The Problem with Intelligence:
Its Value-Laden History and the Future of AI

Stephen Cave

Leverhulme Centre for the Future of Intelligence, University of Cambridge

sjc53@cam.ac.uk

Abstract

This paper argues that the concept of intelligence is highly value-laden in ways that impact on the field of AI and debates about its risks and opportunities. This value-ladenness stems from the historical use of the concept of intelligence in the legitimation of dominance hierarchies. The paper first provides a brief overview of the history of this usage, looking at the role of intelligence in patriarchy, the logic of colonialism and scientific racism. It then highlights five ways in which this ideological legacy might be interacting with debates about AI and its risks and opportunities: 1) how some aspects of the AI debate perpetuate the fetishization of intelligence; 2) how the fetishization of intelligence impacts on diversity in the technology industry; 3) how certain hopes for AI perpetuate notions of technology and the mastery of nature; 4) how the association of intelligence with the professional class misdirects concerns about AI; and 5) how the equation of intelligence and dominance fosters fears of superintelligence. This paper therefore takes a first step in bringing together the literature on intelligence testing, eugenics and colonialism from a range of disciplines with that on the ethics and societal impact of AI.

Introduction

While much attention in the literature on AI has been given to defining the concept of intelligence (Legg and Hutter 2007), very little has been given to interrogating how this concept has been deployed in normative or political contexts. The aim of this paper is to begin to address this, asking not what intelligence is, but rather what the concept of intelligence has historically been used for. It focuses in particular on the use of the concept to establish and legitimate power hierarchies. It then asks how the legacy of this usage might be shaping debates about machine intelligence.

This paper is therefore an assessment of intelligence as a value-laden term — what is sometimes called in moral philosophy a 'thick concept', one that is "both evaluative and descriptive" (Kirchin 2013, 1). It is an attempt at disentangling the evaluative aspect and making it explicit,

so that we understand in full measure what we are doing in using this term (Baisini 2015, pt. 8.1.2); an act of 'consciousness-raising' as Simon Blackburn puts it (Blackburn 2013), or what another tradition calls 'deconstructing' — "exposing a concept as ideological or culturally constructed rather than as natural or a simple reflection of reality" (Collins 2000, 15).

In pursuing this goal, this short paper takes a first step in bringing together two broad bodies of literature: that on the history of intelligence testing, eugenics, scientific racism and colonialism with the burgeoning one on the ethics and societal impact of AI. It begins by briefly laying out relevant aspects of the value-laden usage of the concept of intelligence. It then explores five ways in which this might be influencing thought about AI, ranging from the impact on diversity in the technology sector, to the misdirection of discourse on the risks of AI towards the concerns of those traditionally at the top of the dominance hierarchy and away from those already marginalised.

Intelligence as a Value-Laden Concept: An Overview

The specific aspect of value-ladenness that I will explore in this paper concerns the use of intelligence (and related concepts) to justify social, political and economic hierarchies, both within and across societies. I will argue that the concept of intelligence has played a crucial role in what Patricia Hill Collins calls a "matrix of domination" that has historically preserved the power of a White, male elite. This matrix is formed by intersecting "structural, disciplinary, hegemonic, and interpersonal domains of power" (Collins 2000, 18). Concepts such as intelligence are deployed as part of the "hegemonic ideologies" of superiority and inferiority that legitimate other parts of the matrix (Collins 2000, 284), where by 'hegemonic', Collins means they are "seen as natural, normal, and inevitable" (Collins 2000, 5).

In briefly exploring the ideological usage of the concept of intelligence, I will sometimes refer to related concepts such as 'mental ability' or 'faculty of reason'. To be related in the relevant sense, a concept must fulfil two criteria. First, it must refer to higher cognitive faculties. Second, it must refer (whether usually or in the relevant context) to a faculty that admits of degree. One of the crucial historic shifts in thinking about cognitive faculties was away from the idea of a faculty that all humans possess equally and wholly, at least at birth (such as 'mind') towards the idea of a faculty that some people possess more than others (Gonzalez 1979, 45). Although the rise of the word 'intelligence' in the nineteenth and twentieth centuries is entwined with this shift, earlier writers at times used cognate terms in ways that also imply this gradation.

The Origins of Intelligence as Ideology

The association between mental faculties and a right to power is an ancient one. Famously, in the *Republic*, Plato proposed (through the mouth of Socrates) that the ideal ruler of the ideal state would be the 'philosopher king', he "who has the gift of a good memory, and is quick to learn, — noble, gracious, the friend of truth" (Plato 2008, bk. VI, 487a). This was a novel idea at the time, competing with ideas of democracy, monarchy, inherited aristocracy, tyranny and others. Yet only one generation later, his pupil Aristotle used gradations of mental aptitude to justify a sociopolitical hierarchy that covered both different groups of humans and the rest of the natural world. In the *Politics*, he argues "that some should rule and others be ruled is a thing not only necessary, but expedient; from the hour of their birth, some are marked out for subjection, others for rule" (Aristotle 1905, bk. 1, V). What marks out the rulers is possession of reason; while one who lacks it "is a slave by nature". Those men who use their mind have this more than those who work with their bodies; men have it in general more than women; and humans more than other animals. In all cases, it is better for the lower sorts that they should be under the rule of a master (Aristotle 1905, bk. 1, V).

We therefore see, at the dawn of Western philosophy, an identification of mental aptitude with the privileged, male human, and an argument for his right, because of this aptitude, to dominate over other groups. It is noteworthy that, whereas Plato was recommending a radical new politics, Aristotle was effectively defending the power relations that were then the status quo. What he presents is, in Collins' terms, a hegemonic ideology, one that naturalises and normalises these power relations, and does so based on claims about mental aptitude.

However, this view was far from predominant in the subsequent two thousand years. The enormous influence of Christianity even in its Platonised, Augustinian form, meant that other virtues, such as piety, were valued over mental aptitude. At the same time, sovereignty was justified more by divine or inherited right than by intellectual superiority.

For a number of reasons, this began to change in Europe in the seventeenth century. Some of these reasons have to do with the rapid intellectual developments of this time, including the rise of the rationalist project, empiricism and modern science. But the most important developments were political.

The first part of this was the republican movement's drive to replace principles of hereditary rule with, in Thomas Jefferson's words, "a natural aristocracy" based on "virtue and talents" (Carson 2006, 11). The second was the need to provide a moral and intellectual justification for aggressive colonial expansion from Europe, with its associated conquest, pillage and enslavement (Adas 1989, 199). In this context, the argument offered by Aristotle, that some people because of their superior intellect were born to rule and others less gifted born to be ruled, was hard to resist. We therefore see in this period the construction of race "to create and maintain distinctions between different members of the Homo sapiens species that lend a suprahuman explanatory ground… to these hierarchies" (Weheliye 2014, 28). Foremost among the criteria of difference was mental ability (Carson 2006, 76).

The implications of placement on this hierarchy were of the utmost consequence: because those at the bottom of the ladder were deemed mentally inferior — in the words of Rudyard Kipling, "Half-devil and half-child" — they were judged unqualified to rule over themselves and their lands. It was therefore perfectly legitimate — even a duty, "the white man's burden" as Kipling put it — to destroy their cultures and take their territories (Kipling 1899). According to the historian Michael Adas (here quoting the influential Victorian cleric and author Frederic Farrar) this perceived backwardness 'explained and justified the decimation or (in the case of the Tasmanians) the utter extermination of "primitive peoples" who had not "added one iota to the knowledge, the arts, the sciences, the manufactures, the morals of the world"' (Adas 1989, 204).

This association of hierarchies of mental aptitude with racism and colonialism is further entangled with a dichotomous conception of civilisation versus nature, and the role of technology in mediating that dichotomy. For colonial European powers, superior science and technology were not only the means for conquest, but part of its justification, as they demonstrated the superiority of their intellect and culture (Adas 1989, chap. 4). The very purpose of science and technology was domination over nature and its subjection to human needs (Leiss 1972). This in turn equated to cultural success: "human civilisation was virtually synonymous with the conquest of nature" (Thomas 1984, 25). Inhabiting the opposite of this civilised state were those peoples deemed inferior in intellect and correspondingly pre-technological, and who were therefore savages. The *mission civilisatrice* of the Western powers was to use their superior wits and technology to tame this wild nature — both lands and people — and put it to productive use. The matrix of domination therefore

deployed conceptions of intellectual rank to legitimate exploitation of both nature and those 'races' considered to be in a state of nature; while this very technological capacity to exploit was in turn part of the proof of intellectual superiority.

It is hard to find a European or North American thinker of the 18th and 19th centuries who, if opining on race, did not assert the intellectual superiority of White people (Gould 1981, 32; Carson 2006, 88). The only real debate was whether this was cultural or biological. The idea that it was biological was politically more attractive: it implied that Whites were innately superior, and other races would not be able to challenge this superiority even if given access to similar cultural resources. However, the age of empiricism demanded numbers to support these claims (Gould 1981, 74). Consequently, we see in this period the development of what is sometimes called racial science, or scientific racism. Although it is important to remember that this logic was applied not only to different races, but also to gender and class. We should therefore think of it also as scientific sexism and classism, often operating in conjunction.

The first sustained attempt was in the form of craniometry — measuring the size of the skull. This was taken to be a proxy for mental ability and so was supposed to provide objective grounds for the well-established racial and social rankings. By the mid-nineteenth century, this enterprise had collapsed under the weight of the facts (Gould 1981, chap. 3). Nonetheless, these efforts drove the ideological and scientific imperatives to coalesce and "make intelligence seem a singular, real, measurable, physical entity" (Carson 2006, 78). The ground was therefore prepared for the first systematic attempt to directly measure intellectual ability — by the English scientist Sir Francis Galton. His best-known work was the 1869 book *Hereditary Genius*, in which he set out to demonstrate that mental ability was hereditary, and that it determined people's success in life.

Galton was also the inventor of the term 'eugenics', "the science of improving stock… to give to the more suitable races or strains of blood a better chance of prevailing speedily over the less suitable" (Galton 1883). He believed that the poor and marginalised were poor and marginalised because they were intellectually inferior. They should therefore for the good of 'the race' be discouraged from breeding. His work gave expression to the mood of the privileged classes in Britain and around the world (Levine 2017, 12). The term eugenics became the name of a worldwide movement, with intelligence testing the primary tool for determining who was fit to pass on their genes and who was not. As historian Philippa Levine puts it: "nothing was more important for eugenicists than intelligence… intelligence was the key variable" (Levine 2017, 25). Until this time, the late nineteenth century, the word 'intelligence' had "remained largely in the backwaters of English-language discourse" (Carson 2006, 79). When it began to rise to widespread usage, it was as part of the eugenicist

ideologies of White supremacy, colonialism, classism and sexism.

The Science and Politics of Intelligence in the Twentieth Century

The first test to resemble modern IQ tests (in being a battery of diverse short activities designed to quickly assess different aspects of reasoning) was developed by the French psychologist Alfred Binet in 1905. It created the notion of 'mental' age in order to determine which children were so behind that they should be given special education. But it was not designed for fine gradations or rankings. This changed when the test made it across the Atlantic and was picked up by a group of influential American psychologists who were also active members of the eugenics movement. Through their efforts, within two decades the idea of a hierarchical, unilinear intelligence "had become a term of central importance within American psychology and, to a certain degree, American culture" (Carson 2006, 159).

A landmark in this development was the publication in 1916 by Stanford psychologist Lewis Terman of a revised and expanded version of Binet's test. Known as the Stanford-Binet, its updated version is still in use today. Thereafter, intelligence became associated with IQ (a term coined by German psychologist William Stern and developed by Terman), "understood as innate, quantifiable mental ability" (Carson 2006, 183). Like others in this group, Terman was deeply sexist, classist and racist, and believed that the tests would demonstrate that the established order, in which White, educated men formed the elite, was right and just. He concluded from his studies "that the whole question of racial differences in mental traits will have to be taken up anew and by experimental methods. The writer predicts that when this is done there will be discovered enormously significant racial differences in general intelligence" (Terman 1916, 92). He became a leading member of the American Eugenic Organization, anxious that those he deemed inferior "constitute a grave problem because of their unusually prolific breeding."

Another member of this group was Carl Brigham at Princeton, who created the first SAT (which originally stood for Scholastic Aptitude Test) to test pupils for college-readiness. He was one of the psychologists involved in the group's great coup: the administration of IQ tests to US Army recruits during World War I. Over the course of the war, 1.75 million people were tested. Brigham used this data to write his highly influential 1923 book *A Study of American Intelligence*. He concluded: "It is also possible to make a picture of the elements now entering into American intelligence. At one extreme we have the distribution of the Nordic race group. At the other extreme we have the American negro. Between the Nordic and the negro, but closer to the negro than to the Nordic, we find the Alpine and Mediterranean types" (Brigham 1923, 196). When he developed the SAT, it was not to ensure all peoples,

regardless of race or gender, had an equal chance of college entry, but to ensure that the Ivy League Universities remained predominantly White in the light of increasing immigration.

Later in life, Brigham recanted, admitting his entire analysis of inherited racial difference was baseless (Gould 1981, 232–33). But by then his work had contributed to the rise of the eugenics movement, to the perpetuation of racial segregation, and to much stricter immigration laws — which for example severely curtailed the number of eastern and central European Jews who were accepted into the US just as Nazism was gaining its hold. As Nazi Germany took racial science and eugenics to new extremes, it correspondingly fell out of favour in the mainstream in the US and UK. But this ideology did not go away, and has resurged periodically throughout the post war period, for example, in the controversy around the 1994 book *The Bell Curve: Intelligence and Class Structure in American Life* (Herrnstein and Murray 1994). The association between intellectual superiority, right to rule, and the White male continues to be perpetuated explicitly in right-wing circles today (Saini 2019).

The history of the science of intelligence is therefore the history of "the abstraction of intelligence as a single entity, its location within the brain, its quantification as one number for each individual, and the use of these numbers to rank people in a single series of worthiness, invariably to find that oppressed and disadvantaged groups—races, classes, or sexes—are innately inferior and deserve their status" (Gould 1981, 25). Of course, the historic use of the concept of intelligence by some to further a particular ideology does not impugn the entire psychometric enterprise. I am not in this paper taking a stance one way or another on the complex and contested science of general intelligence. But I hope to have made the case that, regardless of the validity of the science, intelligence is a highly value-laden term that has been implicated in a matrix of domination since its rapid rise to widespread usage just over a century ago.

Implications of the Value-Ladenness of Intelligence for AI Ethics

In the remainder of this paper, I will suggest some ways in which this value-laden legacy of intelligence might be interacting with both the development of AI and the debate around its associated risks and opportunities. I will note five lines of enquiry (there are surely many more): 1) how some aspects of the AI debate perpetuate the fetishization of intelligence; 2) how the fetishization of intelligence impacts on diversity in the technology industry; 3) how certain hopes for AI perpetuate notions of technology and the mastery of nature; 4) how the association of intelligence with the professional class misdirects concerns about AI; and 5) how the equation of intelligence and dominance fosters fears of superintelligence.

My assumption in making these links is that the long, value-laden history of the concept of intelligence, with its associated biases, continues to inform beliefs today. As noted above, in some circles, racialized and gendered ideas of an intelligence hierarchy are explicitly propounded (Saini 2019). Pertinent to AI, it is worth noting that the racialisation of intelligence has historically been particularly stark in the US, home of Silicon Valley, "because of the centrality of race in American culture" (Carson 2006, 4). But even where such claims are not made explicitly, implicit bias is widespread (Doris 2018). Such bias includes perceptions about innate intelligence: associations between different racial groups and different degrees of intelligence have been found in groups ranging from US college students to Dutch high school teachers (van den Bergh et al. 2010; Okeke et al. 2009; Amodio and Devine 2006). While the most egregious abuses might lie behind us, the concept of intelligence remains far from decolonised.

1) Claims about AI, and the Fetishization of Intelligence

The first point to note is that some of the debate around AI and its potential impact uncritically assumes the central importance of intelligence in the human story. For example, an influential 2014 newspaper article by Stephen Hawking, Max Tegmark, Stuart Russell, and Frank Wilczek argued that:

> The potential benefits [of AI] are huge, since everything that civilization has to offer is a product of human intelligence; we cannot predict what we might achieve when this intelligence is magnified by the tools AI may provide, but the eradication of disease and poverty are not unfathomable.

The article goes on to highlight the potential dangers, arguing "success in creating AI would be the biggest event in human history… [but] might also be the last, unless we learn how to avoid the risks" (Hawking et al. 2014). This claim that intelligence is the root of "everything that civilization has to offer" was repeated in an 'open letter' prepared by the Future of Life Institute, of which Max Tegmark is the director. That open letter has been signed by some 8,000 people, including many famous AI researchers and technologists, such as Demis Hassabis and Elon Musk (Tegmark 2015).

However, this claim for the primacy and centrality of intelligence is highly contentious. Indeed, at most points in the last two thousand years, it would very likely have been considered patently false. For example, in the eighteenth century, "industriousness and frugality" were considered paramount qualities (Gonzalez 1979, 51), while into the nineteenth, the range of talents considered central to civilisation were "multiple and diverse" (Carson 2006, 3), with as much emphasis on moral virtues as intellectual ones. Even Francis Galton believed that the triad of mental ability

with zeal and hard work were the key to human achievement. Today, mainstream psychometricians do not argue that intelligence encompasses everything that matters in human society. Ian Deary writes in his introduction to the topic that intelligence tests "do not measure creativity or wisdom… personality, social adroitness, leadership, charisma, cool-headedness, altruism, or many other things that we value" (Deary 2001, 16). Unless we define 'intelligence' so widely as to encompass all human talents, it therefore seems false to claim that "everything that civilization has to offer is a product of" it.

More importantly, in the light of the ideological baggage carried by the term, this fetishization of intelligence is potentially harmful, foregrounding as it does a quality historically associated with (and consistently deployed to benefit) one particular privileged demographic. Needless to say, the dominant voices in debates about the risks and opportunities of AI, such as the authors of the article above, have largely belonged to exactly that demographic. It is therefore unsurprising that the term intelligence has been uncritically promoted, and the extent to which it is part of a matrix of domination correspondingly obscured.

2) The Fetishization of Intelligence, and Diversity in the Technology Sector

This leads to the second point. The fetishization of intelligence and related concepts such as brilliance or genius) impacts on some sectors more than others. A recent study in the US showed that practitioners in some disciplines, including computer science, emphasise the importance of 'brilliance' more than practitioners in other fields, such as psychology (Leslie et al. 2015). Those fields that emphasise the importance of pure brilliance over other attributes have fewer women and African Americans. The authors argue that this arises from the "pervasive cultural associations linking men but not women with raw intellectual talent" — associations whose history we explored in the last section. There are a number of mechanisms by which these associations might have impact: people in these fields might exhibit biases against women; women might themselves be vulnerable to stereotype threat; or women might simply decide that these fields are not for them. The authors demonstrate that these same arguments apply to African Americans.

While that study looks at computer science as an academic discipline, we have reason to think that the harm caused by the fetishization of intellectual brilliance extends to the commercial high-technology sector. The proportion of women doing technology-related jobs in the leading AI companies is notoriously low: ranging from just under one in three of tech-related employees at Netflix to one in five at Google and Microsoft, and even lower at Twitter and Uber (Richter 2019). At the same time, there is a strong myth of meritocracy in Silicon Valley as a community that values only the quality of the coding, not background, privilege or connections (Thompson 2019, chap. 6). However, as the study by Leslie et al mentioned above shows, when a community believes itself to value only intellectual brilliance, it becomes susceptible to bias in perceptions of what kind of person possesses such brilliance. A 2010 study, "The Paradox of Meritocracy in Organizations," found that in cultures that espouse meritocracy, managers "show greater bias in favor of men over equally performing women." The authors even suggest that "ironically, working in an environment that highlights meritocracy might make individuals believe that they are fair and objective, and as a result, make them more likely to display their biases" (Castilla and Benard 2010).

These studies show that although the more explicit uses of the concept of intelligence to establish dominance hierarchies are confined to the last century, the underlying ideology persists. Different groups are perceived to have different degrees of intelligence, and these perceptions determine their access to positions of influence, including in the field of AI. In response to a highly controversial 2017 memo in which a Google employee argued that women are not as suited to coding as men because they were less systematic thinkers and more neurotic, Cynthia Lee, a lecturer in Computer Science at Stanford, described "the background of endless skepticism that every woman in tech faces" (Lee 2017). We can hypothesise that the current hype around AI, and well-publicised proclamations of the paramount importance of intelligence by senior members of the field, will perpetuate this cycle of exclusion.

3) AI and the Mastery of Nature

The remaining three points all relate to how the value-ladenness of the concept of intelligence might be shaping — or distorting — perceptions of the impact of AI.

First, the opportunities. We can ask how the centuries-old association between intelligence, technology and the mastery of nature could be affecting perceptions of AI in this moment of environmental crisis. The open letter quoted above evidences this association in the claim that "we cannot predict what we might achieve when this intelligence is magnified by the tools AI may provide, but the eradication of disease and poverty are not unfathomable." Eradicating disease and poverty are of course two of the main motivations for mastering nature, and both would require very substantial manipulation of our bodies and the natural world. Similarly, AI is sometimes portrayed as the answer to global warming: Demis Hassabis, founder of DeepMind, for example, has spoken of his goal to "solve intelligence, and then use that to solve everything else," including "climate change" and "energy" (Burton-Hill 2016).

But the presentation of AI as the solution to our problems ignores a large body of literature, going back at least to Adorno and Horkheimer's *Dialectic of Enlightenment*, that blames the will to technological mastery for our current environmental crisis (Horkheimer and Adorno 1972 [1944]). Critics in this tradition would argue that the belief that we can simply build better tools to master an increasingly rebellious nature is symptomatic of the kind of thinking that caused the problem in the first place. Indeed, the reality is that AI is currently contributing to the climate crisis — one recent study showed that training a single neural network model produced 300,000 kilograms of carbon dioxide emissions (Strubell, Ganesh, and McCallum 2019; quoted in Dobbe and Whittaker 2019). The notion that AI will further our mastery of nature and so permit us to engineer a solution to climate change could be seen as a palliative thought that undermines motivation for what really needs to be done, which is to change the attitudes and behaviours that underpin unbridled consumption (a challenge sometimes known as 'moral hazard', see The Royal Society 2009, sec. 4.3). While some argue that the goal of dominating nature is intrinsic to the technological project in general (Adas 1989), we can hypothesise that the close association of the concept of intelligence with this project means that a technology pitched as intelligent ('AI') will fuel this particular approach to the natural world.

4) AI and the Threat to Middle Class Jobs

Now to the fears. It is a common trope in discourse on the impact of AI that it will engender a new wave of automation which, unlike previous waves, will threaten middle class jobs. Martin Ford, for example, in his 2015 book *Rise of the Robots: Technology and the Threat of Mass Unemployment* has a chapter on 'White-Collar Jobs at Risk' and discusses the consequences of AI and related technologies for "many skilled professionals—including lawyers, journalists, scientists, and pharmacists" (Ford 2015, xv).

Taking account of the value-laden history of intelligence can help us to understand these fears. If developments in computer technology are being portrayed as *intelligent*, and intelligence has long been considered the primary marker of the professional elite, then it is natural to conclude that these intelligent technologies will take their place in that elite, displacing the incumbents. Hence concern is directed to the fate of those incumbents (e.g., 'Will A.I. Put Lawyers Out of Business?' Sahota 2019). I do not want to deny that such concerns could be partially legitimate; that "increasingly intelligent algorithms threaten higher-skill occupations" (Ford 2015, 59). But we can ask whether the concern for the existing professional elite is proportionate. Much research shows that it is much more likely to be those who are already marginalised who will be most affected by AI and data-driven technologies (O'Neil 2016; Eubanks 2017; Noble 2018; Benjamin 2019). We should therefore be wary of attention given to the impact of AI on unemployed lawyers as opposed to "people of colour, migrants, unpopular religious groups, sexual minorities, the poor, and other oppressed and exploited populations" (Eubanks 2017, 6), who are not rated highly by the ideology of intelligence, but will be most targeted by intelligent machines.

5) Superintelligence and the Dominance Hierarchy

Fifth: from *R.U.R.*, Karel Čapek's 1921 play which launched the term 'robot', through *2001: A Space Odyssey* to the inexhaustible *Terminator* film franchise, we in the West have long fantasised about intelligent machines rising up against us. When we reflect on how the idea of intelligence has been used to legitimate dominance hierarchies since ancient Greece, and in particular in recent centuries, this fear becomes easier to understand. As we saw, Aristotle's argument, enthusiastically adapted for the logics of colonialism and patriarchy and class privilege, was that the intellectually superior are by nature masters, and the intellectually inferior by nature slaves. With this ideology forming part of the Western cultural backdrop, it is understandable that we fear intelligent machines will enslave us. (Indeed, H.G. Wells makes this connection, with aliens instead of AI, in *War of the Worlds* (Wells 1898)).

This perspective also helps us to understand a phenomenon identified by Kate Crawford: that "currently the loudest voices debating the potential dangers of superintelligence are affluent white men" (Crawford 2016). It is exactly this group whose privileged position has for centuries been justified by their claim to superior intelligence. They therefore have the most to lose by the appearance of new entities that claim to be even more intelligent. Indeed, one might say powerful AI would create for them a dilemma: either cede their privileged position to the super-smart machines, or abrogate the ideology of intelligence on which their privilege is based. Others outside this elite, on the other hand, might be less concerned about such creations, as they are already being oppressed by those purporting to be superior beings.

Conclusion

At the outset, I argued that insufficient attention has been given by those thinking about AI to the extent to which the field's foundational concept of intelligence is 'thick' or value-laden. Albeit briefly, I hope to have brought to the fore the central role that intelligence has historically played in the logics of colonialism, racism and patriarchy. Of course, other discourses have also been conscripted into these agendas (e.g., that of the civilised and the savage). But intelligence is notable both for having played such a central role in these matrices of domination and because its rise to widespread usage is so entangled with them. I do not make the quixotic suggestion that the term should therefore be abandoned. Rather, I hope to have shown that critical evaluation of the value-ladenness of intelligence can and should inform debates about the ethics and impact of AI[1].

References

Adas, Michael. 1989. *Machines as the Measure of Men: Science, Technology, and Ideologies of Western Dominance*. Ithaca: Cornell University Press.

Amodio, David W., and Patricia G. Devine. 2006. Stereotyping and Evaluation in Implicit Race Bias: Evidence for Independent Constructs and Unique Effects on Behavior. *Journal of Personality and Social Psychology* 91 (4): 652–61.

Aristotle. 1905. *Aristotle's Politics*. Trans. Benjamin Jowett. Oxford Translation Series. Oxford: Clarendon Press.

Baisini, Claudia. 2015. *Facts Are Facts. Intelligence Objectively Measured*. MPhil Thesis, University of Cambridge.

Benjamin, Ruha. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code*. Medford, MA: Polity.

Bergh, Linda van den, Eddie Denessen, Lisette Hornstra, Marinus Voeten, and Rob W. Holland. 2010. The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. *American Educational Research Journal* 47 (2): 497–527.

Blackburn, Simon. 2013. *Disentangling Disentangling*. Edited by Simon Kirchin. Oxford: Oxford University Press.

Brigham, Carl Campbell. 1923. *A Study of American Intelligence*. Princeton: Princeton University Press.

Carson, John. 2006. *The Measure of Merit: Talents, Intelligence, and Inequality in the French and American Republics, 1750-1940*. Princeton: Princeton University Press.

Castilla, Emilio J., and Stephen Benard. 2010. The Paradox of Meritocracy in Organizations. *Administrative Science Quarterly* 55 (4): 543–676.

Collins, Patricia Hill. 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment, Revised Edition*. New York: Routledge.

Crawford, Kate. 2016. Artificial Intelligence's White Guy Problem. *The New York Times*, 25 June 2016.

Deary, Ian J. 2001. *Intelligence: A Very Short Introduction*. Oxford: Oxford University Press.

Dobbe, Roel, and Meredith Whittaker. 2019. AI and Climate Change: How They're Connected, and What We Can Do about It. *Medium*. 17 October 2019.

Doris, Keith Payne, Laura Niemi, John M. 2018. How to Think about "Implicit Bias". *Scientific American*, March 2018.

Eubanks, Virginia. 2017. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. NY: St. Martin's Press.

Ford, Martin. 2015. *The Rise of the Robots - Technology and the Threat of Mass Unemployment*. London: Oneworld Publications.

Galton, Francis. 1883. *Inquiries into Human Faculty and Its Development*. London: Macmillan.

Gonzalez, Gilbert. 1979. The Historical Development of the Concept of Intelligence. *Review of Radical Political Economics* 11 (2): 44–54.

Gould, Stephen Jay. 1981. *The Mismeasure of Man*. Rev. and Expanded. New York: Norton.

Hawking, Stephen, Max Tegmark, Stuart Russell, and Frank Wilczek. 2014. Transcending Complacency on Superintelligent Machines. *Huffington Post*. 2014.

Herrnstein, Richard J., and Charles A. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

Horkheimer, Max, and Theodor W Adorno. 1972. *Dialectic of Enlightenment*. Trans. J. Cumming. New York: Herder and Herder.

Kipling, Rudyard. 1899. The White Man's Burden. *The Times*, 4 February 1899.

Kirchin, Simon. 2013. *Introduction: Thick and Thin Concepts*. Edited by Simon Kirchin. Oxford: Oxford University Press.

Lee, Cynthia. 2017. James Damore Has Sued Google: His Infamous Memo on Women in Tech Is Still Nonsense. *Vox*. 11 August 2017.

Legg, Shane, and Marcus Hutter. 2007. A Collection of Definitions of Intelligence. *Frontiers in Artificial Intelligence and Applications* 157: 17–24.

Leiss, William. 1972. *The Domination of Nature*. New York: George Braziller.

Levine, Philippa. 2017. *Eugenics: A Very Short Introduction*. Oxford: Oxford University Press.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Okeke, Ndidi A., Lionel C. Howard, Beth Kurtz-Costes, and Stephanie J. Rowley. 2009. Academic Race Stereotypes, Academic Self-Concept, and Racial Centrality in African American Youth. *Journal of Black Psychology* 35 (3): 366–87.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. NY: Crown.

Plato. 2008 [1871]. *Republic*. Trans. Benjamin Jowett. Project Gutenberg.

Richter, Felix. 2019. Female Workers in the Tech Industry: The Tech World Is Still a Man's World. *Statista Infographics*. 8 March 2019.

Sahota, Neil. 2019. Will A.I. Put Lawyers Out of Business? *Forbes*. 9 February 2019.

Saini, Angela. 2019. *Superior: The Return of Race Science*. Boston: Beacon Press.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. 2019. Energy and Policy Considerations for Deep Learning in NLP. *ArXiv* 1906.02243.

Tegmark, Max. 2015. AI Open Letter. Future of Life Institute. 2015. https://futureoflife.org/ai-open-letter/.

Terman, Lewis M. 1916. *The Measurement of Intelligence*. Cambridge, MA: The Riverside Press.

The Royal Society. 2009. *Geoengineering the Climate: Science, Governance and Uncertainty*. RS Policy Document, 10/09. London: The Royal Society.

Thomas, Keith. 1984. *Man and the Natural World: Changing Attitudes in England, 1500 - 1800*. London: Penguin Books.

Thompson, Clive. 2019. *Coders: The Making of a New Tribe and the Remaking of the World*. New York: Penguin Press.

Weheliye, Alexander G. 2014. *Habeas Viscus: Racializing Assemblages, Biopolitics, and Black Feminist Theories of the Human*. Durham, NC: Duke University Press.

Wells, H.G. 1898. *The War of the Worlds*. London: William Heinemann.