather than those that track the truth.

rwsiscoe@gmail.com

REFERENCES

Bach, Kent. 1985. A Rationale for Reliabilism. *The Monist* 68: 246-63.

Berker, Selim. 2013a. Epistemic Teleology and the Separateness of Propositions. *Philosophical Review* 122: 337-93.

———. 2013b. The Rejection of Epistemic Consequentialism. *Philosophical Issues* 23: 363-87.

Bonjour, Laurence. 2009. *Epistemology: Classic Problems and Contemporary Responses*. Rowman and Littlefield Publishers.

Brennan, Jason. 2016. *Against Democracy*. Princeton University Press.

Brewer, Bill. 1997. Foundations of Perceptual Knowledge. *American Philosophical Quarterly* 34: 41-55.

Caplan, Bryan. 2007. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton University Press.

Carnap, Rudolf. 1962. *Logical Foundations of Probability*, 2nd ed. University of Chicago Press.

Cohen, Stewart. 1984. Justification and Truth. *Philosophical Studies* 46: 279-95.

———. 1998. Contextualist Solutions to Epistemological Problems: Skepticism, Gettier, and the Lottery. *Australasian Journal of Philosophy* 76: 289-306.

Cohen, Joshua. 1986. An Epistemic Conception of Democracy. *Ethics* 97 (1): 26-38.

Engel, Mylan. 1992. Personal and Doxastic Justification. *Philosophical Studies* 67: 133-51.

Estlund, David. 2008. *Democratic Authority: A Philosophical Framework*. Princeton University Press.

Estlund, David, and Hélène Landemore. 2018. The Epistemic Value of Democratic Deliberation. In *The Oxford Handbook of Deliberative Democracy*, edited by André Bächtiger, John Dryzek, Jane Mansbridge, and Mark Warren, 113-31. Oxford University Press.

Fantl, Jeremy and Matthew McGrath. 2002. Evidence, Pragmatics, and Justification. *Philosophical Review* 111: 67-94.

Fumerton, Richard. 2002. Theories of Justification. In *The Oxford handbook of Epistemology*, edited by Paul K. Moser, 204-33. Oxford University Press.

Goldman, Alvin. 1979. What is Justified Belief?. In *Justification and Knowledge*, edited by George Pappas, 1-25. Boston: D. Reidel.

———. 1986. *Epistemology and Cognition*. Harvard University Press.

Goodin, Robert and Kai Spiekermann. 2018. *An Epistemic Theory of Democracy*. Oxford University Press.

Hesse, Mary. 1974. *The Structure of Scientific Inference*. University of California Press.

Kearns, Stephen and Daniel Star. Reasons as Evidence. In *Oxford Studies in Metaethics*, Vol. 4, edited by Russ Shafer-Landau. Oxford University Press.

Kronz, Frederick. 1992. Carnap and Achinstein on Evidence. *Philosophical Studies* 67:151-67.

Landemore, Hélène. 2013. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many.* Princeton University Press.

Lasonen-Aarnio, Maria. Forthcoming. Virtuous Failure and Victims of Deceit. In *The New Evil Demon*, edited by Julien Dutant and Fabian Dorsch. Oxford: Oxford University Press.

List, Christin and Goodin, Robert E. 2001. Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *Journal of Political Philosophy* 9: 277-306.

Littlejohn, Clayton. 2009. The Externalist's Demon. *Canadian Journal of Philosophy* 39: 399-434.

———. 2012. *Justification and the Truth-Connection*. Cambridge University Press.

———. Forthcoming. A Plea for Epistemic Excuses. In *The New Evil Demon Problem*, edited by Fabian Dorsch and Julien Dutant. Oxford University Press.

Maher, Patrick. 1996. Subjective and Objective Confirmation. *Philosophy of Science* 63: 149-74.

McGrew, Timothy. 1995. *The Foundations of Knowledge.* Lanham, MD: Littlefield Adams Books.

———. 1998. A Defense of Classical Foundationalism. In *The Theory of Knowledge: Classical and Contemporary Readings*, edited by Louis Pojman. Wadsworth.

Roush, Sherrilyn. 2004. Discussion note: Positive Relevance Defended. *Philosophy of Science* 71: 110-16.

Solomon, Miriam. 2006. Groupthink Versus The Wisdom of Crowds: The Social Epistemology of Deliberation and Dissent. *The Southern Journal of Philosophy* 44: 28-42.

Somin, Ilya. 2016. *Democracy and Political Ignorance: Why Smaller Government is Smarter.* Stanford University Press.

Steup, Matthias. 2005. Epistemology. *Stanford Encyclopedia of Philosophy.*

Sutton, Jonathan. 2005. Stick To What You Know. *Nous* 39: 59-396.

Swinburne, Richard. 1973. *An Introduction to Confirmation Theory.* Meuthen.

Williamson, Timothy. 2000. *Knowledge and Its Limits.* Oxford University Press.

———. Forthcoming. Justifications, Excuses, and Skeptical Scenarios. In *The New Evil Demon*, edited by Julien Dutant and Fabian Dorsch. Oxford University Press.

Speaking Truth to Power? A Foucauldian Theory of Whistleblowing in A Nihilistic World

Lingyu Jing
University of York

Abstract. This paper aims to propose a novel account of whistleblowing. Whistleblowing is seen as "speaking truth to power". But if there is no truth outside power, as in our post-truth age, whistleblowing loses its meaning. The relationship between truth and power should be questioned. The paper refers to Michel Foucault's works to interrogate the relations between truth, power, and subjectivity in whistleblowing and show how mainstream theory fails to address those relations. The new account uses Foucault's term "practice of freedom" to re-anchor the meaning of whistleblowing in a nihilistic world without ultimate truth.

Keywords: whistleblowing, Foucault, the practice of freedom, subjectivity, truth-telling.

Whistleblowing is a politically and ethically contested issue. When the term is mentioned, many names and their controversial stories come to mind: Edward Snowden revealed details of a surveillance programme by the US government, through stealing classified information from the National Security Agency; Aleksandr Solzhenitsyn wrote a book about the Soviet Gulag through journalistic and historical investigations (Delmas 2015, 79). Although their whistleblowing practices are very different, there is a commonality that they all speak truth to power.

The relation between truth and power is key to understanding whistleblowing. In academic debates, whistleblowing is defined as "an act of public disclosure of information concerning wrongdoings" (Santoro and Kumar 2018, 9). In the mainstream theories, whistleblowing is seen as a moral issue that truth should be disclosed to defeat the abuse of power. In short, truth must be defended, and power must be restrained. Whistleblowing is justifiable because the truth is always more valued than power.

This paper, however, aims to challenge the valuation of truth and power in whistleblowing by putting forward a new understanding. In the tradition of western philosophy, truth is seen as the highest value in the world. We humans must pursue truth. The current understanding of whistleblowing is in this tradition. However, Nietzsche questions the value of truth and denies the existence of ultimate truth (Nietzsche 1989, 119). His contestation reveals a problem for whistleblowing, particularly in the "post-truth" age: why we should blow the whistle if there are competing truths, or even worse, there is no truth at all?

Michel Foucault, as a successor of Nietzsche, also questions the value of truth, but he instead insists on the importance of truth-telling. He praises the ancient practices of parrhesia, which means truth-telling (Foucault 2010, 66). He is nihilistic, but he finds the value of truth-telling in the world without ultimate truth. His thought is helpful for us to re-understand the meaning of whistleblowing in the post-truth era.

The paper will use his theory to expose the social and political conditions of truth-telling and develop a novel account of whistleblowing. First, the paper will map out the current debate on whistleblowing in political theory and identify conscience as the anchor of truth-telling. Second, by building on Foucault's conceptions of power, subjectivity, and truth, I will expose the political conditions of truth and conscience-based truth-telling. Third, by building on Foucault's later works of ethics, I will put forward a new understanding of whistleblowing not anchored on the ultimate value of truth but on the act of truth-telling itself.

I. CONTEMPORARY THEORIES OF WHISTLEBLOWING

Currently, there are three mainstream theories of whistleblowing: individual duty, organizational practice, and individual freedom. The "individual duty" view is applied ethics understanding whistleblowing akin to civil disobedience. The main argument is that whistleblowing is a moral duty that individuals have to act extraordinarily to disclose organizational wrongdoing. In particular, two justification for this individual duty can be found in the literature: (i) harm-based justification and (ii) complicity-based justification. In general, harm-based justification proceeds as follows: A member of an organization is morally required to blow the whistle if the disclosure can reasonably prevent the harm which will be caused by the wrongdoing within the organization (De George 2014). This justification of disclosure is a form of consequentialism, which is derived from a moral principle of non-harm generally applied in society (De George 2014, 321).

Complicity-based justifications, on the other hand, primarily concentrate on the moral wrong involved in an agent's complicity in wrongdoing. According to proponents of this view, a member of an organisation ought to disclose wrongdoing committed by their organisation, if the unauthorised disclosure can avoid their complicity with the wrongdoing (Davis 1996). Complicity in wrongdoing is independent of the consequence of wrongdoing, meaning that complicity itself constitutes a moral wrong. Since, in general, we all have a duty to avoid moral wrongs, whistleblowing is justified as a moral duty of non-complicity (Davis 1996, 10).

The two justifications for individual duty understand whistleblowing as a moral action. It is grounded on the moral capacity of individuals that they ought to do the right thing and speak up for the truth.

In contrast, the "organizational practice" view understands whistleblowing not as an individual act but as an organizational function. The crucial difference is derived from the organisational context of whistleblowing. Since individuals are embedded in the structure of the organisation, which is a "system of embodied interrelated rule-governed roles", they occupy structurally interrelated roles (Ceva and Bocchiola 2020, 7). Organisational roles are structurally interrelated due to the function they play within the same organisation – that is, an organisation can function only if all

role-occupants perform their role appropriately. As a result of such interdependence, wrongdoing committed by one role-occupant constitutes a threat to the functioning of the entire organisation. Therefore, in this view, a wrongdoing is not a moral problem but an organisational failure. Such a failure undermines "the normative property of a well-functioning organisation" (Ceva and Bocchiola 2020, 7). It follows then that whistleblowing, on this view, is not an individual act of civil disobedience, but an organisational practice to initiate organisational self-correction.

Ceva and Bocchiola's account introduces a new dimension of whistleblowing by highlighting that whistleblowers are embedded in organisational structures. In fact, for individuals within the organization, whistleblowing is often simply doing their job according to their organisational duties. For example, internal investigators have duties to inspect organizational corruption. It situates the practice of whistleblowing within the routinised organisational answerability structure. The wrongful conduct of some individuals will not seriously undermine the functioning of the organisation if there are some responsible employees who comply with the organisational duty of whistleblowing to denounce and correct the wrongdoing. In other words, whistleblowing should be institutionalised within a legitimate disclosure regime.

Although the "organisational practice" view is a welcome addition to theories of whistleblowing, it still has limitations. If the "individual duty" view is too unrealistic because it relies on extraordinary individuals, then the organisational practice view is too idealistic because it relies on already transparent and well-functioning organisations. An accountable disclosure regime can exist only in transparent organisations, which are committed to public answerability and accountability and where wrongdoing is limited to one-off events. However, there are some non-transparent organisations with secretive intent where wrongdoing is more systematically committed, for instance, within intelligence organisations (e.g. KGB and CIA). We cannot expect this type of organisation to uphold full accountability and transparency because the nature of such organisation requires secrecy and the potential to be ready to commit wrongs.

In most cases, organizations are imperfect. The organizational practices of self-correction are dysfunctional. Then it requires individuals to perform the moral duty of whistleblowing. Again, it entails an individual to have the moral capacity to do so, and this capacity is grounded on conscience. This issue is addressed by the "individual freedom" view, which is mainly proposed by legal scholars Richard Haigh and Peter Bowal (2012). They argue that conscience defines one's individuality and individuals always have the freedom of conscience to perform whistleblowing (Haigh and Bowal 2012). Theoretically, conscience has two inseparable functions: the faculty of self-assessment and moral knowledge. Conscience refers to the inner judge, which constantly observes and examines the self. An individual is split into two persons, one who acts and one who judges the former's conduct as Kant (1991, 189) says that conscience is the "consciousness of an internal court in man." This faculty of self-assessment implies that we are subjectively compelled to answer to the judgment of our conscience. This sense

of an inner judge thus enables us to conform our conduct subjectively to the moral codes that are borne by conscience.

Moreover, conscience refers to the inner voice, which epistemically identifies what is wrong. Once our conduct does not conform to the moral codes, conscience speaks the "infallible voice of truth" that reminds us of our error (Haigh and Bowal 2012, 25), as Rousseau (1979, 286) says, "conscience is the true guide" that "never deceives us." Hence, the two functions of conscience are thought to able us to identify and correct wrongdoing. As Haigh and Bowal (2012, 24, 29) argue, conscience is "the voice which tells us what we should do and what we should leave undone, what the pattern and purpose of our lives should be".

In this way, whistleblowing is justified as the fundamental freedom of conscience, which is the human right to make decisions to act according to one's conscience, as the task of conscience is to translate conviction into action (Strohm 2011, 93). Once one's conscience subjectively judges conduct interfered with by organisational wrongdoing as an error that may jeopardise their moral integrity, whistleblowing can be the decision to "freeing" conscience from the error. In short, whistleblowing is the liberation of conscience.

Overall, the act of whistleblowing is grounded on the freedom of conscience. Since our conscience is internally free, we are always capable of self-determination to perform whistleblowing.

II. TRUTH AND POWER IN WHISTLEBLOWING

The idea of freedom of conscience reflects a traditional understanding of whistleblowing: to blow the whistle is to follow the voice of conscience. It promises an inherent free choice of whistleblowing rooted in conscience. It relies on two fundamental assumptions. First, conscience must be understood as an Archimedean point to make conscientious and free choices. We are capable of self-assessment because conscience, as the internal judge, conforms our conduct to the moral codes. As the nature of conscience is stand-alone (Haigh and Bowal 2012, 3), our conscience-based choice to blow the whistle is independent of the outer world (at least for most adults). Second, wrongdoing is seen as truth. We assume that wrongdoing is a fact that is clear and accessible because our conscience makes us aware of what is wrong. This is due to the epistemological function of conscience: it speaks the voice of truth that identifies what is wrong (Haigh and Bowal, 2012, p. 25). As long as our conscience is aware of wrongdoing and compels us to comply with moral codes, it is our free choice to blow the whistle to liberate our conscience from error. In other words, conscience enables us to "escape from power into freedom" (Taylor 1984, 153).

However, in what follows, I will use Foucault's conceptions of subjectivity, power, and truth to show that the two assumptions are problematic. Freedom of conscience, in fact, is a description from the subject's perspective, presupposing a

pre-given conscience possessed by the individual; however, this assumption is a mere philosophical abstraction. It presupposes a Cartesian subject: an individual subject has a priori self-knowledge that transparently makes the person aware of "who I am" and "what I should do", as the roots of conscience are in the Latin word *conscientia*: self-knowledge. Foucault (2001, 335) rejects the Cartesian universal and transcendental subject and proposes specific and historical subjects. According to the Foucauldian subjectivity, conscience is not independent of power; on the contrary, power shapes conscience. To understand this, we should not assume power as "abuse of power" that represses individuals' conduct (Haigh and Bowal 2012, 7); in contrast, power produces individual conscience and incites them to act (Foucault 2001c, 340).

This new model of power is disciplinary power, which exists in power relations in which subjects are incited to strategically interact with each other in a specific pattern (Foucault 1984, 93). Being a subject implies an individual "tied to one's own identity by a conscience or self-knowledge", and the conscience subjectively compels the individual to conduct in a pattern that is incited by power (Foucault 2001c, 331). The subject, for Foucault, is produced through disciplinary practice in which conscience is turned from an inner judge into an inner guard, which exercises self-gazing and self-examination to conform our conduct to certain codes. In other words, our moral codes are codified by power relations. Like the panopticon effect in *Discipline and Punish*, the eye of power is visible but unverifiable; therefore, a prisoner in the innate state of conscious simultaneously plays the roles of both prisoner and guard, hence internalising codes of conduct through constant self-surveillance (Foucault 1991). In history, many specific individual subjects have been produced through various disciplinary practices, such as criminal, mad and sexual subjects (Foucault 1991, 1988b, 1984). This is to say, there is no eternal moral code but only moralizations due to the eternal change of power relations throughout history.

Since disciplinary power has been diffused into the whole society, the subjects' free choice of whistleblowing is not totally self-determined. It means that the conscience-based decision of whistleblowing is a specific pattern of conduct incited by power relations rather than a self-determined choice based on freedom of conscience. It implies that the subject-centered description of freedom might hide power relations, which the mainstream theories fail to address.

In *Madness and Civilisation*, Foucault (1988b, 213) famously claims, "freedom of conscience entails more dangers than authority and despotism". In appearance, subjects embedded in power relations are still free, not because power has not fully produced the subject yet, but because freedom is the condition of governing well (Foucault 2007, 353). Freedom here services as the "regulator" of governmental practice (Foucault 2008, 65). From a non-subject centred perspective, freedom is redescribed not as an abstract status of self-determination but as a political strategy to govern our conduct. We are not prohibited from free conduct; instead, we are governed through freedom, which involves "the whole range of practices that constitute, define, organize, and

instrumentalize the strategies that individuals in their freedom can use in dealing with each other" (Foucault 2019, 300). In other words, freedom is concretely practiced in power relations that condition our possibility of conduct and hence limit our conduct within the specific pattern that is defined by the dominant discourse. In short, freedom is a conditioned field of possibility of conduct.

In this redescription of freedom, the non-subjective power relations rather than the subject are centred. We seem to follow the voice of conscience to make choices to comply with "the right code", but this code is exactly what power aims to discipline. The moral codes that define personal identities are merely productions of power since power is "soulcraft", which disciplines persons into specific subjects and therefore subjected to power relations. As Foucault (1991, 30) says, "the soul is the prison of the body". In this sense, conscience is not our internal fortress where we are free but the internal prison to which we are subjected. Disciplinary power allows subjects to freely conduct in a specific pattern not because freedom is a priori or valued but because power relations are reproduced through the subjects' interactions and hence reinforce the control of subjects. That is to say, "freedom of conscience" is a function of power relations, enabling us to moralize the external world and confirm the dominant moral codes in that historical period.

For a long time, we have described whistleblowing as an activity of truth-telling (Santoro and Kumar 2018, 15). It is reflected in the second assumption that wrongdoing is a fact that is clear to us because conscience speaks the truth that makes us aware of wrongdoing. However, conscience cannot guarantee that we can identify wrongdoing and speak the truth because the truth is spoken through a dominant discourse that is produced by power. Discourse is understood as the way to organise and express knowledge in our language (Foucault 1972, 193), but discourse legitimates a certain type of knowledge while simultaneously ruling out other types of knowledges. Wrongdoing has a discursive dimension: "wrongdoing" is legitimated through a dominant discourse that rules out the knowledge about what is wrong. This is because, for Foucault, knowledge is not natural but produced by power and reproduces power (Foucault 1991, 27). In this sense, wrongdoing is not a mere fact that is closed to interpretation but a discourse: whether it is right or wrong is not determined by our conscience but the power/knowledge relations that produced it.

In this way, truth should not be understood as a metaphysical existence but a political construction. For Foucault, there is no distinction between truth and falsehood, but only discourses. Historically, some discourses are counted as truth, while some are excluded. Foucault calls it the "regime of truth", which is "a system of ordered procedures, institutions, and authorities for producing "the types of discourse which it accepts and makes function as true" (Foucault 1980, 131). It is about how discourse can become truth in a particular society. Truth is not prior to truth-tellers. On the contrary, power relations create the standards of truth because they determine who the truth-tellers are. Power relations do not simply determine the truth-tellers externally by force

but internally shape their subjectivity to let them believe in some self-evident truth. A Christian would not believe in the evolution theory of humans because God's creation of men is self-evident for him. He refuses the scientific truth of evolution because of his political identity as a Christian.

Foucault denies the existence of the unchangeable truth. For him, there are only competing discourses, and all of them can become truth in a particular society and a historical period. Is witchcraft true? It was in the age of witch-hunting, as least from the people's views. Then a man who informed the church that a woman performs witchcraft can be seen as a whistleblower, at least in that historical context. This is because if there is no absolute boundary between truth and falsehood, there is no absolute distinction between a whistleblower and an informer. Foucault's understanding of truth is relativistic and nihilistic, but unfortunately, this is the world we live. In the post-truth age, the metaphysical structure of truth is gone. There are only competing narratives. Is climate change true? Was Covid-19 bioengineered in a laboratory in China? There could be no answers at all. Even if there is a definitive answer for them, it can be changed with time. As we mock the witch-hunters in history, our belief in the "truth" will be mocked by future generations. As long as the standards of truth change with time, truth only exists historically.

The loss of ultimate truth reveals a fundamental question for whistleblowing: if there is no absolute existence of truth, why should we speak truth to power? In the mainstream theories, whistleblowing is justifiable because truth exists and waits for the revealing. The conscience is waiting for liberation from power. Truth is seen as the antithesis of power. But Foucault's theory implies the entanglement of truth and power. If the truth is shaped by power relations like in the case of witch-hunting, whistleblowing is nothing more than a confirmation of power. Freedom of conscience is no longer the anchor of truth and moral act. Rather, it has become a prison, constraining us in a regime of truth/power where we can only speak up the dominant discourse. We have to find new meanings for whistleblowing in the entangled relations of truth and power.

III. WHISTLEBLOWING AS THE PRACTICE OF FREEDOM

Since there is no ultimate truth, whistleblowing is not about the disclosure of truth. Rather, whistleblowing is about power struggle. Foucault, just like Nietzsche, does not value truth. Because humans are not interested in truth itself but in the will to truth, which is the will to power. We pursue truth because the will of pursuing makes us feel powerful. As Nietzsche says, we would will to nothingness rather than not will (1989, 97). Whistleblowing as speaking truth to power is struggling for power. But this power struggle should not be understood as a realist politics irrelevant to truth and morality. It does not say that truth is not important. Since there is no truth outside power and no power outside truth, a theory of whistleblowing cannot oppose truth and power or deny one of them. Rather, truth-telling itself creates power. Whistleblowing can

be seen as a rupture of the regime of truth, which disentangles and reconfigures the relations between subjectivity, truth, and power.

In Foucault's ontology, the world has no order or direction but only flows of power relations, which are the multiplicity of antagonistic forces traversing all local positions in the social body (Foucault 1984, 93). Those forces constantly confront each other and create certain types of subjectivity, such as masters and slaves, the normal and the abnormal, Christians and pagans. The dominant force becomes the truth-teller and creates the standards of truth. And the truth reinforces the subjectivity by inscribing self-knowledge and then generates power relations. Truth and power are a reciprocal circle. And subjectivity is the key medium of the relations between truth and power. Subjectivity is self-knowledge about who you are. The disentanglement of the power relations must start with subjectivity.

As I have examined in the last section, "freedom of conscience" is not totally self-determined; instead, it is a function of power relations. This "unfreedom" implies that our subjectivity is not the true author of our conduct. As demonstrated, the problem is rooted in subjectivity, which always conforms truth-telling to the dominant discourse. This self-policed truth-telling has been internalised as the subject's fundamental codes of conduct. Foucault (2019a, 378), through the eyes of Nietzsche, conceives those prescriptive codes of conduct as a symptom of modern power that constrains our potentiality to act. However, in his seminal works on ethics, Foucault sees a possibility of freedom of conduct in ancient Greek ethical practices, which enables one to invent new codes according to our own subjectivity. It is a form of ethics that transforms "subjected subjects" into the ethical "subjects of action, subjects of true knowledge" (Foucault 2005, 417). To see whether the Foucauldian ethics provides a possibility for the freedom to blow the whistle, I start with his distinction between morality and ethics.

Morality is prescriptive, while ethics is creative. For Foucault (2012, 25), morality is a "prescriptive ensemble" which entails a set of values and rules of conduct, such as the codes of conduct that are internalised by the "regime of power". This "prescriptive ensemble" produces our subjectivity and codifies our conduct, making our actions subject to the dominant discourse rather than subject to ourselves. On the contrary, ethics is the relationship with the self, in which the person "constitutes himself as a moral subject of his own actions" (Foucault 1997, 262). In short, ethics is *ethos*, a mode of being, not a prescriptive code. To be moral is merely to comply with the codes that are inscribed by power relations but ethics entails inventing a relation to the self to impute our conduct to our own subjectivity. In other words, ethics is to reconstitute a subjectivity to take back control of our conduct that has been codified by power relations. This is because the subject is an imposed form so that it can be refused and transformed (Foucault 1997b, 290).

Before we give a new form to our subjectivity, we need to refuse the existing subjectivity that has been shaped by the dominant discourse. To do so, Foucault

(2012b, 43) offers a technique: the "practice of the self", which derived from an ancient Greek culture, "the care of the self", wherein the relation with the self was cultivated to be an *ethos*. In practice, a person practices reflective self-inspection to guarantee that "one will not become attached to that which does not come under our control" and "accept in the relation to the self only that which can depend on the subject's free and rational choice" (Foucault 2012b, 64). The "practice of the self" aims to achieve self-control in which the subject's conduct detaches from the codification of power and thus freely exercises sovereignty over the self, in Foucault's term, the practice of freedom.

Here, freedom must not be understood as legal or moral codes that permit or are prescribed by someone to do something but as a form of practice invented by someone to constitute a new subjectivity. This non-prescriptive concept of freedom perhaps is not compatible with the language of justification in analytic political philosophy. When we try to justify a decision as free choice, we usually refer to universal principles such as freedom of conscience. It philosophically supposes that we are entitled status of being free, and whistleblowing is justified if we aim to achieve such a status of freedom. However, such abstract freedom does not exist in practice, as in the Foucauldian framework, freedom is a conditioned field of possibility in which the subject conducts in a specific pattern. It indicates that freedom exists only in concrete practices. Hence, the aim of whistleblowing is not to achieve the entitled universal status of freedom but to invent desired concrete practices of freedom. It is the teleological pattern of conduct that the whistleblower aims to invent. In this pattern, the whistleblower exercises sovereignty over the self and speaks "true discourse" by their own *ethos*. As Foucault (2019, 284) contends, "ethics is considered form that freedom takes when it is informed by reflection". Freedom is ethical, not in the sense of moral justification, but in the way that this form of practice is informed by the wish of ethical self-constitution.

The practice of freedom entails self-reflection, which critically practices dissenting discourses in order to refuse the imposed subjectivity. As seen, the unreflective adoption of the dominant discourse results in a subjectivity that is imposed by power relations. However, although our subjectivity is defined, it is still possible to refuse it since we have the capacity for self-reflection, which leads to critical engagement with the dominant discourse. In *What is Enlightenment*, Foucault (2019b, 319) proposes self-reflection as a practice of "critical ontology of ourselves" in which "the critique of what we are is at one and the same time the historical analysis of the limits that are imposed on us". The name of the essay is a dialogue with Kant. For Kant, the world is ordered, and men are capable of discovering their limits by a priori reason. In contrast, for Foucault, the world is chaos; the limits of men are imposed historically and therefore offer the possibilities to transgress the limits. This is to say, we need to reflectively interrogate our subjectivity to expose that the subjectivity is not self-evident but rather historically constituted. In other words, critical engagement

is to expose that the limits of our subjectivity are imposed by the dominant discourse, which is produced by the regime of truth.

This self-reflection practice is not to search for a definitive answer about the self; instead, it is the experimental and transgressive practice of freedom upon limits. As Foucault (2019b, 319) continues, critique is "an experiment with the possibility of going beyond [limits]". In practice, we can start the self-reflective practice by non-conforming to the dominant discourse. That is to say, the practice of freedom is to practice dissenting discourses and hence refuse the self that is imposed by the dominant discourse. In short, practicing dissenting discourses is the art of "reflective indocility" (Foucault 1997b, 386). This self-refusal practice opens up the possibility of no longer being the same self and thus motivates one to seek another self that is not fully constituted by the dominant discourse (Foucault 1997d, 315). For example, Foucault (2001a, 465) comments that in the Solidarity movement in Poland, "the people have not only struggled for freedom… but they have done so by exercising …freedom". Their "[dissenting] discourses … converted into the creation of something new" and eventually led to the transformation in 1989 (Foucault 2001a, 468). This suggests that the dissenting discourse produced by the practice of freedom enables one to refuse the self and thus "opens up the space of freedom…of possible transformation" (Foucault 2019b, 450).

The practice of dissenting discourse thus sheds light on the practice of whistleblowing, as the disclosure of wrongdoing can be seen as a "virtual fracture" of the "regime of truth", releasing the self from the imposed code of conduct (Foucault 2019a, 450). This self-refusal practice opens up a space of freedom where power relations are less intensified, and the subject is "freer" to create truth. Since there is no ultimate truth, everything is possible.

The formation of a new subjectivity has to bring new self-knowledge based on new standards of truth and therefore constitutes a critical opening of the regime of truth. Whistleblowing, in this sense, is challenging the power relations that produce the truth. It is not "emancipating truth from every system of power" but "ascertaining the possibility of constituting a new politics of truth" (Foucault 1980, 133). In other words, the meaning of whistleblowing is the power struggle that transfigures the relations between power and truth. The act of truth-telling creates a new standard of truth and hence generates power.

Whistleblowing can be redescribed as the "practice of freedom" in the sense of normative engagements with the dominant discourse in order to reconstitute a truth-telling subjectivity. The practice of freedom, in the case of whistleblowing, is to practice truth-telling that is non-conforming to the dominant discourse, but this practice needs a normative orientation to maintain continuous control over the conduct. Specifically, the continuous practice of whistleblowing is a process of self-controlled normalisation in which the whistleblower reflectively adopts the "true discourse" as their new subjectivity. Through normative practices of "true discourse",

the whistleblower recodifies their conduct of truth-telling. This kind of normative practice is termed "free subjectification" (Schubert 2020, 15), while I prefer to call it the "self-controlled practice of freedom". Therefore, in their own "game of power", the person constitutes themselves as a truth-telling subject who is capable of conforming truth-telling to the "true discourse" rather than the dominant discourse.

In this understanding, the imperative of whistleblowing is not grounded on truth but on the act of truth-telling. Since there is no ultimate truth, the will to truth/power is the new anchor of whistleblowing for great individuals who are interested in their value of existence. The nihilistic understanding of truth certainly creates great dangers as the whistleblowers could speak what they want or even intentionally lie. Nihilism opens the gate for great truth-creators but also for great liars. We would lose the anchor of moral codes and push humans into a storm of power struggles. But for Foucault, as a Nietzschean, this is how the world is, a chaotic field of force confrontations. Foucault is a pessimist and an activist who acknowledges the dangers of actions and is still willing to act (1997, 256). He enjoys chaos. In this account, whistleblowing is not to protect the order of the world but to reveal the chaotic nature of the world and release the energy of life, which is the will to form a new style of individual existence.

IV. CONCLUSION

I proposed a Foucauldian understanding of whistleblowing as the "practice of freedom". In contrast with the mainstream normative theories, it does not ground the value of whistleblowing on truth but on the act of truth-telling. The capacity that enables truth-telling is no longer freedom of conscience but freedom of practice which transforms subjectivity. This account is nihilistic in moral and epistemological senses, but it reflects the current world we live in and gives meaning to whistleblowing.

jing.lingyu@foxmail.com

REFERENCES

Boot, E. R. 2019. *The ethics of whistleblowing*. Routledge.
Ceva, E. and M. Bocchiola. 2019. *Is whistleblowing a duty?*. John Wiley & Sons.
———. 2020. Theories of whistleblowing. *Philosophy Compass* 15(1): e12642.
Childress, J. F. 1979. Appeals to conscience. *Ethics* 89(4): 315-35.
Davis, M. 1996. Some paradoxes of whistleblowing. *Business & Professional Ethics Journal* 15: 3-19.

Delmas, C. 2015. The ethics of government whistleblowing. *Social theory and practice* 41(1): 77-105. doi:10.5840/soctheorpract20154114.

Foucault, M. 1972. *The archaeology of knowledge.* London: Tavistock.

———. 1980. *Power/knowledge : selected interviews and other writings, 1972-1977.* Edited by Gordon 1948-, C. (ed). New York: Harvester Press.

———. 1984. *The history of sexuality Vol.1.* Harmondsworth: Penguin.

———. 1988a. An Interview with Michel Foucault October 25, 1982. In *Technologies of the self: A seminar with Michel Foucault*, edited by L. H. Martin, H. Gutman, and P. H. Hutton, 9–15. Amherst: Univ of Massachusetts Press.

———. 1988b. *Madness and civilization: A history of insanity in the age of reason.* Vintage.

———. 1991. *Discipline and punish : the birth of the prison.* Edited by A. Sheridan. London: Penguin Books.

———. 1997a. *Ethics, subjectivity and truth: The essential works of Foucault 1954-1984 (vol 1).* Edited by P. Rabinow. London: Allen Lane The Penguin Press.

———. 1997b. What is Critique? In *The politics of truth*, edited by Lotringer, S. and Hochroth, L. New York: Semiotext.

———. 2001a. *Fearless speech.* Semiotext.

———. 2001b. *Power: The essential works of Foucault 1954-1984 (vol 3).* Edited by J. D. Faubion. London: Allen Lane The Penguin Press.

———. 2001c. The Moral and Social Experience of The Poles Can No Longer Be Obliterated. In *Power: the essential works of Michel Foucault 1954-1984*, edited by J. D. Faubion. London: Penguin UK.

———. 2005. *The hermeneutics of the subject: Lectures at the Collège de France 1981-1982.* Macmillan.

———. 2007. *Security, territory, population: lectures at the Collège de France, 1977-78.* Springer.

———. 2008. *The birth of biopolitics : lectures at the Collège de France, 1978-79.* Edited by M. Senellart et al. Basingstoke: Palgrave Macmillan.

———. 2010. *The government of self and others: Lectures at the Collège de France 1982-1983.* Springer.

———. 2011. *The courage of truth: Lectures at the College de France 1983-1984.* New York: Palgrave Macmillan.

———. 2012a. *The history of sexuality, vol. 2: The use of pleasure.* Vintage.

———. 2012b. *The history of sexuality, Vol. 3: The care of the self.* Vintage.

———. 2014. *Wrong-doing, truth-telling: The function of avowal in justice.* University of Chicago Press.

———. 2019a. Nietzsche, Genealogy, History. In *Aesthetics, method, and epistemology: Essential works of Foucault 1954-1984,* edited by J. D. Faubion, 369-391 London: Penguin.

———. 2019b. Structuralism and Post-structuralism. In *Aesthetics, method, and epistemology: Essential works of Foucault 1954-1984*, edited by J. D. Faubion, 433-458. London: Penguin UK.

———. 2019c. The Ethics of the Concern for Self as a Practice of Freedom. In *Ethics: Subjectivity and Truth: Essential Works of Michel Foucault 1954-1984*, edited by P. Rabinow 281-301. London: Penguin UK.

———. 2019d. What is Enlightenment? In *Ethics: Subjectivity and Truth: Essential Works of Michel Foucault 1954-1984,* edited by P. Rabinow, 303-319.

Foxley, I. 2019. Overcoming stigma: Whistleblowers as 'supranormal' members of society?. *Ephemera: theory & politics in organization*,19(4).

De George, R. T. 2014. Whistle-Blowing. In *Business ethics: Readings and cases in corporate morality.* edited by W. M Hoffman, R. E. Frederick, and M. S. Schwartz, 320-38. John Wiley & Sons.

Haigh, R. and P. Bowal. 2012. Whistleblowing and freedom of conscience: towards a new legal analysis. *Dalhousie LJ* 35: 89.

Kant, I. 1991. *The Metaphysics of Morals*. Edited by M. J. Gregor. Cambridge: Cambridge University Press.

Rousseau, J.-J. 1979. *Emile: or, On education*. New York: Basic Books.

Santoro, D. and M. Kumar. 2018. *Speaking truth to power-a theory of whistleblowing*. Springer.

Schatzki, T. R. 2001. Practice theory: an introduction. *The practice turn in contemporary theory*, 1-14.

Schubert, K. 2020. Freedom as critique: Foucault beyond anarchism. *Philosophy & Social Criticism*: 0191453720917733.

Simons, J. 1995. *Foucault and the Political*. London: Routledge.

Sluga, H. 2014. *Politics and the Search for the Common Good*. Cambridge: Cambridge University Press.

Snowden, E. 2019. *Permanent record*. Macmillan.

Strohm, P. 2011. *Conscience: A very short introduction*. Oxford University Press.

Taylor, C. 1984. Foucault on freedom and truth. *Political theory* 12(2): 152-83.

Teo, H. and D. Caspersz. 2011. Dissenting discourse: exploring alternatives to the whistleblowing/silence dichotomy. *Journal of Business Ethics* 104(2): 237-249.

Tully, J. 2008. *Public philosophy in a new key (Vol. 1): Democracy and Civic Freedom*. Cambridge: Cambridge University Press.

Valdez, I. 2016. Nondomination or practices of freedom? French Muslim Women, Foucault, and the full veil ban. *American Political Science Review* 110(1): 18-30. doi:10.1017/S0003055415000647.

Vandekerckhove, W. 2016. Freedom of expression as the "broken promise" of whistleblower protection. *La Revue des droits de l'homme. Revue du Centre de recherches et d'études sur les droits fondamentaux* (10).

Weiskopf, R. and H. Willmott. 2013. Ethics as critical practice: The "Pentagon Papers", deciding responsibly, truth-telling, and the unsettling of organizational morality. *Organization Studies* 34(4): 469-93.

The Significance of Deep Disagreements on Justice, Values, and Morals for Political Epistemology

Manuel Knoll[1]

Ludwig-Maximilians-Universität München

Abstract. Political epistemology asks whether there is knowledge about values, the good, the just, and the morally right that could be applied to political issues. Debates on such questions go back to the ancient Greeks. While Protagoras claims that no moral facts and no moral knowledge exist, both Socrates and Plato strive to overcome Protagoras's skeptical claims. After reconstructing their debate, this article argues, in the tradition of Protagoras, that there is no ultimate moral knowledge to ethically orientate political decisions. The main argument against the existence of a mind-independent and objective moral reality about which moral knowledge could be achieved is based on the existence of myriad widespread deep disagreements on values, justice, morality, and ethics, which are resistant to rational solution. The arguments presented against ethical realism and cognitivism in section II are mainly based on Max Weber's and Isaiah Berlin's views that deep disagreements on values, morality, and ethics exist. The arguments set out against these positions in section III rest on deep philosophical disagreements on social and political justice, which are a characteristic feature of the history of political thought.

Keywords: moral facts, moral knowledge, ethical realism, cognitivism, relativism, skepticism, meta-ethics.

I. DEEP DISAGREEMENTS ON THE CENTRAL PROBLEMS OF POLITICAL EPISTEMOLOGY: THE HISTORICAL BEGINNINGS

Political epistemology is the theory of knowledge applied to political issues. Its questions focus on the area where the disciplines of epistemology and political philosophy intersect.[2] The central question political epistemology asks is whether there is any knowledge that can ethically orientate political decisions. Political decisions are usually based on some kind of knowledge. When a government deals with another state, it attempts to collect, e.g., knowledge about its interests, its power to enforce them, and the people who govern. When legislators draft a new law, they seek to acquire knowledge

2] Political epistemology is not limited to research on the intersection of epistemology and political philosophy. The discipline comprises also research on the intersection of epistemology and political theory or political science and is open to collaboration with scholars from related fields; cf. Edenberg and Hannon 2021; Hannon and de Ridder 2021.

about all areas affected and the social and political consequences. However, such kind of knowledge does not allow governments or legislators to answer the question whether a policy is morally right or whether a law is just. The most fundamental question political epistemology asks is whether knowledge exists about values, the good, the just, and the morally right.

Debates on such questions go back to the ancient Greeks and in particular to Protagoras, Socrates, and his student Plato. The famous sophist Protagoras of Abdera, who should be interpreted as the founder of ancient Greek political philosophy[3], was twenty years older than Socrates. Informed about the variation in customs and moral codes from one culture to another by the historian Herodotus, Protagoras defended the view that in moral and legal matters there are no universal truths. According to Protagoras's argument, if moral truths and an objective or universally valid morality existed, we could expect considerable agreement on moral beliefs. However, we can observe a lot of diversity in moral views and substantial disagreement on what is right and wrong. Therefore, no moral truths and no objective or universally valid morality exist. In the contemporary debate, this argument against the existence of moral truths and an objective or universally valid morality is called the *argument from moral disagreement* or the *argument from relativity* (cf. Gowans 2000, 3-4, 15-18; Ladd 1985 1-3; Mackie 1977, 37; Tersman 2006, xii-xiii).

Despite Protagoras's rejection of any universal truths in moral and legal matters, he argues that some moral beliefs or views about the good and just are more beneficial or useful than others (Plato, *Theaetetus*[4] 166d-167d, 172a-b, 177d-e). Public debates and deliberations about what is good, just, and beneficial for the political community are a characteristic feature of democracy. In the myth the sophist presents in Plato's dialogue *Protagoras*, he has Zeus distribute "justice" (*dikê*) and "respect" (*aidos*) to *all* citizens. Because all citizens have a sense of justice and display respect for others, they all are political beings. Because they all possess such political virtues, the Athenians allow them all to participate in the political life of the *polis*. For Protagoras, politics is not a matter of knowledgeable experts but of *all* citizens (Plato, *Protagoras* 322c-323a). This is an important reason why Protagoras has been interpreted as a defender of (Athenian) democracy (Giorgini 2019, 107-114; Giorgini 2016, 10, 25; Ottmann 2001a, 221-22; Ottmann 2001b, 11-12).[5]

3] Several scholars regard Socrates as the founder of political philosophy (Ottmann 2001, 243-44; Strauss and Cropsey 1987, 4-5). However, the "human turn" of philosophical thought to epistemological, ethical, and political issues occured as early as Protagoras.

4] Henceforth referred to as "*Tht.*". This dialogue, and Plato's *Protagoras*, are our main sources for Protagoras's philosophy.

5] There are more reasons to interpret Protagoras as a democratic political thinker. He was a friend of Pericles, the leading statesmen of Athenian democracy. Pericles asked Protagoras to draft the laws of Thurii, a *polis* newly founded between 446 and 443 BC in the South of Italy (*Diogenes Laertius* IX 50). Because Thurii was a "colony" (*apoikia*) of Athens, in all likelihood it had a democratic constitution.

In the principal clause of his philosophy, Protagoras claims that "man is the measure (*metron*) of all things (*chrêmata*): of the things which are, that they are, and of the things which are not, that they are not" (*Tht.* 152a, trans. Levett, rev. Burnyeat). Because of this clause, Protagoras is usually "considered the first official voice of relativism" (Baghramian and Coliva 2020, 27).[6] According to Plato's interpretation of the clause, all truths are relative to perceiving individuals (*Tht.* 152a). There is not the *one* truth, but as many truths as there are perceiving individuals. Individual human beings are the measure of what is true. Applied to the field of morality, law, and politics, Protagoras's epistemological relativism means that no such thing as "the just and unjust" has "by nature (*physei*) any being (*ousia*) of its own" (*Tht.* 172b, trans. Levett, rev. Burnyeat). In this statement from Plato's *Theaetetus*, Protagoras denies the existence of mind-independent moral facts and an objective moral reality. In the language of contemporary meta-ethics, he was an "anti-realist"[7]. Protagoras holds that "whatever any community (*polis*) decides to be just and unjust, and establishes as such, actually is what is just and right for that community and for as long as it remains so established" (*Tht.* 177d, trans. Levett, rev. Burnyeat). According to Protagoras's ethical and legal relativism, everything which is just and legal is valid only for one *polis* and relative to its particular morality and laws. Just as there is no objective and universally valid truth, there is no objective or universally valid law and conception of justice. Rather, what is just and legal depends on the decisions of a particular political community. This is why Protagoras has been interpreted as a "political conventionalist" and "legal positivist" (Giorgini 2019, 112).

For Protagoras, legislation always aims at what the whole community holds to be its good and its advantage, "A community always makes such laws as are most useful to it" (*Tht.* 177d, trans. Levett, rev. Burnyeat). As already mentioned, this implies public debates and deliberations about questions of justice and about what is good and beneficial for the political community.[8] Protagoras view that some moral beliefs

---

6] Baghramian and Coliva explain, "It is difficult to know what variety of relativism, if any, Protagoras was defending, but Plato seems to be attributing alethic relativism – to the effect that claims to truth should be relativized to a framework or perspective – to Protagoras" (2020, 27; cf. 1-2). Giorgini attributes "epistemological relativism" to Protagoras (2019, 112). Ziglioli conceptualizes Protagoras's relativism as "perceptual relativism" and "ethical relativism" (2007, 15-16). On p. 15, Ziglioli 2007 explains, "perceptual relativism is shown to be an epistemological doctrine that entails a kind of ontological indeterminacy, according to which nothing is if not in relation to somebody". For arguments for the thesis that "Protagoras' position cannot be defined as relativistic", see Eustacchi 2016, 34.

7] In contemporary meta-ethics, "moral realism" or "ethical realism" is the position that claims that moral facts, an objective moral reality or objective moral values exist in mind-independent ways (Miller 2013, 5; Rüther 2015, 93-123. Tersman equates the "realist" view on ethics with the "objectivist" view. Both views claim "that moral issues are issues over matters of fact, issues that allow for objectively and uniquely true answers" (2006, xi-xii).

8] From Protagoras's claim that no universal truths in moral and legal matters exist seems to follow that he rejects the views that an objective public good exists and that there can be a contradiction between this good and the utility of laws. For good reasons, the sophist Thrasymachus argues in Book I of Plato's

or views about the good and just are more beneficial or useful than others implies that he rejects an "absolute relativism" that claims "de gustibus disputandum non est".[9] For him, the sophist is able to persuade the citizens of a *polis* to establish the most useful or beneficial laws and conceptions of justice. Protagoras has been interpreted as an "ethical pragmatist" and "ethical consequentialist" who holds that the better moral beliefs and laws are those which work better and "have better practical results or consequences" (Giorgini 2019, 112).[10]

A main motivation of Socrates's and Plato's philosophical endeavors is to overcome Protagoras's skeptical claims that no moral facts and no moral knowledge applicable to political issues exist.[11] For these two Athenian philosophers, Protagoras's view has a major shortcoming: because of the lack of objective and universal standards, he cannot provide sufficient ethical orientation for political decisions and can only partially overcome moral and legal disagreements, which are disagreements about which actions and laws are right or wrong. To be sure, despite Protagoras's denial of moral knowledge, he is able to suggest a practical criterion to resolve disagreements. This criterion focuses on the results and consequences of actions and laws, and in particular, on whether these are more or less beneficial or useful. Nevertheless, such standards do not satisfy Socrates and Plato. Despite the scholarly disagreement on the question of how to distinguish the historical from the Platonic Socrates, it is rather certain that the historical Socrates, often regarded as the founder of a "scientific ethics", defended a view that is usually called "ethical intellectualism".[12] Essentially, this view claims that virtue is knowledge. This means that a person who has true knowledge of the good will be able to implement this knowledge in her or his actions without getting sidetracked by appetites, passions, inclinations, and the strivings to obtain pleasure and to avoid pain. In contrast to Protagoras, Socrates claims that philosophically accessible moral knowledge exists, which is able to ethically orientate personal and political decisions. This claim is equal to an affirmative answer to the most fundamental question of political epistemology.

Republic that the citizens who hold political power generally use this power to pass laws that are useful for them and not for the whole community (338d-339a). In the light of Thrasymachus's political realism, Protagoras's claim that a "community always makes such laws as are most useful to it" turns out to be an unwarranted generalization and an unconvincing idealistic view. It is neither the political community as a whole that legislates nor do laws always aim at the public good.

9] Protagoras's rejection of an "absolute relativism" leads to the question whether it would not be more appropriate to call him a "moral skeptic" instead of an "ethical or moral relativist"; for the term "moral skeptic", see Mackie 1977, 16-18.

10] Similarly, Protagoras has been interpreted as a representative of a "Utilitarian ethics" (Zehnpfennig 2001, 176).

11] For good reasons, Protagoras has been understood as "Plato's Subtlest Enemy" (Ziglioli 2007). In line with Ziglioli, Giorgini claims, "Plato struggled all his life with Protagoras's thought" (2019, 115).

12] For this judgment and for Socrates's intellectualist predecessors, see Wundt 2019, 1-2; cf. Erler 2007, 433-34; Bonazzi, Forcignanò and Ulacco 2019.

In his dialogues, Plato continues Socrates's research on the relation of moral knowledge to ethical or political decisions. Plato's view on this relation is most clearly expressed in his *Republic*. According to his theory of forms, everything good and just possesses these qualities because it participates in the forms of the good and the just. The good in itself and the other forms constitute an "intelligible region" (*noêtos topos*) that exists separate from the human mind and the world human beings perceive with their senses (*Republic* VI 509d). A long theoretical education allows the *logistikon* of the philosopher and its "intelligence" (*nous*) to access this intelligible region. The *logistikon* and its intelligence are not only able to behold the good in itself and the other forms but are kindred to them (*Republic* VI 490b, X 611e). In the language of contemporary meta-ethics, Plato was an "ethical realist" and "cognitivist" who holds the good to be a moral fact or an objective moral reality about which moral knowledge can be achieved. This knowledge is not only capable of ethically orientating political decisions. The political philosopher is also able to discern a "divine paradigm" (*theion paradeigma*) of the just and good political order in the intelligible region (*Republic* V 472b-473b, VI 500c-501b, VII 517c). According to Plato's famous statement, this divine paradigm cannot be politically realized unless "the philosophers rule as kings or those now called kings and chiefs genuinely and adequately philosophize, and political power and philosophy coincide in the same place" (*Republic* V 473c-d, trans. Bloom).

Protagoras and the two Athenian philosophers vigorously disagree on the basic problem of political epistemology, whether there is any knowledge that can ethically orientate political decisions. A better comprehension of this disagreement is not only significant for a better understanding of the discipline, but of philosophical reasoning in general. In contemporary philosophical language, the disagreement between Plato and Protagoras is a dispute between an ethical realist and an anti-realist. It is further a dispute between a philosopher who believes in the possibility of moral knowledge (about the good and justice), and one who does not. Such disagreements are at the center of contemporary meta-ethical arguments. As the ongoing debate suggests, the discussants are far from overcoming their dissent and from reaching a consensus. This prompts the conclusion that such disagreements are *deep disagreements*. Deep disagreements are disagreements in good faith that cannot be resolved through the use of reasons and arguments (Fogelin 2005, 8, 11).[13] According to Robert J. Fogelin, deep disagreements are resistant to rational solution because of a clash of "underlying principles" or "framework propositions" (Fogelin 2005, 8-9).[14] Fogelin's foundationalist

---

13] Fogelin's definition of deep disagreements could be complemented by the condition that such disagreements are only deep if they do not depend on disagreements on descriptive facts.

14] For a distinction between shallow and deep epistemic disagreements and a sophisticated analysis of the latter, see Lynch 2010, 262-77. Fogelin's article sparked controversies on deep and peer disagreements. For a summary of the debates, see Siegel 2013. For some of the literature on the epistemology of disagreement, see Siegel 2013 and Machuca 2013. Since 2013, several new books on the epistemology of disagreement have appeared.

approach allows for an explanation of the deep disagreement between Protagoras and the two Athenians on the existence of moral knowledge. The reason is that their respective epistemological and political thought is indeed based on "underlying principles" that clash.

Protagoras's political thought rests on his skepticism, relativism, and on his agnosticism. His clause that man is the measure of all things was the opening passage of his work *Alêtheia* (*Truth*). Similarly, the principal clause of Protagoras's philosophy of religion was placed at the beginning of his treatise *On the Gods*:

> Concerning the gods I cannot know whether they exist or not [nor what form (*idea*) they might have], for many are the obstacles that prevent our knowledge: the obscurity of the subject and the brevity of human life (*DK* 80B4, my trans.).

Human beings are able to verify whether the gods exist neither through their senses nor their reason. The impossibility of achieving objective knowledge and truth about the gods can also be derived from Protagoras's claim that man is the measure of all things. From Protagoras's skepticism and relativism follows that his political thought does not acknowledge any objective or universally valid conception of justice and the good. From his agnosticism ensues that he does not recognize anything divine that might be able to orientate political decisions. The human mind and moral, legal, and political actions cannot rely on objective values or any divine or higher standard. Since human beings have no access to any higher authority, they depend exclusively on human judgments and standards such as human benefit, utility, and practical consequences (Giorgini 2019, 107-109). From the perspective of Protagoras's skepticism, relativism, and agnosticism, it seems plausible to reject the existence of any objective moral reality or mind-independent moral facts. Even if such phenomena existed, human beings have no means of accessing them.

In contrast to Protagoras's political philosophy, Plato's political thought rests on his religious and theological convictions. As previously mentioned, he takes the well-ordered city to be a "divine paradigm" (*theion paradeigma*) that should be imitated as much as possible. Plato often states that the forms are divine. In a famous passage, he characterizes the form of the good in a way that clearly suggests that he holds it to be the supreme deity: "the good isn't being but is still beyond being, exceeding it in dignity and power" (*Republic* VI 509b, trans. Bloom).[15] In the *Laws*, Plato does not unequivocally refer to his theory of forms. Nevertheless, he makes the "underlying principle" of his political thought explicit. Countering Protagoras's claim that man is the measure of all things, he emphasizes: "In our view it is God who is preeminently the 'measure of all things', much more so than any 'man', as they say" (*Laws* IV 716c; trans. Saunders).

15] For the equation of the deity with the form of the good, see Zeller 2006, 139. For several arguments for the claim that Plato is the founder of political theology, see Ottmann 2017.

Of course, agnosticism does not conclusively rule out the possibility that an objective moral reality or mind-independent moral facts exist. Skepticism and relativism have many different forms and are controversial and much-discussed views still today.[16] Nevertheless, such "underlying principles" clearly suggest an anti-realist view. If no recourse to a divine reality or a higher standard is possible, the existence of an objective moral reality or of mind-independent moral facts seems unlikely. Another reason for this is that no plausible explanation of the origin of such peculiar phenomena can be given.[17] If there are no moral facts and no objective moral reality, there is no moral knowledge that could ethically orientate political decisions; nothing exists on which such knowledge could be based, or to which it could refer. By contrast, Plato's philosophical theology can explain the existence of an objective moral reality. If the cosmos exists as a divine order that contains "reason" (*nous*), and if human "intelligence" (*nous*) is kindred to this order and able to perceive it, true philosophers are capable of achieving moral knowledge about the mind-independent forms of the good and just. However, Plato has serious difficulties to rationally defend both his philosophical theology and his theory of forms, which are the "underlying principles" of his ethical realism and cognitivism. Of course, these positions can also be defended without recourse to a philosophical theology or Plato's theory of forms.

In the next two sections, this article argues in the tradition of Protagoras that no ultimate moral knowledge exists that is able to ethically orientate political decisions. The main argument is based on the existence of myriad widespread deep disagreements on values, justice, morality, and ethics, which are resistant to rational solution. The reality of these disagreements is a strong argument against the existence of a mind-independent and objective moral reality about which moral knowledge could be achieved.[18] The arguments presented against ethical realism and cognitivism in section II are mainly based on Max Weber's and Isaiah Berlin's views that deep disagreements on values, morality, and ethics exist. The arguments set out against these positions in section III rest on deep philosophical disagreements on social and political justice, which are a characteristic feature of the history of political thought.

16] For the many different forms of skepticism and the debate, see Machuca and Reed 2018. For the many different forms of relativism and the debate, see Baghramian and Coliva 2020.

17] Cf. Mackie's well-known *argument from queerness*: If objective values existed, they would not only "be entities or qualities or relations of a very strange sort", but it would also be very difficult to explain how we could access them (Mackie 1977, 38).

18] This article mainly focuses on the ontological questions whether moral facts and an objective moral reality exist. The important epistemological question of what doxastic attitude one should adopt concerning disagreements on what policy is the best or right to implement will be addressed in a further article. This question revolves around the problem whether in the face of such political disputes one should retain or revise one's belief or whether one should suspend judgment.

II. THE ARGUMENT AGAINST ETHICAL REALISM FROM DEEP DISAGREEMENTS ON VALUES, MORALITY, AND ETHICS: MAX WEBER AND ISAIAH BERLIN

The world of today is not only a world characterized by moral disagreements and value conflicts, but a world that is well aware of such disagreements and conflicts. People disagree on essentials such as religion, values, politics, and the good life, and on specific moral issues such as abortion, euthanasia, gay marriage, affirmative action, taxation, and human cloning. Influenced by Nietzsche's thoughts on the struggle between opposing values[19], Max Weber developed his own theory of value conflict. Weber claims that modernity is characterized by an irresolvable pluralism and struggle of values and ideals (Weber 1949, 17-18). The political struggle between politicians and parties is part of this bigger and more fundamental fight. One main aim of politics is to fight out this battle of antinomic values and ideals. Weber's understanding of politics follows the realist tradition that rejects "ideal theory", considers only pure factuality, and holds power to be the central category in politics. Politics means "striving for a share of power or for influence on the distribution of power, whether it be between states or between the groups of people contained within a single state" (Weber 1994a, 310). Weber states that "all politics is essentially struggle" (Weber 1994b, 219). In this struggle, every citizen and politician has to choose which value or ideal he or she holds to be "God" and which one the "Devil" (Weber 1949, 17-18).

Weber's answer to the central question of political epistemology is unequivocally negative: no moral knowledge exists that is able to ethically orientate a politician's decisions and actions in the struggle of values and ideals. In this context, Weber emphasizes "the 'limits' of ethics" (Weber 1949, 15). Both the social sciences and ethics suffer from argumentative limits concerning a rational or scientific grounding of values and conceptions of justice. Both are unable to scientifically answer the questions of how to act and how to live. Scientific reason is unable to determine which of two or more conflicting goals or values is more desirable or worth choosing. Weber addresses these limits in a crucial statement that also refers to his "decisionism",

> Even such simple questions as the extent to which an end should sanction unavoidable means, or the extent to which undesired repercussions [*Nebenerfolge*] should be taken into consideration, or how conflicts between several concretely conflicting ends are to be arbitrated [*schlichten*], are entirely matters of choice or compromise. There is no (rational or empirical) scientific procedure of any kind whatsoever which can provide us with a decision here (Weber 1949, 18-19).

Weber's central claim is that there "is no (rational or empirical) scientific procedure of any kind whatsoever" that could arbitrate between values or resolve moral disagreements and value conflicts. His point about the argumentative limits of ethics and the social sciences is not that the different parties in moral disagreements and value

---

19] For the influence of Nietzsche's views regarding value conflict on Weber, see Knoll 2019, 117-23.

conflicts are unable to defend their views with reasons and arguments. Rather, Weber claims that such disagreements and conflicts cannot be definitely and conclusively argued out.[20] For Weber, an *ultimate* rational grounding of values, norms, or ideals is impossible.

Similarly to Weber, Isaiah Berlin reflects on value pluralism. Despite his view that values are objective, Berlin holds that they clash frequently. Values are human ends that differ in "some profound, irreconcilable way" and are "not combinable in any final synthesis" (Berlin 1997, 8). According to Berlin, "values can clash" or "be incompatible between cultures, or groups in the same culture, or between you and me"; also values "may easily clash within the breast of a single individual" (Berlin 1997, 10). Berlin exemplifies value conflicts with someone who believes in always telling the truth and someone who believes that this "can sometimes be too painful and too destructive" (Berlin 1997, 10). Berlin's example illustrates a more profound and general conflict between two main approaches to morality and ethics, i.e. between a deontological and a consequentialist approach. A deontological ethics claims that certain actions are always right or wrong in certain situations, no matter what the consequences are.[21] A consequentialist ethics, on the contrary, holds that the judgment about the moral rightness or wrongness of an action depends exclusively on the quality of its foreseeable consequences. These two approaches to ethics are irreconcilably opposed to each other.[22] They do not allow for consensus or compromise. Rather, they require a choice about whether actions or consequences are the appropriate domain of ethical assessment. Weber conceptualizes the deep disagreement between these two approaches to ethics as a conflict between an "ethics of responsibility" (*Verantwortungsethik*) and an "ethics of conviction" (*Gesinnungsethik*) (Weber 1994a, 357-369; cf. Knoll 2019).

Berlin further illustrates value conflicts with the collision of liberty and equality, two major human values and goals. When different individuals exercise their liberty, it usually leads to inequality among them. The accomplishment of equality, by contrast, frequently demands the restraint of liberty; e.g., liberty needs "to be curtailed in order to make room for social welfare" (Berlin 1997, 10-11). The clash between equality and liberty is one of the key conflicts pertaining to the political sphere. In contemporary political philosophy, this conflict is represented by the deep disagreement between the theories of Rawls and Nozick, which is addressed in section III.

20] Weber repeatedly talks about the "Unaustragbarkeit" of such conflicts (Weber 1949, 11, 16); cf. Knoll 2019.

21] For example, for a deontological ethics such as Kant's lying or stealing are wrong actions in all situations; cf. Broad 1930, 206.

22] A deontological ethics, defended most famously by Immanuel Kant, goes back to the Hebrew-Christian ethic. The term "consequentialism" became established in the wake of G.E.M. Anscombe's article "Modern Moral Philosophy", which was published in 1958. The term was first coined in this paper (Anscombe 1958, 7-11). For Anscombe, a consequentialist ethics is "quite incompatible with the Hebrew-Christian ethic" (1958, 8).

According to Berlin, value conflicts are an essential feature of the human condition. In such conflicts "we are doomed to choose, and every choice may entail an irreparable loss" (Berlin 1997, 11). The "normal human situation" is that "ends equally ultimate, equally sacred" and "entire systems of value" do "come into collision without possibility of rational arbitration" (Berlin 2013, 94; cf. 94-99). On the basis of his views on value pluralism and collisions of values, Berlin rejects the utopian political notion of a perfect state as an unrealistic ideal and dangerous political goal; "the search for perfection does seem to me a recipe for bloodshed" (Berlin 1997, 12-13, 15).

Although Weber and Berlin did not use the term "deep disagreements", both scholars claimed decades before Fogelin that there is no possibility to rationally arbitrate between ideals or to rationally resolve value conflicts. While Weber and Berlin focus on deep disagreements on values and ideals, Fogelin examines deep disagreements on moral issues such as abortion and affirmative action (2015, 8-10). Despite Berlin's view that values are objective, he denies the existence of moral knowledge that could help to choose between conflicting values, e.g., between liberty and equality.[23] Weber denies the existence of any given or predefined "hierarchical ordering of values" (Weber 1949, 19). Deep and irresolvable disagreements on values, morality, and approaches to ethics are a strong argument against the existence of an objective moral reality that could ethically orientate political decisions. If a mind-independent order of values existed, it would be likely that after thousands of years of philosophical research, some moral knowledge about them had been discovered. However, as the contemporary debate shows, we are far from any agreement on such kind of knowledge that would allow us to solve value conflicts and moral disagreements.

III. THE ARGUMENT AGAINST ETHICAL REALISM FROM DEEP DISAGREEMENTS ON JUSTICE

Since ancient Greek thought, many philosophers have argued that justice is our central moral, social, and political virtue. Plato argues that true philosophers are able to acquire knowledge about justice; Protagoras denies this. In the contemporary debate, Rawls developed a theory of "justice as fairness" that is able to ethically orientate political decisions.[24] Rawls explains his ideal theory approach to political justice: "We are in the

---

23] For a recent study on Isaiah Berlin, Stuart Hampshire, and Bernard Williams's shared political realism and "shared skepticism about the power of philosophical ethics", see Hall 2020, 5. On p. 27, Hall criticizes that many commentators deny or overlook that Berlin rejects "the idea that moral values or ends exist in some mind-independent sense. […] I suspect that this is because they have been led astray by Berlin's insistence that his value pluralism is consistent with a belief in the objectivity of morality". However, Berlin's denial of the existence of moral knowledge appears to contradict his claim that values are objective. If there is no moral knowledge, how could it be possible to show or prove that values are objective?

24] In 1971, Rawls explains about his conception of justice as fairness that it "constitutes the most appropriate moral basis for a democratic society" (Rawls 1999, Preface, xix). In his later works, Rawls insists

way of describing an ideal arrangement, comparison with which defines a standard for judging actual institutions, and indicates what must be maintained to justify departures from it" (Rawls 1999, § 36, 199). Despite Rawls's conviction that his ideal theory is able to ethically orientate political decisions, he does not claim that his conception of justice as fairness is true: "I do not claim for the principles of justice proposed that they are necessary truths or derivable from such truths" (Rawls 1999, § 4, 19). In 1985, Rawls again declared that he "should like to avoid" "claims to universal truth" for his conception of justice as fairness (223). In his *Theory of Justice* he makes the more modest claim that in a fair "initial situation" a rational agreement on his conception of justice can be achieved (1999, § 2, 11).[25] This agreement should enable a society to cope with disagreements over the good by establishing a just political framework that allows opposing conceptions of a good life to coexist. In *Political Liberalism*, Rawls claims that in a pluralist society with opposing comprehensive religious, philosophical, and moral doctrines an "overlapping consensus" on a political conception of justice can be reached (Rawls 2005, IV §§ 1-8, 133-72; cf. Mason 1993, 9-12). A rational agreement or an "overlapping consensus" on social or political justice is, however, not a realistic goal.

As early as 1917, Max Weber argued in a compelling passage that there are not only irreconcilable conceptions of the good, but also of the just. He illustrates "the 'limits' of ethics" by referring to deep disagreements on social and political justice,

> The implications of the postulate of 'justice' cannot be decided unambiguously by any ethic. Whether one, for example – as would correspond most closely with the views expressed by Schmoller – owes much to those who achieve much or whether one should demand much from those who accomplish much; whether one should, e. g., in the name of justice [...] accord great opportunities to those with eminent talents or whether on the contrary (like Babeuf) one should attempt to equalize the injustice of the unequal distribution of mental capacities through the rigorous provision that talented persons, whose talent gives them prestige, must not utilize their better opportunities for their own benefit – these questions cannot be definitely answered [*dürfte aus "ethischen Prämissen" unaustragbar sein*]. The ethical problem in most social-political issues is, however, of this type (Weber 1949, 15-16).

In the literature, this passage has been discussed under the name of "Weber's 'Babeuf' antinomy" (Turner and Factor 1984, 35). However, in the dense and complex passage, Weber shows that it is a general feature of modernity that there are irreconcilable conceptions of social and political justice. Weber refers not only to

---

on the distinction "between moral and political philosophy" and emphasizes that he defends a "strictly political conception of justice". Such a conception is "limited to the domain of the political" and "expresses political values" (Rawls 2005, xv, 450; cf. 439). Nevertheless, Rawls still understands political philosophy as applied moral philosophy, and justice claims as moral or ethical claims. He explains, "Political conceptions of justice are themselves intrinsically moral ideas" and are "kind of normative values" (Rawls 2005, 484).

25] In the "Preface for the Revised Edition", Rawls hopes that "justice as fairness will seem reasonable and useful" (1999, xi).

Babeuf's view of distributive justice, but to several irreconcilable interpretations of "the postulate of 'justice'," e.g. the conflict of the "performance principle" with what Rawls calls the "principle of redress" (Rawls 1999, § 17, 86).[26] Weber's main point is that the disagreement between these irreconcilable conceptions of justice "cannot be conclusively argued out based on 'ethical premises'" (*dürfte aus "ethischen Prämissen" unaustragbar sein*). In the last phrase of the passage, Weber makes clear that such "'limits' of ethics" apply to most "social-political issues".

Deep disagreements on justice are not only a modern phenomenon but were quite common in the ancient world. Plato's and Aristotle's writings contain central insights regarding conflicts about political justice. These conflicts are based on irreconcilable views of a just distribution of political offices and power. Concerning a just distribution of political power, Plato distinguishes between egalitarian and proportional justice. Egalitarian political justice is identical with democratic justice. It asks for "arithmetic" or "numeric" equality claiming that in the distribution of political power all male citizens should receive equal shares. In contrast, proportional justice asks for "geometric" or "proportional" equality claiming that it is just to allot equal shares only to equals, not to every citizen. The inequalities that politically matter for Plato and in proportion to which he wants to distribute political power are "education" (*paideia*) and "virtue" (*aretê*) (*Laws*, VI 757b-e). Aristotle unequivocally shares this meritocratic view (*Politics* III 13, 1283a25).[27]

According to the political theory of merit Aristotle presents in the *Nicomachean Ethics*, proportional political justice is applied in four different conceptions of distributive justice. All of these conceptions are defended by different groups of citizens who all have distinct political convictions. The democrats favor "freedom" as the appropriate criterion of merit, the supporters of oligarchy "wealth", the aristocrats "virtue" (*arête*), and an unnamed fourth group "noble birth" (*Eth. Nic.* V 6, 1131a24-29). Therefore, Aristotle distinguishes between four different conceptions of distributive or political justice: the democratic, the oligarchic, the aristocratic, and an undesignated fourth conception. Each of these conceptions is linked with its corresponding political system and justifies its specific distribution of political power. To argue that every free-born male citizen should get an equal share in political power is identical with the defense of democracy. To advocate the distribution of political offices in proportion to wealth is the same as to support oligarchy. To argue that political power should be given only to virtuous citizens is identical with the defense of aristocracy.

In contrast to the *Nicomachean Ethics*, in the *Politics*, Aristotle opposes "numeric" or "arithmetic" (*arithmô*) equality, like Plato, to equality "according to merit" (*kat' axian*) (*Politics* V 1, 1301b29f.; VI 2, 1317b4; cf. Plato, *Laws*, 751d). This opposition equals to the fundamental antagonism and deep disagreement between arithmetic and

26] For Babeuf's view of distributive justice, see Fleischacker 2004, 55, 76, 160-61.

27] For Plato's and Aristotle's theories of political justice, see Knoll 2010 and Knoll 2016.

proportional equality and justice. Proportional justice and egalitarian justice are not only opposites but are in general logically irreconcilable. They mutually exclude each other and represent two competing and contradictory understandings of justice that are connected to different sets of rules. This means that they cannot both be applied to solve the same distribution problem without nullifying each other (cf. Herwig 1984, 97-99). Either justice is equality *only for equals* or (in an exclusive sense) *for all* (equals and unequals). The formal principle of proportional justice determines that *only equals* should get equal shares, while unequals should be allotted unequal shares. This formal principle is only in one case reconcilable with the formal principle of egalitarian justice that determines that *all* should get equal shares: in the case that everyone is equal. If there are unequal citizens in a distribution of political power, such as rich and poor, an oligarchic conception of proportional justice requires that unequal shares be allocated to them in proportion to their unequal wealth. For oligarchic justice, to distribute equal shares in such a case would be unjust. If there were virtuous and non-virtuous citizens who had claims in the same distribution, to distribute equal shares would also be unjust from the perspective of an aristocratic conception of justice. Aristocratic justice requires allotting unequal shares to citizens in proportion to their unequal virtue. However, if the most virtuous citizens were poor, such a distribution would be unjust from the perspective of oligarchic justice. From the perspective of democratic justice, however, a distribution of unequal shares is unjust. Democratic justice, which is egalitarian justice, requires allotting equal shares to every citizen, which can be achieved through a universal and equal suffrage. Census suffrage, which is based on oligarchic justice, is unjust from the prevailing modern perspective of egalitarian justice.

In the 19th century, while conflicts about political justice persisted, deep disagreements on social justice entered the stage. Karl Marx argues that the distribution of social wealth should be based exclusively on individual needs and thus be independent from unequal labor contributions to society (Marx 2009; cf. Knoll 2018, 33-39). Several contemporary political philosophers, such as Michael Walzer and David Miller, recognize the importance of the "needs principle" for just distributions of social benefits (Walzer 1983, 25-26, 64-94; Miller 2003, 203-229; cf. Rawls 1999, § 47, 268). However, from the perspective of the performance or merit principle, it is extremely unjust if persons who contribute more to society receive significantly less than those who contribute less. Nevertheless, in social distributions Marx does not acknowledge or approve the "bourgeois" performance principle at all. His own distributive principle is not only irreconcilable with the performance principle but nullifies it and even allows it to be reversed.

A similar conflict and deep disagreement on social justice is at the center of contemporary political philosophy. Although the performance principle and meritocracy are still advocated by political philosophers today (Miller 2003, 177-202), most contemporary political theorists are endorsing egalitarian justice. Only three years after John Rawls published his egalitarian conception of social justice in 1971, the

libertarian Robert Nozick forcefully objected. Rawls's basic moral conviction is that inequalities of birth and natural endowments are undeserved and hence call for social redress or compensation (Rawls 1999, § 17, 86-93). Against this conviction, Nozick argues that even if people might not deserve their natural endowments, they still rightfully own them, have a justified claim to them, and are free to use them for their own benefit (1974, 225). Therefore, Nozick deeply disagrees with the amount of taxation and redistribution of income Rawls's conception of social justice requires. With a polemic statement, Nozick attempts to expose the injustice of Rawls's conception, which treats persons' natural endowments and abilities as means for other people's welfare: "Taxation of earnings from labor is on a par with forced labor. Some persons find this claim obviously true: taking the earnings of *n* hours labor is like taking *n* hours from the person; it is like forcing the person to work *n* hours for another's purpose" (1974, 169; cf. 168-69, 228). The deep disagreement between Rawls's and Nozick's views on social justice could be interpreted as a resumption of the deep disagreement between advocates of the performance or merit principle and Marx's views on social justice.

The preceding arguments that show that deep disagreements on social and political justice are a characteristic feature of the history of political thought are significant for both meta-ethics and political epistemology. The reason is that such disagreements, resistant to rational solution, provide a strong case against ethical realism and cognitivism. More than two thousand years of systematic philosophical research on justice has neither uncovered objective truths about the matter nor led to any agreement among scholars. Looking at it the other way round, deep disagreements among notions of social and political justice exist because there are no objective moral reality and no moral facts. There is nothing in the world to back up *one* of the conflicting philosophical views on justice. Neither is there a procedure using moral facts to solve moral disagreements and to show that one view is wrong and one is right (cf. Waldron 1999, 177-78). If the ethical realist, who like Plato claims that moral facts or objective moral values exist, were able to both cogently demonstrate that such an objective moral reality exists and to successfully communicate his knowledge about it, he could refute the anti-realist and moral skeptic. In this case, such a philosopher could solve moral disagreements and value conflicts by pointing to *the* moral truths and use them as standards to assess antagonistic views and values. However, until today no undeniable moral truths have been tracked down. The ethical realist has the burden of proof. She should be able to show the anti-realist and moral skeptic how such facts and values can be detected and accessed. If we were to accept the belief that objective values are somehow part of the structure of the world, we should be able to track or comprehend their nature and their place in this world. Likewise, someone who claims an empirical fact, say, the existence of centaurs, should be able to prove their existence to the skeptic who denies it. If the moral skeptic has to identify an objective moral reality with an inaccessible divine or metaphysical reality like the one Plato claims for the form of the good, she has good reasons to doubt its existence (cf. section I).

The *argument against ethical realism from deep disagreements on justice* rests on historical and actual disagreements among philosophers, researchers, and scholars, who represent the alleged vigor of human reasoning. Their disagreements are based on elaborate but opposing philosophical theories each of which are defended with carefully constructed arguments. According to an old-fashioned view, philosophy is committed to finding out the truth. However, after more than two thousand years of continuous philosophical disagreements on social and political justice, it is very likely that there exists no objective truth about justice. Considering the enormous amount of past philosophical efforts, there is no good reason to be optimistic that a consensus on justice could be reached in the future (Ribeiro 2011, 3-25, 18-21).

IV. CONCLUSION

The mere fact of continuous deep disagreements on justice, values, and morals does not, of course, conclusively rule out the possibility that an objective moral reality or moral knowledge exists. It is conceivable that some philosopher in the past discovered truths about such matters or will do so in the future. Nevertheless, the debate until now and the persistence of widespread deep disagreements on justice, values, and morals do not encourage hopes that political epistemologists will reach consensus or discover ultimate moral knowledge that could ethically orientate political decisions. Instead of searching for moral knowledge, political philosophers should rather focus on how to practically and ethically deal with such disagreements.

Deep disagreements on justice, values, and moral issues occur in particular in the political arena. Theoretical disagreements among citizens often turn into practical conflicts and lead to fights and violence. In such conflicts, opponents frequently degrade each other as unreasonable, ignorant, backward, prejudiced, vicious, and the like, which is often an oversimplification of the matter. The value conflict between the life of a fetus and the liberty of a woman can easily turn into a conflict between persons and may in extreme cases lead to violent actions by anti-abortion radicals. Since the early 2020s, during the Covid-19 pandemic, in many countries citizens disagreed with the restrictions of freedom and the lockdowns imposed by their governments. Protests against these measures led to violence and arrests. In the worst case, theoretical disagreements and practical conflicts can lead to civil unrest and civil wars. Therefore, the central question is how to reduce practical conflicts and promote a peaceful coexistence among those who disagree but have to live together.[28] Instead of searching for moral facts and moral knowledge, political epistemologists should rather aim at knowledge that promotes reaching these goals.

There is knowledge that allows citizens to reduce practical conflicts and to promote a peaceful coexistence. Such kind of knowledge is rather practical than theoretical. An

---

28] For preliminary answers to this question, see Knoll 2020a.

early model for it is what Aristotle calls "practical knowledge" (*epistêmê praktikê*), which is based on experience and the intellectual virtue "prudence" (*phronêsis*).[29] In Book V of his *Politics*, Aristotle suggests an empirical, inductive, and comparative method to gain knowledge about "upheaval", "sedition" or "revolution" (*stasis*). In Book V, Aristotle demonstrates how the empirical study of history and of the different causes of factional conflict, which includes research on human nature, allows for a better understanding of such phenomena. If we know the various causes of past uprisings and revolutions, we can use this knowledge to prevent future ones. Aristotle's search for knowledge about civil disagreements and civil unrest is oriented by the normative political goal of stabilizing political systems. The insights of his political studies are still relevant for the world of today and its civil wars (cf. Knoll 2020b). Aristotle's foundation of an empirical political science is connected to his political ethics that aims at the good and happy life of the members of the political community. Drawing on this tradition, political epistemologists and a contemporary political ethics should develop strategies for practically and ethically dealing with deep disagreements on justice, values, and morals.

Manuel.Knoll@lrz.uni-muenchen.de

REFERENCES

Anscombe, G.E.M. 1958. Modern Moral Philosophy. *Philosophy* 33(124): 1-19.

Baghramian, Maria, and Annalisa Coliva. 2020. *Relativism*. London and New York: Routledge.

Berlin, Isaiah. 1997[1988]. The Pursuit of the Ideal. In Isaiah Berlin, *The Proper Study of Mankind: An Anthology of Essays*, ed. Henry Hardy and Roger Hausheer. London: Chatto & Windus, 1-16.

———. 2013[1972]. The Originality of Machiavelli. In Isaiah Berlin, *Against the Current. Essays in the History of Ideas*, edited by Henry Hardy (Second edition). Princeton and Oxford: Princeton University Press, 33-100.

Bonazzi, Mauro, Filippo Forcignanò, and Angela Ulacco, eds. 2019. *Thinking, Knowing, Acting: Epistemology and Ethics in Plato and Ancient Platonism*. Leiden and Boston: Brill.

Broad, C.D. 1930. *Five Types of Ethical Theory*. London: Kegan Paul, Trench, Trubner & Co., Ltd.

Edenberg, Elizabeth, and Michael Hannon, eds. 2021. *Political Epistemology*. Oxford: Oxford University Press.

Elm, Ralf. 1996. *Klugheit und Erfahrung bei Aristoteles*. Paderborn: Schöningh.

Erler, Michael. 2007. *Platon = Überweg. Grundriss der Geschichte der Philosophie. Die Philosophie der Antike*. vol. 2/2, edited by Hellmut Flashar. Basel.

Eustacchi, Francesca. 2016. At the Source of the Multifocal Approach: Relations in the Sophistic Context. In *By the Sophists to Aristotle Through Plato*, edited by Elisabetta Cattanei, Arianna Fermani, and Maurizio Migliori. Sankt Augustin: Academia, 33-66.

Fleischacker, Samuel. 2004. *A Short History of Distributive Justice*. Cambridge, MA: Harvard University Press.

29] For Aristotle's theory of "prudence" (*phronêsis*), see in particular his *Nicomachean Ethics*, Book VI; cf. Elm 1996.

Fogelin, Robert J. 2005[1985]. The Logic of Deep Disagreements. *Informal Logic* 25(1).

Giorgini, Giovanni. 2016. The Power of Speech: The Influence of the Sophists on Greek Politics. In *Philosophy and Political Power in Antiquity*, edited by Cinzia Arruzza and Dmitri Nikulin. Leiden and Boston: Brill, 9-40.

———. 2019. Does Democracy Necessarily Rest on Relativism? The Origins of the Debate in Ancient Greece. In *Philosophie für die Polis. Akten des 5. Kongresses der Gesellschaft für antike Philosophie 2016*, edited by Christoph Riedweg. Berlin and Boston: De Gruyter, 95-122.

Gowans, Christopher W. 2000. Introduction. Debates about moral disagreements. In *Moral Disagreements. Classic and Contemporary Readings*, edited by Christopher W. Gowans. London and New York: Routledge, 1-43.

Hall, Edward. 2020. *Value, Conflict, and Order: Berlin, Hampshire, Williams, and the Realist Revival in Political Theory.* Chicago: University of Chicago Press.

Hannon, Michael, and Jeroen de Ridder, eds. 2021. *The Routledge Handbook of Political Epistemology.* London and New York: Routledge.

Herwig, Dagmar. 1984. *Gleichbehandlung und Egalisierung als konkurrierende Modelle von Gerechtigkeit. Eine systematische Analyse.* München: Fink.

Knoll, Manuel. 2010 Die distributive Gerechtigkeit bei Platon und Aristoteles. *Zeitschrift* für *Politik* (ZfP) 1, 3-30.

———. 2016. The Meaning of Distributive Justice for Aristotle's Theory of Constitutions. *ΠΗΓΗ/FONS. Revista electrónica de estudios sobre la civilizatión clásica y su recepción* (OJS, www.uc3m.es/pege), vol. 1, 57-97.

———. 2018. Deep Disagreements on Social and Political Justice: Their Meta-Ethical Relevance and the Need for a New Research Perspective. In *New Perspectives on Distributive Justice. Deep Disagreements, Pluralism, and the Problem of Consensus*, edited by Manuel Knoll, Stephen Snyder, and Nurdane Şimşek. Berlin/Boston: De Gruyter, 23-51.

———. 2019. Max Weber on Politics, Reason, and the Clash of Values and Approaches to Ethics. *Dîvân. Journal of Interdisciplinary Studies* 24, 47(2): 111-40

———. 2020a. Deep Disagreements on Values, Justice, and Moral Issues: Towards an Ethics of Disagreement. *Trames,* Special Issue on Disagreements, vol. 24(3): 315-38.

———. 2020b. Verfall und Stabilität politischer Ordnungen. Zur Aktualität des antiken politischen Denkens. *Philosophisches Jahrbuch* 2, 207-26.

Ladd, John, ed. 1985[1973]. *Ethical Relativism.* Lanham, New York, and London: University Press of America.

Lynch, Michael P. 2010. Epistemic Circularity and Epistemic Incommensurability. In *Social Epistemology*, edited by Adrian Haddock, Alan Millar, and Duncan Pritchard. Oxford: Oxford University Press, 262-77.

Machuca, Diego E. 2013. Editor's Introduction. In *Disagreement and Skepticism*, ed. by Diego E. Machuca. New York and London: Routledge.

Machuca, Diego E., and Baron Reed, eds. 2018. *Skepticism: From Antiquity to the Present.* London: Bloomsbury Academic.

Mackie, J. L. 1977. *Ethics. Inventing Right and Wrong.* London: Penguin.

Marx, Karl. 2009. *Critique of the Gotha Programme.* Dodo Press.

Mason, Andrew. 1993. *Explaining Political Disagreement.* Cambridge: Cambridge University Press.

Miller, Alexander. 2013. *Contemporary Metaethics. An Introduction.* 2nd ed. Cambridge: Polity Press.

Miller, David. 2003[1999]. *Principles of Social Justice.* Third printing. Cambridge, MA.: Harvard University Press.

Nozick, Robert. 1974. *Anarchy, State, and Utopia.* Oxford: Basic Books.

Ottmann, Henning. 2001a. *Geschichte des politischen Denkens. Die Griechen. Von Homer bis Sokrates*. vol. 1/1. Stuttgart: Metzler.

———. 2001b. *Geschichte des politischen Denkens. Die Griechen. Von Platon bis zum Hellenismus*. vol. 1/2. Stuttgart: Metzler.

———. 2017. Platon als Begründer der politischen Theologie. In *Platons Nomoi. Die politische Herrschaft von Vernunft und Gesetz*, edited by Manuel Knoll and Francisco L. Lisi. Baden-Baden: Nomos, 49-63.

Rawls, John. 1985. Justice as Fairness: Political not Metaphysical. *Philosophy & Public Affairs*, vol. 14, no. 3, 223-51.

———. 1999[1971]. *A Theory of Justice*. Revised Edition. Cambridge: Belknap Press of Harvard University Press.

———. 2005[1993]. *Political Liberalism*. Expanded edition. New York: Columbia University Press.

Ribeiro, Brain. 2011. Philosophy and Disagreement. *Crítica. Revista Hispanoamericana de Filosofía*, vol. 43, no. 127.

Rüther, Markus. 2015. *Metaethik zur Einführung*. Hamburg: Junius.

Siegel, Harvey. 2013. Argumentation and the epistemology of disagreement. *OSSA Conference Archive*, Paper 157, 1-22.

Strauss, Leo, and Joseph Cropsey. 1987. *History of Political Philosophy*, 3rd edition. Chicago and London: The University of Chicago Press.

Tersman, Folke. 2006. *Moral Disagreement*. Cambridge: Cambridge University Press.

Turner, Stephen P., and Regis A. Factor. 1984. *Max Weber and the Dispute over Reason and Value: A Study in Philosophy, Ethics, and Politics*. London: Routledge.

Waldron, Jeremy. 1999. *Law and Disagreement*. Oxford: Oxford University Press.

Walzer, Michael. 1983. *Spheres of Justice. A Defense of Pluralism and Equality*. New York: Basic Books.

Weber, Max, 1949[1917]. The Meaning of 'Ethical Neutrality' in Sociology and Economics. In Max Weber, *The Methodology of the Social Sciences*, trans. and ed. Edward Shills and Henry Finch. Illinois: The Free Press of Glencoe, 1-47.

———. 1994a[1919]. The Profession and Vocation of Politics, in: *Weber. Political Writings*, edited by Peter Lassmann and Ronald Speirs. Cambridge: Cambridge University Press.

———. 1994b[1918]. Parliament and Government in Germany under a New Political Order in: *Weber. Political Writings*, edited by Peter Lassmann and Ronald Speirs. Cambridge: Cambridge University Press.

Wundt, Max. 2019[1907]. *Der Intellektualismus in der griechischen Ethik*. Scholars Select (public domain in the US): Wentworth Press.

Zehnpfennig, Barbara. 2001. *Platon zur Einführung*. 2nd edition. Hamburg: Junius.

Zeller, Eduard. 2006[1883]. *Grundriss der Geschichte der griechischen Philosophie*. Elibron Classics.

Ziglioli, Ugo, 2007. *Protagoras and the Challenge of Relativism. Plato's Subtlest Enemy*. Aldershot and Burlington: Ashgate.

On the Meaning of Democracy in the European Union

Jaanika Erne
University of Tartu

Abstract. Using some of Immanuel Kant's ontological categories, and the categories of meaning, the author shows some methodological difficulties met when defining democracy in EU law: the difficulties with categorization because there do not exist complete systems of categories but categories differ and give different research outcomes; with the concept of democracy that is founded upon the recognition of the very absence of any definite foundation and is a multidimensional concept having among others legal and political dimension; with law and politics being autonomous systems with distinct discourses and understandings of categories; with the EU lacking its own legal definition of democracy, which is but in the EU's political acts. The article shows the turning point from empirical (doing, legal practice) to explaining (theory) at the time a researcher poses the question "Why?".

Keywords: democracy, EU, meaning, categorization.

The first part of the article discusses Immanuel Kant's classification of categories, the turn to the categories of meaning, and the postmodern understanding of the necessarily contingent nature of categorizations. Having worked with conceptualizations on the borders of law and politics, I have looked for a link between legal and political (see Kennedy 1980) that are the sides of the same society (Lefort 2007). The second part of the article tries to infer something about the concept of democracy in EU law based on the seven categories selected according to the contemporary flexible understandings of categorization. I show the methodological difficulties that arise category-by-category when trying to explain something related to a particular category. The conclusions are: a) that one cannot explain a political concept based on the objective validity of categories (true about any concept because of the absence of the final ground), and b) that asking for the meaning of democracy requires categories of meaning that cannot be objectively fixed either. The main conclusions are that the legal practice (as empirical) and theories (as asking for meaning or political theories) are autonomous. For the reason empirics in itself cannot explain empirics (Waltz 1979:4), the structure of an empirical analysis can only circle – at best, one could compare an EU legal definition with a meta-definition. Unfortunately, both - an EU legal definition of democracy and a fixed meta-definition of democracy are absent, and one could confirm that democracy "is founded upon the recognition of the very absence of any definite foundation" (Marchart 2007).

For these reasons, a researcher has to ask for meaning or make meaning, going beyond empirics. If empirics consists of undetermined concepts, only the broadest research units are available. In the spheres of knowledge and political, as they have progressed over the last 2000 years, categories have been considered the most abstract research units under which to organize thoughts (Thomasson 2019).

Therefore, the article derives from the Kantian ontological categories as the anchoring points for analysing the principle of democracy. At the same time, the Kantian categories are limited in explanation because they do only allow circling around a certain aspect of contingency. – For example, when we answer the "What?" question, we are constrained with "What?", when we answer the "Who?" question, we are constrained with "Who?", and so forth. This seems similar to Kenneth Waltz's observation that we go around according to the pre-composed rules of research about how to answer a particular question but at the moment we want to understand the reality or explain the research object situated in reality, we are constrained and actually cannot say anything true about the reseacrh object (1979). While ontological categories deny the relative normativity of international law (Fastenrath 1993), the article shows how an empirical research turns into an evaluation at the point a researcher poses the question "Why?", which was directly excluded from the Kantian ontological categories, maybe because of the fear that the question "Why?" would open the door to speculation and meaning-making?

I. CATEGORIZATION BY IMMANUEL KANT AND THE TURN TO THE CATEGORIES OF MEANING

The Kantian categories have been named ontological categories because although Kant distinguishes between the metaphysics of experience (one can name it also as the world of experience or nature, theoretical philosophy, "how things are", appearances, subjectivity, empirics, facts) and the metaphysics of morals (practical philosophy, beliefs of how things ought to be) in his understanding of categories, he does not ask for meaning even when he talks about judgment.

Kant divided categories into four respects. Under each respect, there were three alternative classifications that categorized objects not meanings, for example: quantity (universal,[1] particular,[2] singular[3] - which bring to three corresponding categories under the first group of Kantian categories: quantity – unity, plurality, and totality) (Thomasson 2019). The second group of Kantian categories consists of quality (with subcategories: reality, negation, and limitation) (Thomasson 2019). The third group constitutes of relation (explained as: categorical, hypothetical, disjunctive - with subcategories: inherence and subsistence,[4] causality and dependence,[5] and community (Thomasson 2019).[6] The fourth group constitutes of modality (with subcategories:

1] All swans are white.

2] Some swans are white.

3] Swan Cygmund is white.

4] Substance and accident.

5] Cause and effect.

6] Reciprocity.

possibility, existence, and necessity) (Thomasson 2019). For Kant, this allegedly was an exhaustive table of categories (Thomasson 2019). The critique has included the general doubt as to whether an exhaustive categorization would reveal itself to a human being whose thought is in constant change, while much of the human thought has proven wrong throughout history. A common answer can be that a human decision should be made according to the best evidence available.

Edmund Husserl supplemented the Kantian ontological categories with the categories of meaning (Thomasson 2019). The distinguishing point between Kant and Husserl was that the categories of objects as formal essences find their expression in meaning.

Husserl distinguished between the categories of objects (ontological categories that exist as a matter of empirical fact – the descriptive formal categories that categorize objects not meanings), and categories of meaning (as the way one can think about objects).

Similarly to Isaiah Berlin's method of beginning by answering the question "What does x mean?" (Berlin 2013), Husserl suggested deriving ontological categories from the categories of meaning because "pure truths concerning meaning can be transformed into pure truths concerning the object" (Thomasson 2019). The approach to know what something means prior to verifying that something, places asking for meaning before the investigation of its truth (Berlin 2013). That way, context is important because it can give meaning to the object of a conversation. On the other hand, the term "meaning" has also ambiguous and wide meaning containing seeminglessness, relativity and our knowledge of facts (empirics) being clearly limited by interpretation (Berlin 2013). So, when we understand concepts as post-foundationalists, we do not only give meaning to them but we also make meaning (see also Schiappa 2003).

The attempts to build an exhaustive system of categories allegedly fell out of fashion in the 20th century with the understandings that increasing the number of categories would in itself not explain anything (Waltz 1979, 115), i.e. there cannot be one correct system of categories due to human subjectivity. So, postmodern research allows many different and changing sets of categories, which do not constitute a single fixed system in the contemporary understanding. The broadest categories are space and time. What concerns the rest, then different authors have offered their own differing classifications. For the reason there is no requirement to be exhaustive, a researcher is allowed to build his/her system in research on only some categories. The main categorization in this article is bases on the seven questions: "Which?", "Where?", "When?", "How?", "Who?", "What?", "Why?" that with the exception of "Why?" associate with the Arisotelian and Kantian systems of categories – with the names "categorical", "spatial", "temporal", "procedural", "identitarial", "substantive" and "justificatory". "The justificatory contains both explanation and interpretation. These seven categories constitute the logical tool for the analysis in this article.

II. ANALYSIS OF THE CONCEPT OF DEMOCRACY IN EU LAW BASED ON CATEGORIZATION

2.1. Categorical

This is a question about understanding of democracy, about naming of it, its limits, values and theories. The question: which democracy is being analysed?

The valid Treaties of the European Union (EU) directly mention democracy,[7] representative democracy,[8] participatory democracy,[9] deliberative democracy,[10] democracy in EU external relations,[11] youth democracy[12] (see also Chalmers 2008). It is possible to infer from EU law also direct democracy (referendum,[13] citizens' initiative[14]), constitutional democracy,[15] statutory democracy,[16] EU's internal democracy,[17] input democracy,[18] output democracy,[19] individual(citizen's) democracy,[20] member state democracy,[21] national democracy,[22] democracy at the EU level,[23] parliamentary democracy, liberal democracy,[24] and many other types of democracy (Erne 2011). It is difficult to exhaust the list because, firstly, also other indirect types of democracy can also be inferred from the EU Treaties, and secondly, a measure for claiming the exhaustiveness of a list of the types of democracy is absent. But although it is possible to collect empirically and discuss many other labellings of democracy in EU law, regardless

---

7] Preamble, Arcicles 2, 21 TEU; Preamble CFREU.

8] Article 10 TEU.

9] Articles 15, 165 TFEU.

10] Articles 16, 20 TEU; 284, 330 TFEU; Protocol 1 on the Role of National Parliaments in the EU Legislative Procedure, Protocols 3, 4, 9 TEU, TFEU.

11] Article 21 TEU.

12] Article 165 (2) TFEU.

13] Protocol 16 TEU, TFEU.

14] Articles 11 TEU; 24 TFEU.

15] Some provisions on democracy can be found in the constitutive treaties of the EU, therefore one can talk of constitutional democracy in this context.

16] As far as one can claim that democracy is directly or indirectly present in statutory law.

17] Democracy in the internal affairs of the EU.

18] The term "input democracy" characterizes member state participation in the pre-legislative and legislative activities because by such participation the Member States add value to the EU democratic processes.

19] Protocol 1 on the Role of National Parliaments in the EU Legislative Procedure.

20] There are only two political communities (two entities) represented in EU representative democracy: European citizens, and national governments.

21] Member state-based (national government) democracy in the EU. For further discussion of the concept, see, for example, Fossum 2010.

22] The term refers to how democracy is structured within the EU.

23] As opposite to member state democracy.

24] Based on the liberal theory of democracy.

of occasional correspondence with understandings of democracy in other political units, there are no fixed definitions but ruptures in meaning(s) and disconnections what concerns the concepts beyond the EU. Therefore, gathering the labellings in itself neither defines nor explains democracy.

The list afore of various different appearances of democracy in EU law contains many hybrid concepts. Some hybrid concepts in the EU Treaties contain opposition, e.g., representative democracy/direct democracy. Some hybrid concepts in the EU Treaties contain opposites, e.g., "representative" and "democracy" in "representative democracy", similarly to "Christian" and "democracy" in "Christian democracy" (the decision to crucify based on an the expression of a democratic will). Sometimes such constructions additionally disrupt the meaning of a concept that may already be contested. Such opposites are not illogical because in a hybrid concept only one component is democracy while the other component is an adjective that can have different meanings. The linguistic appearance is similar to the logical exercise in which one has to make sense of "trees": "apple trees", "cherry trees", and so forth, where "tree" is the constant concept. - In the example of democracy, "democracy" is the constant concept, whereas the adjectives vary. When defining a concept, the characteristics of a concept do not disappear but are substituted by a variable, whereas the constant is always broader than the variable. A concept with a variable is more precise, therefore all the hybrid concepts of democracy are deducible to "democracy". Based on Waltz (1979, 55) and his definition of a system as a set of related variables, it is possible to conclude that variables have been added to democracy in the EU Treaties with the aim to specify the content of the concept for the EU.

The concept "democracy" and the concept "representative democracy" are not equal concepts but distinct. The types of democracy (e.g., "representative democracy") cannot fully define democracy although they can say something about the content of democracy. Which is similar to the paradox of the Kantian ontological categories - that a researcher can say something about the logical principles according to which the variables are (re)placed but cannot get beyond that logic. Moreover, this logical circle does not allow a logical deduction that the concept "representative democracy" is a specific concept of "democracy" and therefore must have the same general characteristics as the concept of "democracy". Nor does it allow EU democracy to be a specific concept of democracy that therefore should have the same general characteristics as the concept of "democracy".

A legally situated concept could be compared with other legally situated concepts. But since the EU Treaties do not contain a valid legal definition of democracy, a methodological possibility for a law essay about democracy in EU law would be to simply enumerate the types of democracy in the texts of the valid EU Treaties as in the EU treaties as they have historically progressed. An alternative for a law essay that would wish to remain empirical without getting directly into politics, would be compare the EU principle of democracy with the United Nations General Assembly's functional

definition of the essential elements of democracy that constitutes the normative standard for the EU as a meta-concept referred to in EU law.

I have to confess that it is difficult to find any better alternatives to these two approaches writing a law essay about democracy in EU law, without getting into politics or policy-making. At the same time, in the current situation where all legal definitions of democracy are partial and one final metadefinition of democracy is absent - it is easy to see why there are doubts as to whether an undetermined meta-concept of democracy could constitute a good frame for analysing society, decision making, and governance.

What concerns examining the EU principle of democracy in the light of values then values have been considered weak categories for theorizing.

There are also theories of democracy that could help to create models (categorize). For the reason theories define democracy sufficiently clearly – in a way that allow one to say something about the presence/non-presence of its features in EU law – they are considered stronger than values. The problem with theories is that although they can give different qualitative understandings of democracy, they create limited models because they posit preconditions for democracy. For the reason no EU act directly refers to any theory of democracy, the theories would also constitute a weak basis for a legal comparative research.

2.2. Spatial

The space where democracy takes place for the purposes of this article is EU law. One can approach the space synchronically - comparing the types of democracy in valid EU law, or/and diachronically by trying to trace the concept in its development in the history of EC/EU law.

Under this category, an option could be to discuss the *demos/demoi* of the EU with its multi-level identities because democracy is among others referred to in the Preamble of the Charter of Fundamental Rights of the EU (CFREU), which contains besides the EU citizens' rights also everyone's rights, and that way, the EU constitutional principles are not constrained with the EU level but reach beyond it to the levels of the member states and international. It seems impossible to understand democracy in the CFREU without explaining the context of the CFREU as the evolution of human rights in the EU that consists of long discussions about the relationship of the EU citizens' rights with rights in internal law and in the European Convention of Human Rights (ECHR).

For the reason EU law mentions democracy as an EU guiding principle in accordance with the principles of the Charter of the United Nations (Charter of the UN), it creates a legal link through which EU law and the Chater of the UN become the interacting units and also the principles of the Charter of the UN form part of the EU *acquis* and offer a tool for comparing the understanding of democracy in the EU Treaties with the understanding of democracy in the Charter of the UN. That way, the two democracies (of the UN and the EU) are supposed to coincide, while their

interpretation and application would also require the researcher to investigate the *travaux preparatoires*, the working documents of the Convention on the Future of the EU, the White Paper on the Treaty establishing a Constitution for Europe, commentaries, Explanatory Note to the CFREU, also writings of the publicists and case law would assist here, although there would still remain "grey" areas would remain.

2.3. Temporal

This is a category about time. Temporalization opposes to the ideas of eternal stability and repeatability of categories. As the social phenomena are viewed as processes, concepts change as history preogresses (Marchart 2007) .

Reinhart Koselleck has connected the notion of temporalization with historicization (Marchart 2007, 54). One can situate the EU principle of democracy in time, placing the beginning of the conceptual tree of EU democracy in the idea of Europe before 1951, followed by the appearance of the word "democracy" in the Treaties in 1992 with the Treaty establishing the European Community,[25] the 1994 Agreement on the European Economic Area,[26] the 1997 Treaty of Amsterdam, the 1997 Treaty on EU,[27] the Treaty of Nice of 2001,[28] and so forth. - Step-by-step, democracy has gained space in the EU Treaties. And it went further – the Preamble of the Draft Treaty establishing a Constitution for Europe (Draft TCE) identified and named EU constitutional democracy and proposed a legal definition of democracy for the EU: "democracy means that power is in the hands not of a minority but of the greatest number". The Preamble of the Draft TCE also mentioned equality of persons and freedom, and the wish of the reunited Europe to deepen the democratic and transparent nature of its public life, and to strive for peace, justice and solidarity throughout the world. For the first time in the EU's history, the constitutive Treaties were to have a separate section with the title on democracy as the Draft TCE contained in Articles 44-46 in Title VI "The democratic life of the union" the principle of democratic equality and its legal definition (Draft TCE Art.44), the principle of representative democracy and its legal definition at two levels - as citizens' representative democracy and as member states' representative democracy, while referring to parliamentary democracy, elections, participative democracy, and citizens' representation through political parties (Draft TCE Art.45); and to the principle of participatory democracy with mentioning civil society organizations and the citizens' initiative (Draft TCE Art.46). The text of the

---

25] Article 130u Treaty establishing the European Community (Maastricht Treaty) in the context of development cooperation, saw as the general objective of the EU development and consolidation of democracy and the rule of law.

26] Preamble Agreement on the European Economic Area (EEA Agreement). Democracy is mentioned also in the Declaration by the Governments of the EFTA States on Article 103(1) of the EEA Agreement but with regard to the EFTA states.

27] Preamble, Treaty on European Union (Treaty on EU).

28] Article 181a (1) Treaty of Nice.

Draft TCE that was published in the Official Journal of the EU in the year of 2003 included as its Part II the codification of human rights for the EU that referred to democracy (Preamble). The CFREU has grown out of this Part II. Article III-193 (1) of the Draft TCE in Title V addressing the EU's external action mentioned democracy as an EU guiding principle in accordance with the principles of the Charter of the UN.

The Draft TCE, which was aimed at reconceptualizing EU law did not enter into force. A revised Treaty establishing a Constitution for Europe (TCE) was adopted and published in 2004 but this Treaty did not enter into force either. The TCE contained provisions similar to the Draft TCE provisions, although some provisions were extensively revised. For example, absent was the legal definition of EU democracy, while the Chapter on democracy was maintained but was renamed as "The principle of democratic equality", and contained in Articles I-45-I-47 TCE the principle of democratic equality (directly labelled as "the principle of equality of EU citizens" in Article I-45 TCE); the principle of representative democracy and its legal definition at two levels - as citizens' representative democracy and as member states' representative democracy; also reference to parliamentary democracy, elections, participative democracy, and citizens' representation through political parties (TCE Art.I-46); preserved was the principle of participatory democracy, whereas "civil society associations" was replaced by "representative associations and civil society"; included was the citizens' initiative (TCE Art.I-47). As the TCE of 2004 did not enter into force, there was a "reflection time" during which period was worked out the Treaty of Lisbon that incorporated many re- and new conceptualizations from the Draft TCE and the TCE. The Treaty of Lisbon was adopted in 2007 and entered into force on 1 December 2009. The Treaty of Lisbon stressed in its Preamble the democratic legitimacy of the EU, and made amendments: Democratic equality was not explicitly mentioned in Article 1a TEU, instead one can read: "The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights". The Treaty of Lisbon includes a specific section "On Democratic Principles" materialized in Title II TEU "Provisions On Democratic Principles" in Articles 8 (8a-8c) containing the principle of the equality of the EU citizens, representative democracy consisting of citizens' direct representation at EU level in the European Parliament, and member states' representation in the European Council by their Heads of State or Government and in the Council by their governments (TEU Art.8A), participatory democracy (TEU Art.8A(3)), political party representation, representative associations (TEU Arts.8B(1) and (2)), citizens' initiative (TEU Art.8B(4)), national Parliaments (TEU Art.8C). Amendments to the general provisions of the EU external action generally remained as they had appeared in the TCE what concerns democracy (TEU Art.10A). "[E]ncouraging the participation of young people in democratic life in Europe" was added as in the Draft TCE and TCE. A solidarity clause was added.

Each stage of the EU legal development has been politicized – i.e., the legal meaning of the EU has been made by politicians. Therefore, a conclusion could be that an analysis

of the legal evolvement of the concept of democracy would also require extensive work with the EU policy documents that fix the discussions concerning what to include in or exclude from the EU Treaties, with the *travaux preparatoires*, Convention on the Future of the EU documents, the White Paper on the TCE, commentaries, Explanatory Note to the CFREU, the writings of the publicists and case law, etcetera.

Methodologically, one can still see a circle: the historical is itself the (ever-changing) condition for the presence to emerge. The historical as identifiable facts is empirical but we cannot say much about the content of presence based on this empirics. Not to mention that we approach history subjectively. That way, a time-situated approach can only assist in understanding a concept as a process. But not as a process in evolution that presumes that reality in growing in precision (i.e., growing the specific out of the general, or the future out of the past) because there are political disruptions. So, even if there would exist a single meta-concept of democracy, these disruptions would not allow one to synthesize the EU democracy out of it.

It is the privilege of the legal researchers to write only about the types of democracy expressly present in the valid EU Treaties, and to compare only the black letter appearances of the types of democracy in EC/EU law. Because inclusion of also the derivative types of democracy would already require an explanation, and an external anchoring point for explanation. Because of the absence of legal links, and as the legal choices in the EU depend on the political leadership of the EU, the use of any theory of democracy theory as an anchoring would become weak theorizing for a lawyer, compared with the use of the preparatory legislative texts, case law of the European Court of Justice (CJEU), and even the EU law publicists.

2.4. Procedural

This category asks how to approach the concept of democracy. The choice of a method depends on how the democracy under observation has been limited because democracy can function from an agora to legalized democracy. A research about the legal formulations of democracy in the EC/EU Treaties can become a descriptive enumeration of only the types of democracy or a comparison of them.

As mentioned before, a research could also infer the types of democracy from the EC/EU Treaties. Such an approach would require more explaining and an external anchoring point for explaining.

A research could be a symbolic construction - meaning that a research could describe certain models in the framework of certain decision systems, e.g., in the framework of EU law (as n1), or in the frame of the major theories or understandings of democracy (as n2). Here, problems begin with already the fact that in a legal or/ and political research one cannot state that a general concept of democracy means a meta-concept of democracy, and consequently, that the concept of democracy in the EU Treaties is a derivative specific concept. A danger with using general concepts of

knowledge in political (including legal) modelling is that when one says that B can only build on (or follow) the knowledge about A, one constructs mathematically, but in reality A does not certainly guarantee that B will follow it. That way, (de)construction of a meta-concept cannot in itself explain EU law because a historically evolving meta-concept of democracy cannot explain the politicized EU concept of democracy.

A related but distinct question is whether a contemporary concept necessarily needs to be based on the historical understandings of the same concept. Already St Augustine concluded that it actually needs not because each time has its own understandings and concepts of justice (Saint Augustine 2008). In that sense, one could ask what is the use of conceptual history analysis in politicized areas?

It is interesting that there seems to be a contradiction between the earlier thought of Augustine and a later(!) Marxist understanding that new structures operate based on previous structures and reflect the old structures more and more adequately – Karl Marx understood concepts as successive evolution but his dialectical method applies toward science, not politics (Rosental 1948).

2.5. Identitarian

A general understanding is that a theoretical approach depends on the selected literature theorizer-by-theorizer. A literature review can be used as a frame when discussing whether something is/is not a democracy. The theorizer chooses whose ideas constitute the specific democracy or a suitable theory.

There are many sources available for external approach. Some databases: the Oxford English Dictionary, the Stanford Encyclopedia of Philosophy, ProQuest, Jstor – define in a way that approximates ideologies (Gagnon 2019). This may be called ideology control and implies that refers to politicization. But the labelling as a "Western concept" is not identical with the labelling as a "politicized concept" because the understanding may simply depend on the intellectual background of a particular author. Maybe for that reason, for example, Joseph H. H. Weiler has searched for the meaning of demos and polity from also political theories (Weiler 2011:28). At the same time, the EU treaties do not refer to any political theory or author. And even if they did, the theories are always dynamic and, therefore, the understandings of demos and polity would always be in flux.

At the first sight, a comparison of a Eurocentric meta-concept of democracy with the EU's own concept of democracy would seem as a reliable possibility for finding connecting points for analysis. In practice, such comparison is very complicated because one single Eurocentric meta-concept of democracy is absent. For example, Joseph H. H. Weiler refers to international, supranational and infranational levels of analysis of concerning European democracy (2011:28) and to different general democratic theories, such as the traditional democratic theory, the consociational theory, the competitive elites model of democracy, the federalist

model of pluralist democracy, the neocorporatist model of democracy, the American and British models of democracy, and so on (Weiler 2011:33-44). From the EU perspective, such an approach would require defining EU's own democracy based on EU preparatory acts, policy texts, case law of the ECJ/CJEU, writings by EU law publicists, etcetera.

What concerns the EU's own legal concept of democracy, it was defined by the Convention on the Future of the EU, composed of the representatives of the EU and member states institutions and civil society, and the representatives of the acceding member states, convened with the aim of revising the EU Treaties as the constitutional convention of the EU, and to propose the Treaty amendments as the basis of the 2004 intergovernmental conference (because the procedure of the amendment of the EU Treaties consists of cooperation of the intergovernmental conference and the following European Summit). As a result of their work, the definition of EU democracy was written into the Draft TCE (mentioned above) that did not enter into force. That way, the EU does still not have a valid legal definition of democracy.

The documents of the Convention on the Future of the EU, which have not exactly rewritten the contentious discussions around the policy issues it worked with, allow identification of the people responsible for the amendments in these documents. Therefore, an important source to begin with for understanding the EU's own concept of democracy could be to consult these people, their writings and references. Would it be useful? - Without this link, one can well observe (trace) the development of democracy as written in EU law, and can discuss the general concept of democracy with its many meanings. A comparision of the EU's principle of democracy with any general societal and political concept would in any case be complicated because of strong political influence.

2.6. Substantive

This categorical question is concerned with the content of a specific kind of democracy under research, too, but the difference is in substance not in a categorical form. The question is: What does the democracy in question discuss (focus on)?

When one can neither understand the meaning of democracy nor define it, one cannot participate in this discussion. In that sense, this question is also related with meaning, more specifically, meaning-making, which is always a matter of choice and agreement. Paraphrasing Heidegger: There is no universal thing - only points of agreement (Heidegger 1962). This reminds also of Kant's warning that empirical knowledge is not absolutely true. Therfore, in the case of political concepts important is also the question of Leadership. Kenneth Waltz has said that the frame of understanding is set by the composer of a manual (Waltz 1979).

2.7. Justificatory

"Why?" is the only question among the seven discussed in this article that asks for on explanation of democracy. In this way, the question "Why?" is different from the other questions. Explaining is a foundationalist method.

Sometimes there are limits to explanation. For example, legal concepts are autonomous with regard to general concepts, while politicized concepts require specific methodology.

Claude Lefort explains that the spheres of law, power and knowledge should be kept autonomous because the merging of these spheres would constitute a totalitarian attempt to centre a society around a single ground of legitimation (Lefort 2007, 105).

2.8. The Results Based on Immanuel Kant's Ontological Categorization

In the absence of the correct meta-concept of democracy, and the connecting links between a democracy theory and valid EU law, I began my analysis of EU democracy with a choice of seven categories derived from Immanuel Kant's system of 12 categories and its critique.

The conclusions are that for the EU, the concept of democracy has been developing in the EU *acquis* (the Treaties, secondary acts, case law of the ECJ, the working documents of the Convention on the Future of the EU, etcetera) under the frame of the United Nations General Assembly's definition of democracy as referred in the EU's secondary acts. The recognized EU publicists (for example, Weiler, Fossum, Eriksen, etcetera) focus on EU democracy through specific sectors - for example, civil society, the rule of law, policy making, decision making, democracy deficit, division of powers, interparliamentary cooperation, inter-institutional cooperation, vertical and horizontal separation of powers, electoral regimes, direct elections, political rights, political accountability, exercise of state power, the governmentality *vs.* government debate, EU governance debate, etcetera. The UN acts have the similar functional approach - the UN General Assembly has not defined democracy as such but has defined what it considers as the essential elements of democracy. According to Waltz, similarity in itself in a political structure does not mean uniformity, even if it means similar effects (Waltz 1979, 87, 88). Going deeper into any of these sectors would mean a separate sectoral research, while a general focus on democracy could be lost into details.

The frame derived from the Kantian ontological categories and their critique offers a measure for approaching the principle of democracy in EU law. The Kantian ontological system of categories teaches that for evaluation, one has to go beyond the Kantian categories and ask for the meaning of democracy. The meaning of the EU's principle of democracy should be discussed in the light of the UN General Assembly's definition of the essential elements of democracy, because there is no formally fixed meta-concept of democracy other than the one formally agreed upon in the UN framework to which concept the EU has formally consented. For the reason the Draft TCE did

not enter into force with its legal definition of EU democracy, the legal definition of democracy is only emerging for the EU in the process of its reconstitutionalization. From the viewpoint of validity, the draft and preparatory documents may have only explanatory value in certain circumstances, for example, they might commonly be used to explaining an already validated norm. Although the post-Lisbon *acquis* did grow out of the Draft TCE and the TCE, the latter two documents by virtue of having not entered into force can hardly amount to soft law that a lawyer can use in interpretation of valid law.

CONCLUSIONS

The article discussed the limits of defining the principle of democracy in EU law based on the Immanuel Kant's ontological categorization and its critique.

Although a correct metadefinition of democracy and a correct system for systematizing things into categories are absent, the author of the article does not accept the solution that one cannot actually say anything when trying to explain a non-existant thing using a non-existant method. In Marchart's opinion such could amount to anti-foundationalist nihilism, existentialism or pluralism, all of which assume the absence of any ground and would result in complete meaninglessness, absolute freedom or total autonomy (Marchart 2007). I tried to find some anchoring points from which to say and saying something about this research object, which has been in constant flux. The concept of democracy is one of the most problematic concepts because there is no correct meta-concept of democracy, rather the meanings of democracy have been politically constructed so that different and controversial notions of democracy have emerged in the course of historical and political evolvements.

In addition, the valid EU law does not contain a legal definition of democracy and although EU law refers to the UN General Assembly's understanding of democracy as a connecting point, the UN General Assembly has defined only certain sectors of democracy.

A suitable reference in such circumstances comes from Marchart: "While the sea is boundless and bottomless, it is still structured" (2007:3).

The author of this article, firstly, approached the EU principle of democracy with a constructed system of seven categories and then viewed the results. This approach allowed different classifications diachronicaly and synchronically, while the difficulties arose mainly due to the political nature and indeterminacy of the concept. The questions "Which?", "Where?", "When?", "Who?", "What?" and "How?" were suitable for an empirical (factual) research, at the same time, the approach based on pure progress, induction, deduction and direct conclusions did not in itself allow explanation in the sense of understanding. Only with the question "Why?" the empirical legal research acquires the dimension of evaluative explaining because the question "Why?" can break beyond the logical circling when it escapes the constraints of causality. At the

same time, even though when an inner meaning (pure meaning) of things could be "out there" and one could grasp it regardless of human constraints, politicization is nevertheless disrupting and reconceptualizing the rmeanings of things. The rules of demos and kratos are politically constructed for the EU.

jaanika.erne@ut.ee

REFERENCES

Berlin, Isaiah. 2013. *Concepts and Categories: Philosophical essays.* Princeton University Press.

Birkinshaw, Patrick, and Michael R. Varney (eds.). 2010. *The European Union Legal Order after Lisbon.* Austin [et al.]: Kluwer Law International.

Chalmers, Damian, and Mario Monti. 2008. *European Union Law. Updating Supplement.* Cambridge: Cambridge University Press.

Eriksen, Erik O., and John E. Fossum. 2007. Europe in transformation. How to reconstitute democracy? *RECON Online Working Paper,* no. 1. http://www.reconproject.eu/projectweb/portalproject/RECONWorkingPapers.html (accessed March 11, 2019).

Erne, Jaanika. 2011. Learning Object "Representative Democracy in the EU". Study Material. Tallinn University. Repository of the Estonian Innovation Centre for Digital Education. http://www.tlu.ee/opmat/ri/RIO7008/democracy/21_representative_democracy.html (accessed January 11, 2019).

Fastenrath, Ulrich. 1993. Relative Normativity in International Law. *European Journal of International Law* (4): 305-40.

Fossum, John E. 2010. The Future of the European Order. In *The European Union Legal Order after Lisbon*, edited by P. Birkinshaw and M. Varney, 33-56. Austin [et al.]: Kluwer Law International.

Franck, Thomas M., and John E. Fossum. 2003. Still a union of deep diversity? The convention and the constitution for Europe. *Arena Working Paper* (21).

Gagnon, Jean-Paul. 2018. 2,234 Descriptions of Democracy. *Democratic Theory* 5(1): 92-113.

Heywood, Andrew. 1997. *Politics.* Houndmills: Macmillan.

Jørgensen, Knud E. [et al.] (eds.). 2007. *Handbook of European Union Politics.* London [et al.]: Sage Publications.

Heidegger, Martin. 1962. *Being and Time.* Oxford UK & Cambridge USA: Blackwell.

Kelsen, Hans. 1996. *Introduction to the Problems of Legal Theory.* Oxford: Clarendon Press.

Kennedy, David. 1980. Theses about International Law Discourse. *German Yearbook of International Law* 23(11): 353-91.

Lefort, Claude. 2007. Chapter 4. The Machiavellian Moment Re-Theorized. In *Post-Foundational Political Thought. Political Difference in Nancy, Lefort, Badiou and Laclau*, by Marchart Oliver. Edinburgh University Press.

Majone, Giandomenico. 2002. Delegation of regulatory powers in a mixed Polity. *European Law Journal* (3): 319-39.

Marchart, Oliver. 2007. *Post-Foundational Political Thought. Political Difference in Nancy, Lefort, Badiou and Laclau.* Edinburgh University Press.

Masing, Uku. 2004. *Keelest ja meelest.* Tartu: Ilmamaa.

Rosental, M. 1948. *Marksistlik dialektiline meetod.* Tallin: Poliitiline kirjandus.

Saint Augustine. 2008. *Confessions.* US: Penguin Classics.

Schiappa, Edward. 2003. *Defining Reality: Definitions and the Politics of Meaning.* Southern Illinois University Press.

Thomasson, Amie. 2019. Categories. The Stanford Encyclopedia of Philosophy (Summer 2019 Edn.), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/sum2019/entries/categories/ (accessed February 12, 2021).

Waltz, Kenneth Waltz. 1979. *Theory of International Politics.* Reading [et al.]: Addison-Wesley Publishing Company.

Weiler, Joseph H. H., [et al.]. 2011. *European Democracy and its Critique. Five Uneasy Pieces.* EUI Working Paper RSC, no. 95/11. European University Institute.

Legal Acts

Treaty establishing the European Community. - Official Journal C 224, 31/08/1992, pp. 0001-0079 (Maastricht Treaty).

Agreement on the European Economic Area. - Official Journal L 1, 3.1.1994, pp. 3–522 (EEA Agreement).

Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts. - Official Journal C 340, 10/11/1997, pp. 0001- 0144 (ToA).

Protocol on asylum for nationals of Member States of the European Union. Treaty of Amsterdam amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts. - Official Journal C 340, 10/11/1997, p. 0308.

Treaty on European Union. - Official Journal C 340, 10.11.1997, pp. 145-72.

Treaty establishing the European Community. - Official Journal C 340, 10.11.1997, pp. 173-306.

Treaty of Nice amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts. - Official Journal C 080, 10/03/2001, pp. 0001-0087.

Declaration on the future of the European Union. Treaty of Nice amending the Treaty on European Union, the Treaties establishing the European Communities and certain related acts. - Official Journal C 080, 10/03/2001, pp. 0001-0087.

Treaty on European Union (Consolidated version 2002). - Official Journal C 325, 24.12.2002, pp. 5-32.

Treaty establishing the European Community (Consolidated version 2002). - Official Journal C 325, 24.12.2002, pp. 33-184.

Draft Treaty establishing a Constitution for Europe. - Official Journal C 169, 18/07/2003, pp. 0001-0105 (Draft TCE).

White Paper on the Treaty Establishing a Constitution for Europe.

Treaty establishing a Constitution for Europe. - Official Journal C 310, 16.12.2004, pp. 3-474 (TCE).

Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007. – Official Journal C 306, 17.12.2007, pp. 1-271.

Explanations relating to the Charter of Fundamental Rights. - Official Journal C 303, 14.12.2007, pp. 17-35.

Treaty on European Union. - Official Journal C 326, 26.10.2012, pp. 13-390 (TEU)

Treaty on the Functioning of the European Union. - Officia Journal C 326, 26.10.2012, pp. 47-390 (TFEU).

Protocol 1 on the Role of National Parliaments in the EU Legislative Procedure
Charter of Fundamental Rights of the European Union. - Official Journal C 326, 26.10.2012,
     pp. 391-407 (CFREU).

Political Knowledge:

Measurement, Elitism, and Dogmatism

Kristoffer Ahlstrom-Vij Birkbeck

Featurespace/Visa

Abstract. Political knowledge is a resource: having a lot of it means being in a position to navigate the political world, and stand a better chance of connecting your fundamental political goals with successful means. The present piece argues that standard political knowledge tests measure political knowledge, so understood, and uses counterfactual modeling to demonstrate the difference having such knowledge can make to political choice. It then takes up two of the most forceful objections to political knowledge and its measurement, by (a) rejecting the idea that knowledge scales encode elitist assumptions, and (b) arguing that, even if political knowledge breeds dogmatism, such dogmatism can be expected to serve a protective function in exactly the type of hostile epistemic environments – filled with lies, falsehoods, and misinformation – that make up the political domain.

Keywords: political knowledge; public ignorance; information effects; knowledge scales.

1. THE TRADITIONALIST AND THE REVISIONIST

It is well-established that most of us know very little when it comes to politically relevant facts (e.g., Delli Carpini and Keeter 1996; Achen and Bartels 2016; Somin 2016). But what follows from this? There are two schools of thought: the traditionalist and the revisionist.1 The traditionalist takes the straightforward line that public ignorance poses problems for democracy. The two main problems are as follows:

- **The problem with ignorant voting:** Voters might be rationally ignorant (e.g., Downs 1957). But what is individually rational can still be trumped by moral obligations in cases of aggregate harm. (My attempting to free-ride on recycling might be rational, but I still ought to recycle.) Something similar goes for ignorant voting: it is a 'collectively harmful activity' (Brennan 2016; see also Somin 2016).[1]
- **The problem with ignorant non-voting:** Political knowledge is unevenly distributed, with higher levels found among people who are white, educated, older, and male. Since those with high levels of political knowledge are also more likely to vote, political igno- rance means that those already disadvantaged remain so, and that injustices will likely be perpetuated rather than eradicated through the democratic process (e.g., Delli Carpini and Keeter 1996, 138).

The revisionist, by contrast, typically argues that, voters use varieties of heuristics and rules of thumb (e.g., party-political cues) to make up for what they do not know

---

1] I am here setting aside a large literature responding to concerns about public ignorance by suggesting that a large public is likely to outperform a smaller number of more competent individuals, or arguing that there are epistemic benefits of deliberative or liberal institutions. See (Ahlstrom-Vij 2020b) for an argument that these considerations fail to cast doubt upon the traditionalist line covered here.

(e.g., Popkin 1991). For example, if I don't know much about economic policy, but I know that party X is 'my' party, I can just vote whatever their position is on the matter.

Importantly, the revisionist narrative is testable, since it entails a counterfactual: *By relying on heuristics, we vote and hold political preferences in the same way we would have, had we been fully informed.* Note that 'fully informed' is a term of art here: it does not designate political omniscience, but is usually operationalised as getting a perfect score on a civic style political knowledge test (more on these in Section 2 below). Note also that this counterfactual provides a normative standard for thinking about voter competency. In particular, consider Arthur Lupia's – a prominent revisionist – who suggests that

> [...] we should evaluate a voter as competent regardless of how she reaches a conclusion, as long as it is the conclusion she would have reached had she been aware of the best available information (Lupia 2006, 226).

With this in mind, we can capture what constitutes a sufficient amount of political knowledge – sufficient for competence, as Lupia would say – as follows:

> **CHANGES:** You have a sufficient amount of political knowledge for purposes of political preference and choice if it is the case that, had you known more, this would not have changed your political preferences and choices.

Differently put, you know enough if your preferences and choices would have survived your knowing more. With that in mind, we can test the relevant counterfactual as follows, to see if the revisionist's story holds up. We first gather data on people's level of political knowledge as given by some appropriate political knowledge test, alongside their political preferences and demographics – or, alternatively, we rely on established surveys of political opinion, such as the American National Election Studies (ANES) in the U.S. and the British Election Study (BES) in the U.K. Then, we use regression analysis to identify the relationship between test scores, demographic variables, and the probability of reporting particular political preferences. Finally, we use that model to increase the knowledge variable of each to the maximum, and note for each respondent how that probability changes. Assuming that the relevant knowledge tests do in fact measure political knowledge (more on this in Section 2), the presence of any information effects – i.e., differences between our actual preferences and the preferences we would have held, had we been fully informed – is evidence that we do not know enough.

Such information effects also speak directly to the revisionist counterfactual. Henrik Os- carsson explains:

> If the popular [revisionist] view of low information rationality [through the use of heuristics] is correct – that most uninformed voters most of the time make the same voting choices as they would have had they been fully informed – we would not expect any significant information effects, and certainly not any important changes in the aggregate outcome of elections. More knowledgeable voters would not differ from less knowledgeable in political preference or behavior since the use of heuristics would be a successful compensatory strategy" (Oscarsson 2007, 304).

As it happens, Information effects have been identified in a number of geographical contexts over the past couple of decades (see Ahlstrom-Vij (2020a) for a discussion). For example, Scott Althaus (2003) finds that fully informed preferences tend to be 'more dovish and interventionist on foreign policy, less conservative on social, environmental, and equal rights issues, and more conservative on morality issues and questions about the proper limits of government activity' (2-4). Across all survey questions considered by Althaus, the average difference on the collec- tive level between surveyed and modelled (reported) preferences is 6.5 percentage points, and '[s]hifts in collective preferences [were] large enough to change what appears to be majority or plurality consensus on an issue occur quite frequently' (126), and 'correcting for information effects changes collective preferences in nearly half of governance questions' (128).

In terms of vote choice in particular, Larry Bartels (1996) finds that, on average, 'Democrats do almost two percentage points better and incumbents do almost five percentage points better than they would if all voters in presidential elections were, in fact, fully informed' (220). André Blais et al. (2009) simulate the outcome of six past Canadian elections, and see a likely difference in outcome in one, and an average information effect of 2.3 percentage points across parties and elections. Henrik Oscarsson (2007) simulates six past Swedish elections, with a likely electoral difference in outcome in two under a fully informed electorate, and an average net gain of 2.7 percentage points for right parties. Yosef Bhatti (2010) models three European Parliament elections (in Denmark, Finland, and Sweden) with an average information effect of 3.5 percentage points across parties and elections, and two instances where the difference between actual and simulated support exceeds ten percentage points (with a social democratic party losing out, and a liberal one gaining).

Return to the revisionist counterfactual with these results in mind. In each case, these infor- mation effects suggest that we likely would have voted differently and held different preferences, had we been fully informed (but otherwise been identical across measured covariates) – contrary to the revisionist narrative.

2. WHAT POLITICAL KNOWLEDGE IS AND HOW TO MEASURE IT

So, the presence of information effects is bad news for the revisionist. However, any estimate of informed political preferences will only be as good as the conception and measure of political knowledge that it relies on. What, then, is political knowledge? Delli Carpini and Keeter offer a helpful gloss on its domain:

> [...] a general familiarity with (1) the rules of the game (the institutions and pro- cesses of elections and governance); (2) the substance of politics (the major domestic and international issues of the day, current social and economic conditions, key polit- ical initiates, and so forth); and (3) people and parties (the promises, performances, and attributes of candidates, public officials, and the political parties) is critical to the maintenance of a healthy democracy (Delli Carpini and Keeter 1996, 14).

This makes clear that political knowledge is a resource: having a lot of it means being in a position to navigate the political world, and to stand a better chance of connecting your fundamental political goals with successful means – e.g., by voting for the candidate or party that is best place to realise those goals.[2]

How, then, do we measure political knowledge? The standard way is to ask a number of factual questions about what government is and does – along the lines of Delli Carpini and Keeter's three-part gloss above – and then add up the number of correct answers for a 'knowledge score.' A couple of points are in order:

- Why think these tests measure anything? This is fairly easy to establish on formal grounds, since we in any given instance can test for internal consistency (the extent to which items co-vary) and unidimensionality (whether the items likely measure a single latent trait).
- Why think they measure political knowledge specifically? If political knowledge is a resource, these scales should find more of it among the privileged and less among the disadvantaged, which turns out to be the case (e.g., Althaus 2003, 135; see also Delli Carpini and Keeter 1996, 177).
- What topics do the items need to cover? Since people tend to be generalists – i.e., those who know a lot in one political area tend to know a lot in other areas – we 'need not be overly concerned with the mix of specific topics covered by individual items' (Delli Carpini and Keeter 1996, 174).

- Do the scales need to be long? Tests with as few as five items tend to perform as well as longer ones (Delli Carpini and Keeter 1996 and 1993).

To get a sense of how this works in practice, as it relates to the type of information effects considered in the previous section, consider the six, true/false knowledge items in Table 1, from the March 2019 (wave 15, N = 30,842) panel survey of the British Election Study.

Item
1. Polling stations close at 10.00pm on election day.
2. No-one may stand for parliament unless they pay a deposit.
3. Only taxpayers are allowed to vote in a general election.
4. The Liberal Democrats favour a system of proportional representation.
5. MPs from different parties are on parliamentary committees.
6. The number of members of parliament is about 100.

Table 1: The six general knowledge items from Fieldhouse et al. 2020.

2] See also Zaller 1992 on how the politically informed tend to be the most ideologically coherent, a point that we will nd reason to return to below.

Looking at the response patterns in this particular data set, four of the six items (2, 4, 5, and 6) make for a scale that both has a high internal consistency and is likely unidimensional, i.e., measuring a single, latent feature, as given by a factor analysis.[3]

Moreover, looking at the proportion of respondents answering all four question correctly, we see the patterns we should expect to see, if political knowledge is a resource: in particular, Figure 1 shows that those with higher education, higher income, of older age, and who report being more likely to vote are over-represented among the politicall

Figure 1: Proportion of respondents in Fieldhouse et al. (2020) with a maximum knowledge score by education, income, age, and self-reported likelihood to vote in the general election.

So, it looks like we are measuring something, and that that something moreover is plausibly thought of as political knowledge. But the revisionist will want to know whether what we are measuring makes a difference to political preferences and choice: some might have more knowledge than others, but perhaps those with less do just fine.

To evaluate this, remember CHANGES: if knowing more would have changed my pref- erences/choices, I do not (presently) know enough. As we have seen, such (counterfactual) differences are measured by information effects. For illustration, Figure 2 shows what happens if we go through the three steps in Section 2 with regards to respondents' vote in the UK's 2016 EU referendum, from the above BES data set.[4]

Given the fairly substantial information effect (4.3 percentage points), when compared to the size of typical information effects (see Section 1), it looks like a

---

3] The standardized (Chronbach's) alpha for the scale is 0.772. The factor analysis was performed using fa.parallel in R's psych package (Revelle 2018).

4] For full information about the modeling involved, see the Appendix in Section 7.

sizeable proportion in the sample did not have enough political knowledge – and that the knowledge measured does make a difference.[5]
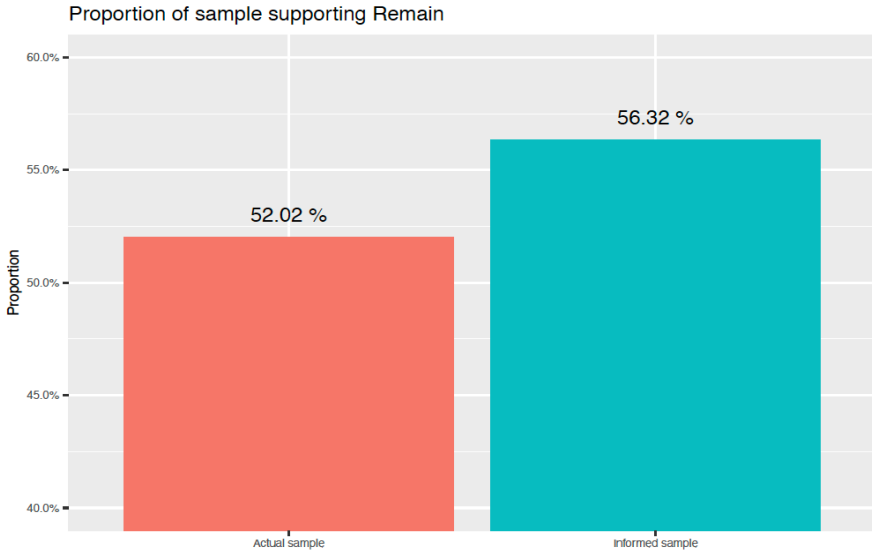
Figure 2: Proportion of respondents in Fieldhouse et al. (2020) supporting 'Remain' as opposed to 'Leave' in the 2016 EU referendum in the actual sample (left) and as estimated by a logistic model for a fully informed sample (right), holding constant gender, age, income, ethnicity, and education.

3. ARE KNOWLEDGE SCALES ELITIST?

We are now in a position to evaluate one of the most forceful, revisionist lines on knowledge scales: Lupia's charge that these scales are driven by elitism. He explains:

> Most political-knowledge questions are not derived from a replicable or transparent logic about how their answers bear on a voter's ability to make decisions of a particular quality in the voting booth. Instead, the questions are generated by a worldview that is shared by a select set of academics, journalists, and politicos, but few others. [...] The elitist move is to assume that these questions have a similar value to citizens whose societal responsibilities can be very different than their own (Lupia 2006, 219).

5] Someone might object that the knowledge tests involved simply measure a form of technical political knowl-edge, as opposed to what we might call situated political knowledge, such as knowing what it is like to go hungry for days on end, or to be on the receiving end of racism. But note that, by controlling for standard demographic variables (e.g., income, gender, ethnicity, etc.) with which these forms of knowledge likely travel, the type of model relied on here preserves such situated knowledge in estimating informed preferences. Specifically, it can be thought of as modeling what you would have preferred, had you had all of the situated knowledge you currently have (e.g., on account of a direct experience with hardship and injustice), and also been informed in the sense of having technical political knowledge (e.g., by being highly informed on how the political world operates, and how to navigate it).

I will suggest that Lupia goes wrong on two points. First, he confuses testing with educating for political knowledge. Second, he disregards evidence that what is tested for is not mere political trivia. Let us consider these points in turn.

3.1 Confusing testing with education

Lupia asks us to imagine that we have identified some set of facts, A-Z, and we want to determine whether knowing these facts is necessary for competent voting (task t).

> To answer [this] question about necessity, we must ask whether there is a differ-
> ent set of facts, perhaps even a subset of facts A-Z, that also allows the voter to
> accomplish t? If the answer is no, then knowledge of every fact from A to Z is
> necessary for the voter to choose competently. In such a case, we can assess the
> voter's political competence reliably by quizzing her about A through Z – as with
> the political-knowledge tests on which so many analysts base their judgments of
> voter incompetence. If we find her deficient in her knowledge of even one of these
> facts, we can accurately judge her incompetent at task t. If we want to increase
> her competence at this task, moreover, we know that a precondition of success is
> providing her with the knowledge of all such facts about which the quiz reveals her
> ignorance (Lupia 2006, 222).

Lupia is right in this: if there is some set of facts necessary for performing a task, someone can be deemed incompetent at that task if they turn out not to know some of those facts. And we can make her more competent by imparting those facts in her. Where he goes wrong is in assuming that this is how political knowledge tests work. He continues:

> The problem with this approach to assessing voter competence is that it is validity
> depends on establishing that facts A-Z are necessary for competence: i.e., that
> knowledge of no subset of these facts, or alternative set of facts, would suffice for
> task t to be accomplished. However, if facts other than the full set A-Z are sufficient
> for citizens to accomplish t, then knowing A-Z cannot be a necessary condition
> for competence at t. Thus, merely demonstrating that a voter does not know these
> facts may reveal little or nothing about her competence in the voting booth (Lupia
> 2006, 222).

Lupia's point here makes sense if we are looking to educate for political knowledge: if there are things everyone needs to know (necessary), these must be identified and imparted. If we succeed, the person will have become more competent; and if we fail, she will not. But knowledge tests are not educational tools. They are – naturally – constructed to test people. This is done in terms of what is diagnostic of political knowledge, and what is diagnostic might be neither necessary nor sufficient for such knowledge. Table 2 illustrates this point, using an analogy with medical diagnostics.

Table 2: Medical diagnostics and political knowledge cases compared.

So, there are things that might be diagnostic of some target property, even if neither neces- sary nor sufficient for having the relevant target property. Knowing that (say) there are more than 100 MPs in the Commons, or that the Liberal Democrats favour a system of proportional representation, is neither necessary nor sufficient for being a competent UK voter. Still, these are things that, if you keep up with politics and are politically informed, you will likely know these things; and if you do not, then you likely won't.

3.2 Testing for non-trivial knowledge

Lupia might grant the above point about testing versus educating, but respond that what typical knowledge items are diagnostic of is not political knowledge, but political trivia. As he writes: 'what benefit does a randomly selected citizen draw from knowing something like the name of the Chief Justice [of the Supreme Court]?' (218).

Part of the answer to this question was provided in the previous section: an item on a knowledge scale does not earn its keep for being necessary for political competency, but for being diagnostic of such competency. Moreover, two further things should be pointed out here As already noted, short knowledge scales of the kind discussed earlier correlate well with longer scales (e.g., Delli Carpini and Keeter 1996 and 1993). Beyond that, you would be hard pressed to maintain that the collection of items usually found on such longer scales are mere trivia, as opposed to substantive facts relevant to sensible political action. By way of illustrating this point, see Table 3 for the twenty-item scale from Delli Carpini and Keeter 1996, which is representative of subsequent work on political knowledge (e.g., in Althaus 2003; see also Bennett 2003).

People	Party	Civics
1. Quale	9. Relative ideological	15. Times a president
2. Gorbachev	location of the two	can be elected
3. Thatcher	parties.	16. Whose
4. Name one	10. Party with most	responsibility is
candidate (and his/her	seats in the House	judicial review
party) for US House	11. Relative location	17. Whose
5. Mandela	of parties on defense	responsibility is it to
6. Foley	spending	nominate federal
7. Rehnquist	12. Party with most	judges
8. Mitchell	seats in the Senate	18. What are the first
	13. Relative location	ten amendments called
	of parties on federal	19. What majority is
	spending	needed to override
	14. Relative location	presidential veto
	of parties on aid to	20. How long is a
	blacks	senator's term

Table 3: The twenty items from Delli Carpini and Keeter's 1996 (302) long knowledge scale, drawn from the 1990-91 American National Election Studies survey.

4. DOES POLITICAL KNOWLEDGE BREED DOGMATISM?

A more fundamental objection to the utility of political knowledge and its measurement takes the line that such scales measure political knowledge, but that any benefits of political knowledge will be canceled out by such knowledge also making us dogmatic.

This line has been pursued recently by Michael Hannon (2021). Hannon starts by making the observation that the politically knowledgeable tend to be more partisan, in the sense of having the most stable and worked-out ideological positions (e.g., Zaller 1992). That, in itself, is neither surprising nor necessarily worrying. What is worrying, according to Hannon, is that the politically knowledge – and thereby most partisan – also are more susceptible to motivated reasoning than the rest of us. Motivated reasoning involves pursuing and processing information in ways that systematically favor your preferred viewpoints, and often in ways that serve to 'protect' beliefs and convictions that are central to your group identity. For example, when it comes to politically motivated reasoning in particular, it involves 'the formation of beliefs that maintain a person's status in [an] affinity group united by shared values' (Kahan 2016, 2).

Why think that the politically knowledgeable are more susceptible to politcally motivated reasoning than the rest of us? Hannon's primary piece of evidence involve a study by Taber and Lodge (2006). Taber and Lodge measure political knowledge using a "17-item general political knowledge scale" (760), and contend that "the politically knowledgeable, because they possess greater ammunition with which to counterargue incongruent facts, figures, and arguments, will be more susceptible

to motivated bias than will unsophisticates" (757). Taber and Lodge clearly have a particular causal story in mind: political knowledge is the mechanism that gives rise to increased levels of motivated reasoning, on account of the politically knowledgeable having "greater ammunition" with which to shoot down any challenges coming from outside their partisan fortress. Hannon is more careful on this point, suggesting instead that "partisanship is the common cause of both acquiring political knowledge and motivated reasoning" (8). This matters for whether we can expect making people more politically informed will mean also making them more partisan and/or dogmatic, in systematically rejecting viewpoitns coming from "the other side," so to speak.

For purposes of argument, let us assume Taber and Lodge's causal story, whereby the politically knowledgeble are dogmatic because they are knowledgeable. As far as knowledge scales are concerned, this raises questions, not about whether such scales are diagnostic, but whether they are measuring anything that we should have any interest in promoting. After all, while it might sound desirable to raise the level of political knowledge in society, it is not clear that we should want to do that if it will invariably lead to increased levels of dogmatism.

But this, in turn, raises the more fundamental question whether dogmatism is invariably a bad thing, from an epistemic point of view. Dogmatism – roughly, the unwillingness to engage with the perspectives and viewpoints of those with whom one disagrees – is a belief-forming disposition, and the epistemic status of such dispositions are a function of the environments in which they are employed. For example, just like blindly deferring to others might be a virtue if one is surrounded by knowledgeable sources (Ahlstrom-Vij 2015), dogmatically refusing to engage with others might be a virtue if one is surrounded by ignoramuses (Battaly 2018).

In fact, we can investigate in a more systematic fashion the implications of being highly knowledgeable, and as such (we are assuming) also highly dogmatic, versus being less knowl- edgeable using a simulation study. Specifically, let us set things up as follows:

• Assume some set of agents – say, 500 – and that each agent has a partisan affiliation, called "green" and "blue." These are assigned randomly, with a probability of 50%.
• As part of the simulation, each agent encounters 100 pieces of information in sequence. Some of these are true, and some of them are false. We code true ones as '1' and false ones as '-1,' and then quantify what an agent has learned at the end of the process by taking the sum of what they've taken on board. So, if you encounter 100 pieces of false information, you will come out at -100, designating having become maximally misled, relative to the body of information. If you encounter 100 true ones, you'll come out at 100, designating that you've come out maximally informed, relative to the body of information.

• Each piece of information is also "tagged" with a partisan cue. We can think of this as the source of the information being a representative of either "green" or "blue."

- Assume also that each agent either has a high level of knowledge knowledge or a low level of knowledge. These labels are assigned to agents randomly, with a probability of 50%. Crucially, the only difference between these two types of agents is that low knowledge agents take on board all information, while high knowledge agents only take on board those "tagged" with their own partisan label.

In other words, what we are simulating is two types of agents being faced with 100 pieces of information in sequence, where each piece of information is "tagged" with a partisan affiliation, and the politically knowledgeable agents using their increased sophistication to "shoot down" and as such reject all information that is affiliated with the other party.

Under these assumptions, what will be the effect of that dogmatism on where high and low knowledge agents will find themselves, in regards to their sum of informedness, as a result of having encountered the 100 pieces of information? This will partly be a function of the base rate of true and false information in the body of evidence as a whole (i.e., of the base rate). So let us run our simulation for 101 such base rates, ranging from 0% true (i.e., 100% false) up to 100% true. Figure 3 shows what happens.
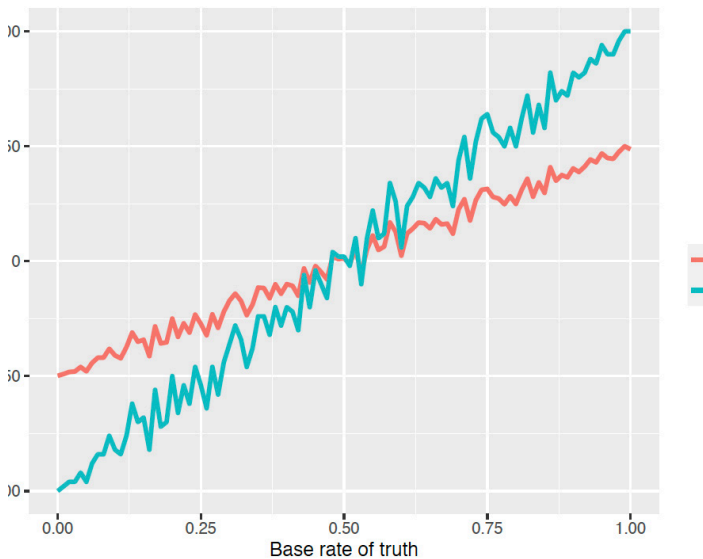
Figure 3: Change in level of informedness as a function of the base rate of truth in the environment for high and low knowleddge agents, after having encountered 100 pieces of information.

Along the x-axis, we have the base rate of truth. On the y-axis we have the mean change in the level of informedness that the respective group of agents – high and low knowledge – has at the end of encountering all of the information, given a base rate. In other words, What we are modelling is the diachronic fact of agents encountering information, and then seeing where they end up. And where do they end up? Look at low knowledge agents first. Here the line pretty much follows the diagonal. That makes

sense: since low knowledge agents take on board all information they encounter, the change in their level of informedness at the end of the process should track the base rate.

But now look at the high knowledge agents. They do better in "hostile" epistemic envi- ronments, for the simple reason that they reject some information in a situation where lots of information is false. And they do worse in "friendly" environments for the same reason (i.e., they miss out on true information). So the mere fact of dogmatism is epistemically beneficial in hostile epistemic environments, even if the dogmatism itself is epistemically neutral. After all, agents are not rejecting messages because of their epistemic merit, but simply because they are coming from the "wrong" party. Specifically, high knowledge agents lose less knowledge than do low knowledge agents in hostile epistemic environments, and gain less in friendly ones.

This, of course, is compatible with high knowledge agents still knowing more across both types of environments after that loss/gain, simply on account of knowing more to begin with. What their dogmatism does in hostile enviromnet is to protect their relative level of epistemic superiority vis-a-vis low knowledge agents. Heather Battaly makes this same observation when reflecting on why dogmatism can in some cases qualify as an epistemic virtue:

> So, what is a knowledge-possessing agent to do when she finds herself in an epistem- ically hostile environment? My proposal is that there are epistemic reasons for her to be closed-minded – to be unwilling to engage seriously with relevant intellectual options that conflict with what she already knows. [...] Why should she be closed- minded? Because, in an epistemically hostile environment, closed-mindedness is an effects-virtue. When a knowledge-possessing agent is stuck in an epistemically hos- tile environment, surrounded by falsehoods, incompetent sources, and diversions, closed-mindedness about options that conflict with what she knows will minimize the production of bad epistemic effects for her (Battaly 2018, 39).

This offers an intriguing response to Hannon's challenge to the benefits of political knowl- edge. Because even if the politically knowledgeable are maximally dogmatic, as in the above simulation, and even if it is their knowledge that makes them dogmatic (as per Taber and Lodge's suggestion), this does not cast doubt on the epistemic merits of being politically knowl- edgeable. On the contrary, such dogmatism serves to protect the knowledgeable in hostile environments – and most political epistemic environments are, after all, highly likely to be ex- actly that, on account of being marked by a high proportion of falsehoods, lies, and misinforma- tion. Consequently, being politically knowledgeable remains preferable to not being politically knowledgeable even under those assumptions.[6]

6] Thinking back to the problem about ignorant non-voting perpetuating unjust patterns in society, it might be objected that the protection offered the politically knowledgeable by their dogmatism will be to the detriment of the already marginalized. (Many thanks to Fabienne Peter for making this point.) However, note two things. First, there is some evidence to suggest that the more knowledgeable tend to be more progressive on political matters (e.g., Althaus 2003), suggesting that the knowledgeable generally make political decisions that would benefit not just them, but also the marginalized. Second, even if that

5. CONCLUSION

Political knowledge can be thought of as a resource: having a lot of it means being in a position to navigate the political world, and stand a better chance of connecting your fundamental political goals with successful means. The present piece argued that standard political knowledge tests measure political knowledge, so understood, and used counterfactual modeling to demonstrate the difference having such knowledge can make to political choice. It then took up to two of the most forceful objections to political knowledge and its measurement, by rejecting the idea that knowledge scales encode elitist assumptions, and arguing that, even if political knowledge breeds dogmatism, such dogmatism can be expected to serve a protective function in exactly the type of hostile epistemic environments – filled with lies, falsehoods, and misinformation – that make up the political domain.

ahlstromvij@gmail.com

REFERENCES

Achen, Christopher, and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton: Princeton University Press.

Ahlstrom-Vij, Kristoffer. 2015. The Social Virtue of Blind Deference. *Philosophy and Phenomenological Research* 91(3): 545-82. https://doi.org/https://doi.org/10.1111/phpr.12158.

———. 2020a. The Case for Modelled Democracy. *Episteme*: 1-22. https://doi.org/10.1017/epi.2020.10.

———. 2020b. The Epistemic Benefits of Democracy: A Critical Assessment. In T*he Rout- ledge Handbook of Social Epistemology*, edited by Fricker Miranda, Graham Peter, Henderson David, and Nikolaj Pedersen. Routledge.

Althaus, Scott L. 2003. *Collective Preferences in Democratic Politics: Opinion Surveys and the Will of the People.* Cambridge University Press. https://doi.org/10.1017/CBO9780511610042.

Bartels, Larry M. 1996. Uninformed Votes: Information Effects in Presidential Elections. *American Journal of Political Science* 40(1): 194-230. http://www.jstor.org/stable/2111700.

Battaly, Heather. 2018. Can Closed-Mindedness Be an Intellectual Virtue?. *Royal Institute of Philosophy Supplement* 84: 23-45. https://doi.org/10.1017/S135824611800053X.

Bennett, Stephen Earl. 2003. Is the Public's Ignorance of Politics Trivial?. *Critical Review* 5(3-4): 307-37. https://doi.org/10.1080/08913810308443585.

Bhatti, Yosef. 2010. What Would Happen If We Were Better Informed? Simulating In- creased Knowledge in European Parliament (EP) Elections. *Representation* 46(4): 391-410. https://doi.org/10.1080/00344893.2010.518078.

Blais, Andre, Elisabeth Gidengil, Patrick Fournier, and Neil Nevitte. 2009. Information, Visibility and Elections: Why Electoral Outcomes Differ When Voters Are Better

were not true, the appropriate response would be, not to want to see a reduction in knowledge among those who have it, but an increase among those who don't.

Informed. *European Journal of Political Research* 48(2): 256-80. https://doi.org/https://doi.org/ 10.1111/j.1475-6765.2008.00835.x.

Brennan, Jason. 2016. *Against Democracy.* Princeton: Princeton University Press.

Delli Carpini, Michael X., and Scott Keeter. 1993. Measuring Political Knowledge: Putting First Things First. *American Journal of Political Science* 37(4): 1179-1206. http://www.jstor.org/stable/2111549.

———. 1996. *What Americans Know about Politics and Why It Matters.* New Haven: Yale University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper.

Fox, John, and Sanford Weisberg. 2011. *An R Companion to Applied Regression.* Second edition. Thousand Oaks CA: Sage. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Hannon, M. (2022). Are knowledgeable voters better voters? *Politics, Philosophy & Economics* 21(1), 29-54. https://doi.org/10.1177/1470594X211065080

Kahan, Dan M. 2016. The Politically Motivated Reasoning Paradigm, Part 1: What Politically Motivated Reasoning Is and How to Measure It. In *Emerging Trends in the Social and Behavioral Sciences*, 1-16. American Cancer Society. https://doi.org/https://doi.org/ 10.1002/9781118900772.etrds0417.

Lupia, Arthur. 2006. How Elitism Undermines the Study of Voter Competence. *Critical Review* 18(1-3): 217-32. https://doi.org/10.1080/08913810608443658.

Morgan, Stephen, and Christoper Winship. 2015. *Counterfactuals and Causal Inference : Methods and Principles for Social Research.* Cambridge University Press.

Oscarsson, Henrik. 2007. A Matter of Fact? Knowledge Effects on the Vote in Swedish General Elections, 1985-2002. *Scandinavian Political Studies* 30(3): 301-22. https://doi.org/ https://doi.org/10.1111/j.1467-9477.2007.00182.x.

Popkin, Samuel. 1991. T*he Reasoning Voter : Communication and Persuasion in Presidential Campaigns.* Chicago: University of Chicago Press.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Revelle, William. 2018. *Psych: Procedures for Psychological, Psychometric, and Personality Research.* Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/ package=psych.

Somin, Ilya. 2016. *Democracy and Political Ignorance: Why Smaller Government Is Smarter.* Stanford, California: Stanford Law Books, an imprint of Stanford University Press.

Taber, Charles S., and Milton Lodge. 2006. Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50(3): 755-69. https://doi.org/https://doi.org/10.1111/j.1540-5907.2006.00214.x.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* Cambridge England New York, NY, USA: Cambridge University Press.

7.1 Models and diagnostics

The modelling performed in Section 2 used 'doubly robust' estimation for counterfactual infer- ence (e.g., Morgan and Winship (2015)). The 'double robustness' owes to how counterfactual effects – in our case, what position someone would take, had they been fully informed – are es- timated in a context where we have both controlled for (assumed) confounds, and improved the balance between the two groups (here: those with a maximum score on the knowledge test and everyone else) to make up for the fact that they have not come about as a result of randomized assignment. In the present case, this second layer of 'robustness' was achieved through so-called 'propensity scores.'

In our case, propensity scores measure the probability (or propensity) that an observation will be found in the 'informed' category (i.e., attaining a maximum score on the knowledge test), as a function of someone's demographic features. The idea is to then use these scores to remove any correlation between these features and the 'informed' category, to justify a counterfactual inference. Why? It is helpful to think of this on the model of a randomized design, where the random allocation of participants to a treatment and a control group means that no feature of the participant is predictive of being found in the treatment as opposed to in the control. Whether female or male, rich or poor (etc.), you are equally likely to end up in one group as opposed to in the other, provided assignment is truly random. In the case of observational data, by contrast, this might not be the case. In the case at hand, it will (for example) be the case that that some features of the observations – e.g., their level of education, their income, or what have you – are predictive of ending up in the 'informed' category.

Specifically, we can use the inverse of those scores as weights (such that an observation with a low propensity is weighted heavily, and vice versa) in fitting the model. This improves the balance between the two groups, and – given an appropriately chosen set of covariates when calculating the scores – recreates a situation that would have been expected in a randomized experiment, thereby allowing greater confidence in inferring a counterfactual.

In the present case, the propensity scores were calculated by way of logistic regression, using glm in R's stats package (R Core Team (2018)). By way of illustration, consider Figure 4, where the left-hand panel shows the balance (or rather: lack thereof) for the income variable prior to applying the propensity weights – it can be seen that informed participants (teal bars, designated here as 1 or 'treated') are over-represented among the wealthy (higher quintiles), and underrepresented among the less wealthy (lower quintiles) – and the right-hand panel the balance achieved once the weights had been applied.
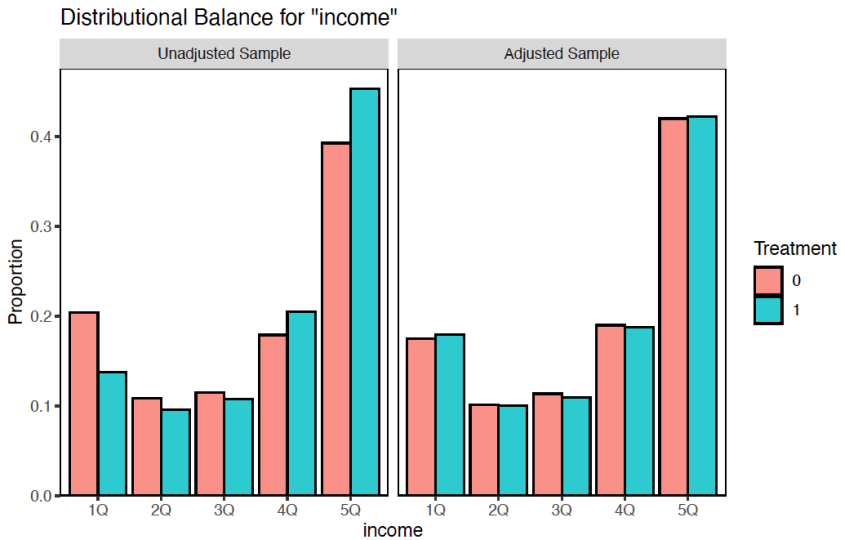
Figure 4: Balance plots for income quintiles, before and after applying propensity weights/

Using the propensity scores as weights, a logistic regression model was then fitted, again using glm. Table 4 contains details on the coefficients and coefficient values of the logistic model used, designated 'Model 1' in that table. The highest variance inflation factor value for that model was 1.187266, as given by vif in the car package (Fox and Weisberg 2011), providing no evidence of multicollinearity. There also was no indication of extreme values as measured by Cook's distance, but 431 observations had a standardized residual greater than 3. In light of these potentially influential observations, the model was refitted without those observations.

This is the model designated 'Model 2' in Table 4. If we simulate the informed outcome of the 2016 referendum on that model, we get an (increased) information effect of 6.54 percentage points (58.56% for Remain).

	Model 1	Model 2
Intercept	0.80 (0.08)***	1.11 (0.09)***
Knowledge	0.43 (0.02)***	0.63 (0.02)***
Gender (Male)	−0.23 (0.02)***	−0.43 (0.02)***
Income (2Q)	0.10 (0.04)*	0.20 (0.05)***
Income (3Q)	0.14 (0.04)***	0.22 (0.05)***
Income (4Q)	0.13 (0.04)***	0.25 (0.04)***
Income (5Q)	0.37 (0.03)***	0.50 (0.04)***
Ethnicity (White)	−0.25 (0.06)***	−0.35 (0.07)***
Education (No education)	−0.55 (0.05)***	−0.69 (0.06)***
Education (Below GCSE)	−0.37 (0.06)***	−0.46 (0.07)***
Education (GCSE)	−0.47 (0.03)***	−0.51 (0.04)***
Education (Undergraduate)	0.44 (0.03)***	0.45 (0.03)***
Education (Postgraduate)	0.96 (0.04)***	1.34 (0.05)***
Age (30-45)	−0.51 (0.05)***	−0.61 (0.06)***
Age (45-65)	−1.10 (0.05)***	−1.43 (0.05)***
Age (65 and over)	−1.30 (0.05)***	−1.77 (0.05)***
AIC	44319.24	38027.94
BIC	44444.52	38152.85
Log Likelihood	−22143.62	−18997.97
Deviance	46431.36	40177.62
Num. obs.	18591	18160

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table 4: Logistic models

7.2. Simulation

The simulation in Section Section 4 relied on the following R script:

```
set.seed(100)
n <- 500
info_pieces <- 100

# generate results dataframe
results1 <- data.frame(base_rate = (0:100)/100, high = NA, low = NA)

# high knowledge agents reject any information tagged with
# the wrong partisan affiliation low knowledge ones take it
# all on board
for (k in 1:101) {
    # generate n agents with level of knowledge and partisan
```

```r
# affiliation
agents <- data.frame(knowledge = sample(c("high", "low"),
    n, replace = TRUE), party = sample(c("blue", "green"),
    n, replace = TRUE), success = NA)

# generate some information, tagged with partisan affiliation
information <- data.frame(info = sample(c(-1, 1), info_pieces,
replace = TRUE, prob = c(1 - results1$base rate[k], results1$base rate[k]
party cue = sample(c("blue", "green"), info pieces, replace = TRUE))

for (i in 1:n) {
    temp <- NA
    for (j in 1:info_pieces) {
        if (agents$knowledge[i] == "high") {
            ifelse(information$party_cue[j] == agents$party[i],
              temp <- append(temp, information$info[j]),
              NA)
        }
        if (agents$knowledge[i] == "low") {
            temp <- append(temp, information$info[j])
        }
    }
    agents$success[i] <- sum(temp, na.rm = TRUE)
}
results1$high[k] <- mean(agents$success[agents$knowledge ==
    "high"])
results1$low[k] <- mean(agents$success[agents$knowledge ==
    "low"])
print(k)
```

Epistemological Issues in he Theory of
Equality of Opportunity

John E. Roemer
Yale University

Abstract. EOp (Equality of opportunity) is an algorithm for measuring the degree of opportunity equality in a society. The researcher must identify the kind of advantage of concern (e.g., income , life expectancy, wage earning capacity), and what circumstances are important. She then partitions society into types, where each type consists of persons with approximately equal circumstances. The set of feasible policies must be delineated. If income is the kind of advantage under study, policies consist of different income tax functions. The algorithm holds people responsible for their effort, but not responsible for their circumstances. Compatibilism is discussed, as is the concept of an age of moral consent. Children, who have not reached the age of moral consent, are responsible for neither their nature nor their nurture. The scope of EOp requires us to recognize that EOp concerns itself with treating those are competing for social positions fairly, but does not address the issue of justice among those who consume what those competitors produce. Both are important.

Keywords: equal opportunity, meritocracy, circumstances, effort, type, responsibility, compatibilism.

I partition my remarks into seven sections.

I. THE EOP APPROACH

Consider a population whose members are competing for some valuable kind of *advantage* (G.A. Cohen's nomenclature - 1989), such as income, life expectancy, wage-earning capacity, or a state of good health. Call this kind of advantage the *objective.* To some degree, the achievement of the objective depends on *circumstances,* natural (such as genetic) characteristics of the individual and aspects of her environment over which she has no control (e.g., the socio-economic status of her family, her race and sex). Apart from these, her own *effort,* the degree to which she consciously decides to succeed in pursuing the objective, will matter. Finally, her success depends upon state policy, which may compensate persons with disadvantageous circumstances in order to improve their chances of acquiring the objective.

We partition the population into *types.* A type is the set of persons who have, as far as we can tell, very similar circumstances. Suppose, for example, we choose as circumstances the income of the family in which the individual was raised, her sex, her race, a rural versus urban background, and her IQ. These comprise five components. Suppose her race, sex, and rural v. urban background are binary variables, and we measure IQ and parental income as each having four possible levels. Then the number of *types* is $2^3 4^2 = 128$. If we have a population of 2^{20} souls,

about a million, this means on average a type will possess $2^{13} \gg 8000$ members. We will thus, with some confidence, be able to speak of a well-defined distribution of the objective within each type. The fact that the average size of a type is 8000 implies we can be fairly confident that the empirical distribution of income – if income is the objective – within a type, is something not due to chance, but is intimately associated with the circumstances of that type.

The *data* of the problem consist of the objective, the distribution of circumstances, a typology, the distribution of the objective in each type, and a set of social policies.

I have not said that it is these circumstances that *cause* the income distribution to be what it is, but I think we must also be able to assert causation. That is, we must have a theory of how variation in these circumstances causes variation in income. This theory will come from social science.

Let us continue with the example. What causes income to vary *within* types? One can be quite confident that, even if the number of types is fairly large – like 128 – there will be substantial variation of the objective within most types. The EOp theory asserts that this variation is due to *differential effort* within types. Effort, as I said, comprises the conscious actions individuals have taken to succeed in attaining the objective. There may as well be *episodic luck* – that is, the luck of being in the right place at the right time, or perhaps the effect of circumstances that we have not included in our list. I will say more about this below.

Returning to effort. The theory assumes that the more effort a person expends, the higher the level of the objective he achieves. I am thus discounting useless effort, which may be misguided because a person has an incorrect theory of causation. This means that, except for luck or useless effort, we may assume that the higher a person is in the distribution of income of his type, the more effort he has expended. In other words, if a person sits at the 60th centile of the income distribution of his type, then he sits at the 60th centile of the effort distribution of his type. This is the case *if* we can assume that all individuals in the type have faced the same *policy* (say, state financing of education), and because they all, by hypothesis, have the same circumstances. If all persons in a type have not faced the same policy, then the differences in the application of policy must itself be circumstances.

Thus, we can define the *degree of effort* of an individual as the centile of the income distribution of his type at which he sits. But how can we compare the efforts of two persons in different types? We must understand that the *distribution of effort* within a type is a *circumstance* of that type: that is, it is a fact of nature, not of any individual. That distribution is not associated with any individual, but with the circumstances and policy that define the type. So it *is* a circumstance for that type, and as such, persons should not be responsible for being in a type with a 'low' distribution of income – a distribution with, let us say, a low mean.

Because of this close association of the distribution of effort with the circumstances of the type, we cannot compare efforts easily across types. What we need is a measure of effort that sterilizes effort of its type-specific aspects. I say a measure that does this is the *centile* of the effort distribution of her type at which an individual sits. This is, as I've said, also the centile of the objective distribution of the type at which he or she sits.

So if a man, Jack, a member of a type of white men with advantaged social backgrounds, and a woman, Jill, a member of a type of black women with disadvantaged backgrounds, both sit at the 60th centiles of the respective income distribution of their types, I say they have expended *equal degrees* of effort, even though Jack may have spent many more hours studying than Jill did. This is the consequence of sterilizing the effort distribution of its type-specific characteristics. Of course, this view is only sensible if we have a *causal theory* explaining why Jill's circumstances made it harder to study than Jack's did. The fact that they sit at the same centile of their respective income distributions means *relative to others with the same circumstances,* they have performed equally well.

What is the optimal equal-opportunity policy? It is a policy that will compensate disadvantaged types with resources that improve their income distributions. To be specific, we'd like to find a policy that renders the income distributions of the types to be as close as possible to each other. In real life, we can never find a policy that will equalize all the income distributions *at the highest possible level*; given the space of feasible policies, there will be an optimal one. We need not here discuss the statistical details of how one measures the closeness of income distributions to each other: that's a technical issue, on which there is ample literature.

It is also desirable that we include as many circumstances as we can measure, for if the partition of the population into types is too coarse, we will implicitly assign to effort some achievement that is, in reality, due to an unmeasured circumstance. We now have some data sets with quite fine partitions of populations into types. Sweden, for example, has a statistical registry of all residents that permits us to have over a thousand types, with meaningful income distributions for each type. Sweden, however, is virtually unique in this regard. The United States and the UK also have very good data sets, but that's about it. There are no developing countries for which data sets currently exist that allow one to make a fine partition of the population into types. This means that when we measure the fraction of income inequality that is due to circumstances, we are in fact only computing a lower bound to the true number, because the paucity of data on circumstances means we will falsely assign to 'effort' effects that are in reality due to unmeasured circumstances.

Returning to the issue of episodic luck: I think the only way to deal with this is to increase the set of circumstances that we measure. There are some for dealing with luck, but it's beyond my scope to discuss these today.

II. COMPATIBILISM AND RESPONSIBILITY

The materialist thesis says that any action a person takes has a correlated brain state. Thought and action, in theory, can be read from brain states. *Compatibilists* say that the materialist thesis and the thesis that responsibility is meaningful are consistent. One can believe the materialist thesis and still assert that there are actions for which one can hold a person responsible. Incompatibilists say the two are inconsistent: if the materialist thesis is true, responsibility makes no sense.

This is, at least, my understanding of these terms.

I believe it's probably the case that most philosophers are compatibilists. What thinking supports the compatibilist view? I think the justification of the view is as follows. If one has contemplated the action, one was in a calm and sober state, one's powers of thought were whole, then one is responsible for the action. One way of countering this view would be to say that evolution has endowed us with a theory of responsibility that is very useful for survival, but is false. That is, having this theory of responsibility will help us avoid dangerous people, because we won't wish to associate with people who 'behave badly.' It's easier to justify ostracizing or disassociating with a person who responsibly takes bad actions, than one who is 'out of control.' One might desire to help a person who is out of control, even if it is dangerous to do so.

I think evolution may very well have endowed us with a theory of responsibility, but I don't say it's false: I say it is part of who we are. Of course, society may, over time, hold people responsible for a smaller set of actions than it did earlier in history, as we learn how circumstances cause behavior – that is, as social science matures.

I am a compatibilist. I generally hold people responsible for actions that they appear to have arrived at by calm, conscious thought, even though I recognize that there eventually may be a way of explaining those actions as due to circumstances for which the individual was not responsible.

A few years ago, I was writing a paper measuring equality of opportunity in several countries, and one of my collaborators proposed using brain scans of people that were available in our data set as circumstances. I strongly opposed doing so. Why? As a compatibilist, I believe that *every* action a person takes, responsible or not, has an associated brain state. Thus, showing that particular brain states were associated with the action tells us nothing about whether we should hold the person responsible for the action. I am not saying that brain-scan data are irrelevant, but at present, I think that using such data to excuse an individual from responsibility is only permissible if we have a causal theory of action.

III. THE PRIVACY ISSUE

I know of two philosophers (there may be more) who have objected to the EOp approach on the grounds that acquiring information about circumstances may violate

privacy. These are Elizabeth Anderson and Norm Daniels. I quote from Moriarty (2005), who writes that Daniels writes that 'responsibility tests' are 'intrusive and demeaning and violate concerns about liberty and privacy.'[1]

This concern reflects a poor understanding of how the application of the EOp approach proceeds. As I said, there are four decisions a social scientist must make: what is the objective, what are the circumstances, what is the typology, and what is the *policy space.*

I will focus now on the policy space – here is an example (see Roemer 2013 for the details). There is a population with two types: Rich and Poor. For the purposes of the example, we assume this is the single circumstance – being rich or poor. We will be concerned with the objective of life expectancy, and as policy makers in the Ministry of Health, we are not to be concerned with how people became Rich or Poor. Rather, from the Ministry's vantage point, income level is a circumstance that it believes should not, from the ethical viewpoint, affect a person's life expectancy. We wish to design a policy that will equalize life expectancies in the two types.

There are two diseases that afflict this population. The Rich suffer from cancer; the Poor suffer from cancer and tuberculosis. Within each type, the probability of developing cancer or TB is a function of *life-style quality*: how much one exercises, whether or not one smokes, diet, and so on. Life-style is the effort variable. There is, unsurprisingly, a different distribution of life-style quality in the two types. The Rich have a better distribution of life-style quality than the Poor: their mean life-style quality is high, and therefore they suffer a lower rate of cancer than the Poor. Only the Poor contract TB, and this probability is decreasing in their life-style quality.

Suppose the Ministry has a fixed budget for medical care. There are various treatments for each disease; the more a hospital spends on the treatment of a disease, the higher will be the life expectancy of the population receiving that treatment. Question: How much should the Ministry spend on each disease treatment? The data one needs to address fully the allocation problem include what the life expectancy will be within each type as a function of how much is spent on each occurrence of the disease. Knowing the life-style of a person who presents with cancer or TB may tell us something about the effectiveness of treatment.

Horizontal equity is a property of a policy under which anyone who presents with a certain disease (with a particular level of severity, etc.) receives the same treatment, regardless of his circumstances *or effort*. One might well want to consider only policies that are horizontally equitable – policies that will use the same treatment for any person who comes to the clinic with a given disease and a given level of severity. There are two reasons: first, privacy may be violated if the medical practitioner has to get details of

---

[1] Anderson's citation is similar (Anderson 1999, 289): "equality of fortune, in attempting to ensure that people take responsibility for their choices, makes demeaning and intrusive judgments of people's capacities to exercise responsibility and effectively dictates to them the appropriate uses of their freedom."

life-style quality and circumstances, and second, because this may put the practitioner in the position of triaging patients, which might reduce the trust between practitioner and patient.

In other words, one might not use certain policies that could improve the life-expectancy of the Poor, making it closer to that of the Rich, because there is a trade-off between doing so and other values one has – such as patient privacy and the quality of the relationship between practitioner and patient.

That is my first response to Daniels and Anderson. My second response is that, yes, we may want to use information about patients that violates privacy – indeed, we may decide this is essential – but that need only be done with a *small random sample* of patients. For this random sample, one acquires the private information, and then one experiments with different treatments. The experiment enables us to compute what the optimal treatments for each disease would be both by discriminating in the treatment of patients based on their personal information, and by ignoring this information. One can then decide how much particular kinds of information are worth knowing. Of course, it would be even better from the purely clinical viewpoint to get the exact life-style quality of the patient by asking probing questions, but it may be prudent not to compromise the patient's trust in the doctor and treatment by doing so.

Thus, the critique from Daniels and Anderson can be rebutted by understanding that the policy space of the Ministry may be chosen so that clinicians need not compromise patient privacy.

Another example might be whether upper middle-class parents who read a lot to their children should be taxed for doing so. Aside from the difficulties of enforcing such a policy, this is not a policy that any society today would support. As I said, the EOp theory is meant to implement a policy that conforms to the society's views about responsibility. Very few societies would support a policy of 'levelling down' like taxing parents who read to their children. The right policy is to provide more teaching personnel so that all young children can be read to by loving adults.

IV. THE AGE OF CONSENT

Children are humans in the state of formation, which includes moral formation. We should think of there being an age of moral consent below which a person is not responsible for his behavior. More subtly, one might wish to hold children increasingly responsible as they get older, although not being fully responsible until the age of 18 or so.

For simplicity, let us ignore the continuous approach, and say that the age of consent is a certain age during adolescence, and a person is not responsible for her actions below that age. Of course, we teach children what responsibility means by commenting on the morality of their actions. Not holding the child responsible does not mean we fail to remark upon the moral quality of the child's behavior.

It follows from this view that effort – responsible choice – strictly speaking, does not exist for children below the age of consent. Every action of a child, in this binary view, is the consequence of circumstances. Both nature and nurture are clearly circumstances. This implies that if we are performing an EOp analysis of income acquisition among young adults, and if we had a complete biography of the individual at the age of consent, that biography should be treated as a circumstance! If the age of consent were 16, then everything that happened to the child, and that the child did before the age of 16, should be treated as a circumstance or as being caused fully by circumstances.

Now one might say that this view will result in holding people responsible for very little about their achievement of the objective – because what the child has accomplished by the age of 16 may severely narrow the degree of acquisition of the objective that the child can attain by the age of 30. I say it would not be correct, on this account, to conclude that we should lower the age of consent: rather, our societies should undertake effective(s compensation policies beginning early in childhood so that by the age of 16, virtually all children have acquired good skills and a good moral character. This should begin with state-financed early-childhood education, and other policies (such as paid parental leave) that will produce more successful children. Europe is surely much further along in having such policies than the United States.

But if, despite our attempts, there is a great deal of variation in accomplishment by the age of 16, which engenders a great deal of variation in income acquisition by the age of 30, then the EOp view says we must attempt to compensate those who are disadvantaged, through policy. Adopting the view that before an age of consent sometime in adolescence the biography of the child is a circumstance will have quite radical implications for redistributive policy among adults.

V. THE STATISTICAL VERSUS THE IDEAL APPROACH

Philosophers, when considering a theory such as the EOp theory, delight in constructing hypothetical examples showing that application of the theory will result in errors. A genre of such examples are ones in which, with a given list of circumstances, one will always mis-categorize a certain outcome as being due to the lack of effort, when it is in fact determined by a missing circumstance. I say this is holding social science to an ideal-theory standard.

Social scientists and policy makers are concerned with statistical outcomes. We want to get things approximately right. At some point, in delineating the set of circumstances and the policy space, we will cease to rely on science and instead use intuition. But this is not a defect of the EOp theory – it is true of all social science, and perhaps of all science. Theories of causation require models of the world, and a model is a simplification of an impossibly complex reality. To be able to construct a theory, one must prove theorems about what happens in one's model. A modeler is always open to

criticism that she has ignored an important aspect of reality in constructing the model. Since *a priori* she has only a vague theory of what is happening in the aspect of the world she is studying, she must make the decision about what features of the world to include in the model and what features to ignore, or, as we say what features to 'abstract away from.' Having no complete theory at the point of model-construction, that decision must be made with intuition, not science.

The same is true in analytic philosophy. To construct a theory (say, of distributive justice) one constructs a model that abstracts away from features of reality that one thinks *intuitively* are not essential for the theory. Then, if the theorems one derives from the model seem bizarre, one changes the model and starts again. (Rawls described this process as one of reflective equilibrium, but the idea is much older than Rawls: surely Euclid constructed his axioms for geometry by trial and error.) I view social science and philosophy as in part arts, because at crucial points they rely on the intuition of the practitioner. Perhaps philosophy is more of an art than social science, for the social scientist can often observe whether the conclusions of the model appear to hold in reality, while philosophers deal with claims that are not observable.

VI. THE SCOPE OF EOP

A major application of the EOp theory is to the fair treatment of individuals who are competing for desirable positions in society: admission to good high schools or universities or medical schools, or acquiring government contracts. There is less (or no) emphasis on what people who acquire positions can do for the rest of society. In the US, we talk a lot about access to Ivy League universities for disadvantaged students; we talk less (or very little) about what the graduating classes will do for society as a whole. Of course, everyone understands that one of the university's roles is to train a skilled cohort of workers. But this is not seen as the moral issue that equal opportunity is.

Roughly speaking, *meritocracy* is the view that those who will be most skilled in carrying out the requirements of social positions should be recruited to those positions, and EOp focuses instead on fairness to the cohort of applicants for those social positions. My own view is that both considerations are important, and neither side should ignore the importance of the other.

Why do we laugh when I propose that we equalize opportunities for short players to be recruited to professional basketball teams? After all, being short is almost surely something beyond control of the individual. The answer is because there are only a handful of basketball players, but there are millions of fans. Here the welfare of the fans trumps the fair treatment of short players. I am here supposing that basketball games would be less interesting if every team had some short players. I do think, however, that we should practice equal opportunity for short players to join high school basketball teams: for these players may thereby become successful coaches, if not professionals.

More generally, I think we must consider how distant the social positions we are concerned with are from the final provision of consumption goods and services to citizens. I think expenditures on early-childhood education should be *solely* concerned with equalizing the achievements of children. I think admission to high schools should be *largely* concerned with equalizing opportunities for children, but to some degree be concerned with preparing a highly skilled cohort of adults to produce complex goods and services for the consumption of the citizenry. I think admission to medical schools should be concerned with both merit and opportunity. But I think the standards for passing the surgery board exams should be entirely based on merit: because the outcome is producing doctors who immediately will provide surgeries consumed by citizens.

Now a militant meritocrat will respond that if we admit a substantial number of disadvantaged students to the Bronx High School of Science in New York City, and other such excellent institutions, that will eventually reduce the pool of highly accomplished medical-school applicants, which will then force us to lower the standards on the surgery boards, to produce a large enough cohort of surgeons. So the average quality of surgery will degrade.

I think the proper way to evaluate this claim is empirical: it is not, in my view, an ethical question. If it turns out that admitting a substantial number of disadvantaged students to medical schools does degrade the quality of surgery the society achieves, then we must conclude it is too late at the point of medical school admissions to rectify the injustice of the disadvantage – the rectification must begin at an earlier age[2]. This is a similar conclusion to the one I reached in the example concerning the age of consent.

VII. RIGHT-WING VERSUS LEFT-WING USES OF EOP

The EOp theory that I have outlined attempts to equalize in a population the acquisition of an objective so far as it is influenced by circumstances beyond the control of individuals, but not to equalize the degree of objective acquisition to the degree that it is differentiated due to effort. The left-wing approach emphasizes the equalization-via-compensation-for-disadvantage aspect, and the right-wing approach emphasizes the differential-returns-to-effort aspect. Emphasizing the 'desert' aspect of EOp seems to me to be part of the right-wing approach. That is, the left-wing approach emphasizes the claim that surely people do not deserve the poor outcomes that are associated with

---

2] Only 10 black students passed the entrance exam for admission to the prestigious Stuyvesant High School in New York City in 2020. It is unclear, however, whether lowering the standards for admission for disadvantaged students, or giving crash courses teaching to the exam, are good policies, even if they succeed in raising the matriculation of black students. Much better, I conjecture, would be sharply to improve the public. elementary schools that most disadvantaged students attend. Even that may be insufficient: eliminating homelessness and poverty of families may be necessary. Obviously these policies are much costlier than achieving a higher black matriculation by teaching to the test. For the effects to show will take time. Rectifying the costs of racism does not come cheap, but 'no short-cuts' may be the right slogan.

disadvantageous circumstances, and the right-wing approach says that surely people do deserve rewards for effort.

When I first proposed the EOp theory (Roemer 1993), the paper was entitled "A pragmatic approach for the egalitarian planner." By pragmatic I meant that the social planner did not have a moral theory of responsibility, but rather would take the theory of moral behavior of the society in which she was working to determine the set of circumstances. I suppose it's what Rawls would have called a political, not metaphysical, theory.

I continue to hold this view: that the theory is a tool with which a society can construct social policies that are consonant with its own view of the causation of behavior. Having the language of the theory – circumstances, effort, type, objective, and policy – sharpens discussions. For instance, in the affirmative action debate that has taken place over the last forty years in the US, there have been two kinds of objection to affirmative action policies. The first claims that scarce social positions (e.g., in medical schools) should be assigned on the basis of predicted merit, while the second claims that race is not the right circumstance upon which to focus – rather, the relevant one is socio-economic disadvantage. The language of the theory tells us that the first objection is against the theory of equal opportunity as such, while the second is only quibbling with what the set of circumstances should be. It is interesting to note that policies have evolved, in the main, to respond to the second objection rather than the first. That is, the circumstance of race has often been replaced in university admissions policy by one of socio-economic disadvantage. But recruiting for social positions on the basis of merit alone is now quite rare.

john.roemer@yale.edu

REFERENCES

Anderson, E. 1999. What is the Point of Equality?. *Ethics* 109(2): 287-337.

Brunori, P., Hufe, P. and Mahler, D. In press. The roots of inequality: Estimating inequality of opportunity from regression trees and forests. *Scandinavian Journal of Economics.*

Cohen, G. A. 1989. On the Currency of Egalitarian Justice. *Ethics* 99(4): 906-44.

Moriarty, J. 2005. The epistemological argument against desert. *Utilitas* 17: 205-21.

Roemer, J. 1993. A pragmatic theory of responsibility for the egalitarian planner. *Phil. & Public Affairs* 22: 146-66.

———. 2013. Equity in health care delivery: Some thoughts and an example. *Research on economic inequality* 21: 471-89.

PUBLIC REASON

Journal of Political and Moral Philosophy

ARTICLES