

A Democratic Critique of Epistocracy with Artificial Intelligence as a Case Study

Joachim Skahjem

FI2900

Institute of Philosophy and Religious Studies, NTNU

Spring 2021

Word Count: 6430

Contents:

Introduction.....	3
Epistocracy.....	4
A hyperintelligent A.I. system.....	5
The Epistocratic Argument.....	7
Political Competence.....	8
The Epistemic Limit.....	10
What is Democracy?.....	11
Tyranny.....	12
The Relevancy Requirement.....	14
Qualitative Decisions.....	15
Conclusion.....	16
References.....	18

Introduction

Plato suggested that society would be best ruled by those most qualified by virtue of their qualification. This is a form of aristocracy, specifically epistocracy. In this paper, I will use David Estlund's critique of epistocracy as the starting point and argue, in opposition to Estlund, that it is on a procedural, and not outcome-centred, basis that epistocracy is rendered unjustifiable. It will be held in this assignment that a procedural approach would place the critique of epistocracy on a firmer basis. This will be illustrated through a hypothetical artificial intelligence system used as a case study. That authority does not follow, even when the epistocrats in question are demonstrably superior in political competence to the median citizen. This, I will argue, is because a democracy is significantly more suited than epistocracy to prevent the ultimate political injustice: tyranny.

Tyranny is a so-called primary bad. Estlund underlines that not all political decisions or results are qualitatively the same. Some, those he calls primary bads, must be avoided at all costs (Estlund, 2008: 163). Whereas Estlund's *Democratic Authority* concludes that democracy can be more accurate than epistocracy in terms of primary bads, I will argue that due to the procedural nature of democracy it is guaranteed to avoid certain primary bads, specifically that of tyranny. For tyranny is, in itself, undemocratic. It violates the requirements of the democratic procedure. This does, of course, rely on a very strict understanding of democracy with a very limited denotation. Similarly, the definition of tyranny is crucial to the argument.

Democracy is defined roughly, and it will be further explicated later, as a form of rule wherein people are only subject to policies whose procedure they themselves were included in. This will be formulated into a relevancy requirement for political inclusion, so that a democratic decision is understood as any decision wherein the decisionmakers consisted of all the relevant parties, and only the relevant parties. Conversely, the failure to meet the relevancy requirement would lead to tyranny. Tyranny is here power as domination, which is to say whenever decisions are so free from a limiting procedure that they can be arbitrarily imposed, such as the authority a master has over a slave (McCammon, 2018). The main argument of this paper is thus twofold. A failure to meet the relevancy requirement is also a failure to safeguard against tyranny (1) and epistocracy, unlike democracy, is unable to meet the relevancy requirement due to its non-procedural nature (2).

Epistocracy

Epistocracy, or expert-rule, is a form of aristocratic rule wherein the basis for authority is knowledge, particularly in the form of political competence (Viehoff, 2016: 406). Plato argued that ruling was like any other profession in society and so the job should be given to those who are the most qualified just as a doctor and a carpenter are hired on the basis of their unique expertise rather than through an electoral process (Plato, 2004: 97). It can be described as a “dictatorship by experts”, and it is therefore an autocratic theory (Estlund, 2008: 24). Autocratic elements can actually be found in democracy as well, for instance toward children and animals who are excluded from the political procedure on an epistemic basis. What makes epistocracy stand out however, is that its standard for epistemic qualification is much higher than that of a democracy. Democracy, on the other hand, has procedure as its high standard for qualification.

Pure proceduralism is not concerned with the inherent justness of any policies, but rather the justness of the manner in which those policies were arrived at, that the procedure itself is considered just (Estlund, 2008: 65-66). This is how Estlund sees proceduralism, and it is in his overly critical assessment that he has assigned democracy a less procedural nature than it ought to have. In other words, Estlund argues from the viewpoint that it is first and foremost democracy’s concern for outcome, and not its proceduralism, that makes it preferable. This becomes clear through his example of determining all policies by a coin flip. It would be the simplest way of arriving at a fair procedure David Estlund argues, but democracy aims, due to its epistocratic elements, for a system that has a greater accuracy rate than the 50% of a coin flip (Estlund, 2008: 66). The underlying idea is that a large enough group of people has a greater chance of being right than wrong as long as each individual has a slightly better chance of being correct than incorrect in any political matter (Estlund, 2008: 15). However, I would argue that democracy’s rejection of the procedural coin flip method is not an example of epistocratic elements within democracy. It is rather a desire for fairness as the coin flip method would fail to grant fairness to any decision. This can be seen if one considers John Rawls’ thoughts on the natural lottery. The natural lottery is a coin flip, or a roll of the dice when there are more than two possible outcomes, and Rawls argues that it is anything but fair (Rawls, 1999: 64). If one’s state of poverty is determined by a coin flip, either in the natural lottery or a case of political proceduralism as presented by Estlund, there is nothing to compel that person to accept the outcome simply because they had an equal chance to the wealthy

person of becoming wealthy. The point at hand is that democracy is not less procedural than a coin flip as Estlund posits, if anything it is more procedural as it has a stricter requirement than that of equal chance. It requires fairness in the procedure through relevant participation as opposed to an external force, here a coin, determining one's fate. Thus, the coin is not rejected in the belief that a group of people is more often correct than a coin flip, but rather that a political decision is not fair if it is made arbitrarily.

There is, however, a clear example of an epistemic element in democracy, namely in the political exclusion of various agents purely on the basis of their political incompetence. Children and animals are, as mentioned, entirely excluded from the political process and are at the mercy of adult human policy the same as the average citizen would be to an epistocrat. This is a traditional epistocratic feature that Glaucon, a character in Plato's dialogue *The Republic*, defended by stating that he would not let his children rule in the city he designed together with Socrates: Kallipolis (Plato, 2004: 230). That was an argument for epistocracy, but democracy excludes children and animals all the same. Adult humans are epistemic autocrats in relation to animals and children. This is not to say that those groups are necessarily oppressed by adult humans, but rather that the only thing hindering oppression and tyranny towards those groups is the ruling group itself. In other words, adult humans as a collective group lacks checks on its power in the same way any epistocratic group would. This is where the supporters of epistocracy would raise a crucial question. If democracy excludes certain groups on the basis of political incompetence in favour of rule by the more qualified, why stop at children and animals rather than continuing all the way into a true epistocracy?

A hyperintelligent A.I. system

A conceivable democratic critique to that question could either be that there are no agents, besides adult humans in relation to children and animals, in society who are demonstrably epistemically superior to others, or that there is such a superiority but it is too insignificant to warrant political exclusion of the less competent parties. When considering traditional epistocracies such a critique would do. Plato suggested that education would lead to, and to an extent reveal pre-existing, epistemic differences in people significant enough to warrant authority (Plato, 2004: 208). Even if that point were to be granted, which I would be hesitant to do, there still remains the problem of demonstrating that some particular individuals are fit to have authority whereas others are not. Finding who, among the educated, are fit to rule and

who are not. Take, for instance, the epistemic exclusion of children found in democracy. A right to vote is often given at the age of 18. There are undoubtedly certain 17-year-olds who are as competent or more competent than certain 19-year-olds or 47-year-olds for that matter. However, since identifying which ones those are is a potentially impossible task, the voting limit is set at a certain age with no exceptions, even in cases where it would make epistemic sense to make exceptions. An epistocratic system of rule would have to find a way to evaluate each and every citizen in an effort to find out who could be included in political participation and who would merely be a subject of the policies determined by the epistocrats. Education or test-taking might narrow the field of eligible actors, but it would not get anywhere near a Platonic ideal of a small class of philosopher-kings that are vastly superior to the masses.

This is where the hypothetical artificial intelligence system comes into play. None of the aforementioned apprehensions are relevant if the epistocrats at hand are in fact demonstrably epistemically superior to others, such as an A.I. system. For the purposes of this paper two premises will be used as unquestionable axioms concerning the state of being of the A.I. system so as to establish it as the most perfect epistocrat it can be. This is to avoid a strawman fallacy wherein flaws in the A.I. system is mistaken for flaws in epistocracy as a system of rule.

The first premise is that this A.I. system has undergone an intelligence explosion. The intelligence explosion is a hypothetical series of events wherein a human-made A.I. system is intelligent enough to create another A.I. system that is more intelligent than itself and so on and so forth to eventually arrive at an A.I. system that is hyperintelligent (Khatchadourian, 2015). This state of hyperintelligence is vastly more intelligent than any human being is, including the combined intelligence of human beings working together, even when using computers and other technological aids. The second premise is that this A.I. system has developed a form of sapience, that is to say human-like consciousness, that allows it to think beyond binary code, so as to avoid claims that it would be unable to comprehend moral problems and similar issues. These two assumptions are the only facts concerning the A.I. system in this thought experiment. Whether it has free will, is capable of emotions, has individuality or is a hive mind and other questions remains unanswered and will not be used as a defence or critique of its claim to rule. It is possible that certain qualities would make this A.I. system a better or worse epistocrat, in any such case it is always granted the system whatever quality makes it a better epistocrat. That is to say more politically competent, and it is distinct from the more holistic claim of being a better ruler. For whether a perfect epistocrat

is a better ruler than a democratically elected one is precisely the question this paper intends to investigate.

The Epistocratic Argument

With the hypothetical A.I. system established it is time to investigate whether this vastly improved epistocrat would have a just claim to rule or if there are valid democratic criticisms that remain sufficient even in light of a potential epistocrat of far superior political competence to any and all human citizens. The epistocratic argument consists of three premises and a conclusion, and the aim of a democratic critique of epistocracy is to either disprove one of the premises or show that the conclusion does not follow from the premises so as to render the conclusion invalid. The following argument for epistocracy is, of course, not the only conceivable way of attempting to justify epistocracy, but it is a fairly useful starting point.

The first premise is the truth claim and it states that there are such things as political truths (Lippert-Rasmussen, 2012: 243). In other words, some political decisions are objectively better than others. A political view being true in this context will simply mean that if x is wrong, right, advantageous, unfair or what the case may be, then it is true that x is wrong, right, advantageous, unfair or what the case may be (Estlund, 2008: 5). The alternative to this would be some level of political relativism and that view would, if granted, disallow epistocracy. If there are no objective standards for politics to aspire to then it would be nonsensical to speak of any actors who are more or less likely to meet those standards than others. While there are certainly examples of political decisions that have very few moral implications and perhaps lack a clear objectivity, there are also a great number of policies that have tremendous moral implications. Decisions concerning war, economy, climate action, healthcare, education, the criminal justice system and a myriad of other topics cannot be seen as simply a matter of opinions wherein no opinion is more or less correct than others. Even if identifying which opinions are which is a much more difficult task. Estlund introduces the concept of qualified points of views, those are viewpoints that, though not necessarily correct, are reasonable to some extent (Estlund, 2008: 33-34). The purpose of that concept is to illustrate that not all political stances share a qualitative value, for instance opinions held by a Nazi on a great variety of topics would be disqualified, and thus disregarded, on the basis of being unreasonable. That standard of qualified points of view also serves as a useful tool concerning the first premise of the epistocratic argument. The truth claim can be thought of as

correct as there are certainly political decisions and views that are more reasonable than others.

The second premise is the privileged knowledge claim. It states that there are some members of society that have more knowledge of political matters than others (Lippert-Rasmussen, 2012: 243). That premise undoubtedly holds true as the A.I. system is already established as far more knowledgeable than humans, so one group in society is significantly more likely to identify correct decisions than any other group. This is coupled with the final premise, the authority claim, to reach the epistocratic conclusion. The authority claim posits that those with privileged knowledge should rule as they have more knowledge of politics than non-epistocrats and then the epistocratic conclusion has been reached (Lippert-Rasmussen, 2012: 243). Since the first premise is held as true and the second becomes true through the imposition of the hypothetical A.I. system, the critique of epistocracy in this paper will attack the authority claim. The claim that knowledge yields authority. However, the democratic critique of an A.I. based epistocracy is only successful if it can justify the political exclusion of children and animals in the face of a more politically competent group (adult humans), while at the same time oppose the political exclusion of all humans in the face of a more politically competent group (A.I. system).

Political Competence

It was established earlier that the A.I. system is undeniably more competent than humans. However, that does not necessitate greater political competence. Political competence is here used as the ability to discern between good and bad policies as well as all other politically centred issues including questions concerning who ought to rule. This very discussion is an example of political competence in that regard. Political philosophy is, however, only a part of political competence and it would be unreasonable to deem those not well-versed in political philosophy as politically incompetent. A traditional Platonic epistocracy might require a political-philosophical education of sorts for each and every epistocrat, but something being a requirement for great political competence is not equal to its omission meaning political incompetence.

This has to do with the false equivalency that arises for the kind of epistocracy that Estlund, as well as this paper, has as the recipient of its critique. It is important to note that the epistocracy discussed in this paper, one that aims to maximize expertise and is based on an idealist Platonic model, does not encompass every epistocratic theory. A certain epistocracy

might see a disqualification of the mildly incompetent as opposed to only the severely incompetent as in a democracy, without requiring the disqualification of the mildly competent in favour of the supremely competent, like a Platonic democracy. It is, I hope, clear at this point that this paper aims only to be a critique of the latter kind of epistocracy. The kind that aims to maximize its expertise and disqualify all but the most competent. It is that form of epistocracy that falls victim to a particular false equivalency.

The “political competence” term used by the epistocrat is, by necessity, different from the one used by the democrat. That is due to where each theory sets the standard for political inclusion. Both the epistocrat and the democrat could agree that complete political incompetence warrants political exclusion, but the requirements for political competence are vastly different. An epistocracy, of the sort this paper deals with, is based on a relative standard whereas democracy, it could be argued, is based on a set standard. The politically competent in an epistocracy are at any given time those that are more competent than everyone else (Estlund, 2008: 30). Thus, if an epistocracy had been implemented prior to the existence of the A.I. system the human epistocrats would be replaced by the A.I. system upon its arrival. The same principal that granted them their authority before the existence of the A.I. system would be the exact same that disqualified them of that authority later on. The fallacy by the epistocrat is to assume that democracy uses this same relative standard when it excludes the politically incompetent, that is to say children and animals. Had that been the case then human adults should naturally be considered incompetent when the A.I. system arrives.

In a democracy, however, adult humans are not regarded as politically competent by virtue of being the most competent actors present, but rather by virtue of being competent enough, being above a certain limit. That limit might be one that requires the ability to identify one’s own best interest as well as an ability to partake in political discourse or discussion, for instance. Those, as well as the ability to identify the interests of the wider society, are mentioned by J.S. Mill as reasons to extend political inclusion to all societal classes, as everyone can learn to identify those interests through political discourse (Schwartzberg, 2015: 193). I exclude the ability to identify the best interests of wider society as that has a moral component. Namely, the assumption that the best of society ought to be strived for, rather than being strictly competency-based, and competency is the only relevant factor at this point.

Children of a certain age and animals are below this limit, and the vast majority of human adults are above it. There are definitely a few inconsistencies concerning this. A 17-year-old

child is likely to be above that particular limit, and if the limit had been raised so as to ensure that the 17-year-old child is definitely below it, then most people of all ages would also be below it. However, the fact that the epistemic limit found in a democracy is imperfectly put, and cannot ever be perfectly placed, does not mean that it is not correct to feature it in any capacity. The most important thing for the democratic critique of epistocracy to work is that the limit it sets is based on a reasonable principle, and that adult humans, as a more or less collective group, are in fact above that limit.

The Epistemic Limit

The epistemic limit for political competence, and thus political inclusion, in a democracy is based on the two aforementioned principles. The first being the ability to identify one's own best interest, and the other being the ability to partake in a political discourse wherein one can be educated or educate others. It is crucial that adult humans are above that limit as well as that being a reasonably set limit in the first place. It is preferable to exclude anyone from the political process who would be better off if someone else determined policies on their behalf, such as children. It is also preferable to exclude those who are unable to partake in a political discourse and as such are unable to persuade others or themselves be persuaded by reasonable arguments, such as animals or very young children. It is not only clear that young children and animals are below this limit, but also that the vast majority of adult humans are above it.

The epistemic limit as it has now been set is not a relative one, and therefore the existence of the A.I. system does not automatically render humans politically incompetent. Some might argue, however, that most adult humans are not politically competent even by the epistemic limit used in a democracy, but that would be based on a conflation between democracy and regimes commonly referred to as democratic. I will elaborate. Humans have, as a group, created poverty in the absence of scarcity, war in the absence of an enemy outside the group itself, severely worsened the conditions on the very planet we live on and created weapons capable of our own mass extinction. These seem like incompetent decisions made by a group not qualified to rule, but they are rather decisions made without sufficient procedure. It is rare that the people who make these decisions are also the ones who are subject to their consequences. Politicians are not the most common casualties of war, policies creating or sustaining poverty very rarely, if ever, affect the rulers who made them and mass polluters are usually not the ones who bear the brunt of any ecological crises. The point being that these decisions, these primary bads, violate the democratic procedure. They could therefore not

occur in a democracy, even though they frequently occur in regimes commonly designated the term “democracy”. Simply put, it is not the incompetency of rulers that most often lead to primary bads, but rather a lack of proceduralism. The ability for rulers to act arbitrarily and not be limited by procedural requirements. Thus, the solution to those kinds of severe problems in society, tyranny among them, is not a maximization of competency at all, but rather a stricter procedure. The solution is democratic rather than epistocratic.

What is democracy?

The form of rule used as an antithesis for epistocracy by Plato has been accused of being a “mobocracy disguised as democracy” (Dahl, 1989: 77). A form of rule wherein large groups of people rule, directly or indirectly, without a systemic process that ensures that minority oppositions are heard, or that certain policies are disallowed in spite of majority support. It is a misconception of democracy as meaning simply majority rule, and nothing more. Most regimes called democracies tend to be more complex than that, and democracy itself is different not only from a Platonic mobocracy, but also actual regimes called democratic. The democracies used to either critique or defend epistocracy tend to be based on these actual regimes.

Democracy as a political-philosophical ideal must be distinguished from actual regimes that are colloquially called democracies. No form of rule is ever the same in theory and practice, and since the A.I. led epistocracy is purely theoretical, the democratic model it is critiqued by should be so as well. Democracy is a rule by the people. In actual regimes there are all kinds of hindrances to democracy such as propaganda/misinformation preventing the people from having all the relevant perspectives, reduced options that only let people vote between some of the many options that exist, voter suppression, a resource-based disparity in a system that allows for lobbyism and so forth. The purpose of clarifying the distinction between a democracy and regimes called democracies is to ensure that a democracy is not criticized for a feature that is not inherent to it, but simply a part of what we may call common democracy, which is to say any political regime traditionally called a democracy.

An elected leader in a common democracy might make a political decision that would send a part of the population into poverty. Even though that part of the population certainly would not approve of it. That decision, however, is undemocratic as it violates the democratic procedure. It would be a mistake to think that the democratic procedure is nothing more than referendums or parliamentary votes. The democratic procedure is a strict process that aims to

ensure that people partake in the decisions whose consequences they are subject to. That renders several policies de facto illegitimate even though they may have majority support, as they lack the relevant majority support. There will be more on that point later as the relevancy requirement mentioned in the introduction will be further explicated.

Tyranny

An understanding of democracy is not complete without a comprehension of its counterpart: tyranny. That is, of course, an assumption made in favour of democracy as any form of rule aims to be the counterpart of tyranny or other things of that nature. However, it is not arbitrary that I call democracy the antithesis of tyranny. It is because democracy, as understood in this paper, is born out of a definition of tyranny. The democratic procedure is democratic if, and only if, it can safeguard against a tyranny. The definition of tyranny is therefore crucial and it was briefly given in the introduction. It is a matter of domination, which is to say when a person, or in this case a state, has the ability to impose its will arbitrarily (McCammon, 2018). An arbitrarily made decision is one made without regard to a procedure that goes beyond “the decision must be made by the ruler to be legitimate” or a limitation such as a law that the ruler cannot just change at will. If the procedure simply requires the ruler (x) to be the one making the decisions affecting the subject (y) then x has no procedural limitations of consequence. There is then no conceivable act that x cannot commit towards y, for there is no one but x to stop x. This is the case even if x is not a singular ruler, x can here be a population majority in a common democracy, or a group of epistocrats, including the artificial intelligence system, and the domination would still remain.

The epistocratic argument presumes that the goal of ruling is to reach the best possible decisions and consequently bases its argument on the notion that it would better accomplish that than democracy. However, the ultimate goal of a democracy is, I will argue, not to maximize the good outcomes, but minimize the bad ones and in particular avoid the very worst one, which is tyranny. This concept is referred to as primary bads, political decisions whose avoidance is more important than any other political consideration (Estlund, 2008: 163). Tyranny is not best avoided by political competency; in fact, it could be argued that successfully creating and sustaining a tyranny requires a great deal of competency. Instead, it is avoided by removing the opportunity. Tyranny is born out of opportunity and any autocratic element is by necessity an opportunity to abuse power. Democracy is not flawless in this respect, for instance it was mentioned earlier that children and animals were entirely at the

mercy of the goodwill of adult humans. That is because adult humans are autocrats, more specifically epistocrats, in relation to children and animals. The justification of that autocracy is, as discussed, that due to the severe political incompetence of children and animals they are better off with political exclusion as their own decision-making would not benefit them. Importantly, the justification of an autocratic element does not necessitate the justification of a complete autocracy, such as an A.I. led epistocracy would be.

It is worth noting that tyranny is prevalent in common democracy, and it takes the form of a tyranny of the majority. A phenomenon often occurs where >50% of the population rules over <50% of the population through an electoral process (Estlund, 2008: 37). While Estlund does not designate this phenomenon as a tyrannical majority, other thinkers such as Burke, and most famously Tocqueville, do and so does this paper in tune with its definition of tyranny (Horwitz, 1966: 293). This is entirely anti-epistocratic in nature, and its radical anti-epistocratic nature could easily justify a complete epistocracy through a false dichotomy. For if the choice is between a tyranny by the politically competent and a tyranny by a fairly random majority, then the choice quickly falls on the former.

Some have argued that a majority has epistemic value by virtue of each individual being slightly more likely to be right than wrong (Schwartzberg, 2015: 192). However, that assumption, even if correct, would lead to an unreasonable dismissal of the opinion of all political minorities, as it assumes that having less support for a standpoint increases its likelihood of being incorrect. It completely ignores the vast history of political consensus wherein views that enjoy majority support now did not do so in the past. Yet, they were as correct, or incorrect, then as now. A better alternative would be Estlund's concept of qualified points of view, so that no point of view is more highly regarded by virtue of the number of people who support it, but only by virtue of the arguments that support it. Thus, a majority has no greater epistemic value than a minority. That would render a tyranny of the majority non-epistocratic, and the epistocrat would conclude, correctly, that an epistocratic tyranny is preferable.

However, that does, of course, assume the false dichotomy of an epistocratic and a non-epistocratic tyranny as the only options. Democracy does not, in theory, lead to a tyranny of the majority or any other tyranny and can therefore be justified over an A.I. led epistocracy if avoiding tyranny is held as the highest priority. This is due to the strict procedure of a democracy and its lone premise of having a rule by the people. If the people, rather than a majority of the people, rule then no one is ruled by anyone else. It would accomplish

Rousseau's ideal democracy wherein each citizen only obey themselves (Estlund, 2008: 37). Or rather it would achieve what Rousseau thought he achieved by simply allowing for unfettered majority rule without designating it as a tyranny. By the definition of tyranny in this paper, however, an unchecked majority is tyrannical due to its capacity for domination. This standard seems impossible, how can a democracy demand that people are only ruled by themselves and not others?

The Relevancy Requirement

The democratic procedure does not require that all people determines their own political fate completely or has a right to partake in the decision-making process of every decision that affects them. For every political decision in any given society can be conceived as having an impact on every single person, also those not in that society. The election of an American President certainly impacts Norway, but that is not a case for Norwegian participation in the electoral process. Instead, it is, as mentioned in the introduction, a matter of relevant participation. The relevancy requirements found in common democracies are geographical and predicated on citizenships. Thus, a French citizen votes on French policies whereas a Norwegian citizen partakes in the political process in Norway. Democracy, as a political-philosophical ideal, must go further than this to avoid domination, which is to say tyranny.

Several political decisions must be disqualified on the basis of not including all the relevant parties. These are decisions that are not disqualified in an epistocracy as the relevancy requirement is purely procedural. For instance, Norway bombed Libya in 2011 as a joint effort with several other NATO members (Ask, 2017). This decision, I will argue, was entirely undemocratic. However, it was not undemocratic due to a lack of majority support in the Norwegian population or in parliament, it was rather due to a lack of support from the most relevant party: the Libyan people. The decision to bomb Libya was undoubtedly one that affected the Norwegian society due to its bearing on Norway's continued NATO membership and relation to the United States. The impact was, however, far greater on the Libyan population and their disapproval vastly outweighs any Norwegian approval. Borders have become artificial limitations on the bearings of democracy, and therefore common democracies often allow for actions that violate the democratic procedure. An A.I. led epistocracy might not have bombed Libya in 2011 had it been in charge of the Norwegian state at the time as it is at this point widely believed to have been a very incompetent decision that only worsened the condition of Libyan society through morally critique-worthy methods

(Ask, 2017). The artificial intelligence system might have recognized this and ruled against the bombings. However, the epistocracy has no procedural limitation hindering it from committing the act. It relies entirely on the decision-making of the rulers in the moment. Whereas a democracy would disallow it entirely. Not due to its outcome, but because of its procedure.

Epistocracy utilizes a competency-based requirement for political inclusion. This same requirement occurs, to an extent, in democracy as has been discussed in relation to children and animals. The point at hand is that the relevancy requirement takes precedence over the competence requirement in a democracy. Even if the avoidance-of-incompetence requirement, the one applying to children and animals, might take precedence over both of them. An A.I. led epistocracy could never pass the relevancy requirement as it would lead to an artificial intelligence system making decisions pertaining to humans. The problem is not the machine-human distinction, but the distinction between ruling epistocrats and ruled subjects. The complete removal of the subjects from the process of decision-making is that which sets the stage for arbitrary impositions of will and thus tyranny.

Qualitative Decisions

It is important to note that the democratic critique of epistocracy is not a strictly procedural one, it is also outcome-based. For the underlying assumption is that primary bads will be made more often in an epistocracy, even one led by a hyperintelligent A.I. system, than in a democracy. A long series of good, but insignificant, decisions pales in comparison to one grave mistake (Estlund, 2008: 163). Take, for instance, the example of slavery. Whenever the politics from the 18th century, for instance, is evaluated slavery is mostly the topic people find relevant, as any number of other potentially good decisions fail to make up for this one terrible one. An epistocrat might suggest that slavery occurred as a result of incompetency and for that reason an epistocracy, particularly one led by someone as competent as the A.I. system, is the best way to avoid qualitatively poor decisions of that magnitude. However, slavery occurred, arguably, not as much due to incompetence as due to opportunity and unreasonable political exclusion. The case of a slave and a master can be viewed as a scenario in which there are two competing groups. In the 18th century, the slave owners, that is to say white people, were included in the political process whereas black people were not. The decision to enslave black people were in other words not made by black people. That is a clear case of unreasonable political exclusion, and had black people been the only people

included in the political decision as democracy would have it, then it stands to reason that the decision to outlaw slavery had been made. In just the same way every other ethnic, geographical, religious and socio-economic group would be the only group involved in the decision regarding their own enslavement. For a slave cannot be tyrannical towards a slave owner. A slave demanding its freedom against the will of, and without the political inclusion of, the master, is not tyrannical, but rather precisely what democracy aims to be.

What actually occurred, however, was far more akin to epistocracy than it was democracy even if it did occur in a common democracy of sorts. White men with property were the only ones included in the political process, so it can by no means be referred to as either a democracy or an epistocracy. It might be best described as a gender-and-race-based timocracy, but the most important point is that it was autocratic, just like an epistocracy is. It was precisely due to this autocracy that it could exclude certain people from the political process and thus had the opportunity to enslave them. An epistocrat might suggest that a political decision as bad as that one would not be made in an epistocracy as long as the epistocrats ruling are sufficiently competent, as the A.I. system would be. However, the point here is not what an epistocrat is likely to do, but rather what they can do by virtue of what the autocratic model allows for. A democracy would inherently disallow slavery or something of the sort as it is by necessity tyrannical for one person to at all times determine what another person should do. It is quite simply the clearest form of domination. A slave owner is an autocrat and a tyrant by definition, so policies like that would be disallowed anyway, and not dependent on the result of a vote or the decision of a ruler. That is why an epistocracy, even one with epistocrats as competent as a hyperintelligent A.I. system, is unjustifiable. For even if procedure is disregarded, as it ought not be, the best way to avoid poor outcomes is not through competency, but through relevant political representation and inclusion by ensuring all political exclusion is relevancy-tested.

Conclusion

In summary, a justification of authority from knowledge depends on competency being the most vital requirement of any ruler, but, as argued in this paper, the avoidance of tyranny is even more significant. The goal of avoiding primary bads, tyranny being merely one example, is one shared with Estlund's *Democratic Authority*. However, this paper differs in approach by basing the critique of epistocracy primarily in procedure, with improved outcomes only

being seen as a natural consequence of a democratic procedure. It is in the end concluded that an epistocracy, even one led by a hyperintelligent A.I. system, is unjustifiable.

References:

- Ask, A. O., 2017. «Norge slapp nesten 600 bomber over Libya i 2011 – nå skal innsatsen granskes». Aftenposten. <https://www.aftenposten.no/norge/i/MxXOR/norge-slapp-nesten-600-bomber-over-libya-i-2011-naa-skal-innsatsen-gr>
- Dahl, R. A., 1989. *Democracy and Its Critics*. New Haven: Yale University Press.
- Estlund, D. M., 2008. *Democratic Authority*. Princeton: Princeton University Press.
- Horwitz, M. (1966). Tocqueville and the Tyranny of the Majority. *The Review of Politics*, 28(3), 293-307. <http://www.jstor.org/stable/1405588>
- Khatchadourian, R., 2015. The Doomsday Invention. *The New Yorker*, A Reporter At Large; 23 November 2015 issue. <https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom>
- Lippert-Rasmussen, K., 2012. Estlund on Epistocracy: A Critique. *Res Publica* 18, 241–258. <https://doi.org/10.1007/s11158-012-9179-1>
- McCammon, C., 2018. Domination. *The Stanford Encyclopedia of Philosophy*, Winter 2018 Edition, Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/win2018/entries/domination/>
- Plato, 2004. *The Republic*. (Translated by C.D.C. Reeve). Indianapolis: Hackett Publishing Company. (Original work published c. 375 BC).
- Rawls, J. B., 1999. *A Theory of Justice: Revised Edition*. Cambridge: Harvard University Press.
- Schwartzberg, M., 2015. Epistemic Democracy and Its Challenges. *Annu. Rev. Polit. Sci.* 18:187–203. <https://doi.org/10.1146/annurev-polisci-110113-121908>
- Viehoff, D., 2016. Authority and Expertise. *The Journal of Political Philosophy*: Volume 24, Number 4, 406-426. <https://doi.org/10.1111/jopp.12100>