

The *Wednesday*

www.thewednesdayoxford.com



Weekly Magazine of the Wednesday Group - Oxford

Editorial

What Has Happened To The Genius?

The romantic age was occupied with the idea of the genius. However, this idea has been eclipsed in the twentieth century and what looked more like supernatural abilities were reduced to more technical capabilities that can be measured with new psychological techniques such as IQ tests. It might be a shock to someone who in the days of Kant and the post-Kantians connected brilliant work with the notion of genius that present philosophy is not concerned with this concept at all. It has been delegated to the sciences, such as biology, psychology and sociology. The metaphysics of the character and role of the genius has been moved from speculative philosophy to technical characterisation by the empirical sciences.

But if the philosophical question has been dropped for being difficult or not so fruitful for some, the empirical data provided by science is more confusing and doesn't give a unified picture as to what genius is. The sciences are more concerned with a set of capabilities and skills. Perhaps this is what we mean when we say someone is a genius. We could mean that the genius has a gift of understanding and can apply ideas faster and in a more skilful way than most. But this has much lower status than the place where Kantian philosophy positioned the genius.

In his short but important book *Genius: A Very Short Introduction*, Andrew Robinson reviews most of the data on genius going back centuries. He thinks the idea in its modern sense began with the Enlightenment. He summarises a definition of the Enlightenment's genius as 'an individual who demonstrates exceptional intellectual or creative powers, whether inborn or acquired (or both)'. It is a good definition. It fits most cases in a variety of fields.

Galton, Darwin's cousin, thought genius runs in families. He was a pioneer in his empirical study of *Hereditary Genius*. But, as Robinson notices, Galton is concerned with talent rather than genius. Psychology

confirmed Galton's findings but did not give it the importance that he gave it. However, genius, unlike talent, is 'the result of a unique configuration of parental genes and personal circumstances.' But 'this configuration never repeats itself in the offspring.'

But will schooling fare better? The answer is negative. The case of Ramanujan was cited. He was self-taught but won the admiration of G.H. Hardy who brought him from India to Cambridge and published joint papers with him. He said of Ramanujan: 'All his results, new and old, right or wrong, had been arrived at by mingled argument, intuition, and induction, of which he was entirely unable to give any coherent account.' Robinson also cites psychological research and statistics which show the decline of genius despite improvements in education, especially higher education, and the number of PhDs. The psychologist Hans Eysenck concludes his book *Genius: The Natural History of Creativity* by saying: 'We probably cannot train it, but we can prevent it from being suffocated by rules, regulations, and envious mediocrity.'

The conclusion may be that genius could be a matter of perspiration and inspiration. This leads to the 'ten-year rule' suggested by psychologist John Hayes thirty years ago which states that a person 'must persevere with learning and practicing a craft or discipline for about ten years before he or she can make a breakthrough.' It may sound a long time but creativity doesn't come easily.

All this and more you can find in Robinson's book but any hope of a philosophical inquiry into the nature and value of the genius is not mentioned. Possibly the common-sense, empirical approach is not fit for such an investigation and we need a more speculative mind and philosophy.

The Editor

ARTIFICIAL
INTELLIGENCE

Ethics in Automation

How much moral weight can we give to technology and automation?

ROB ZINKOV

2 **M**uch discourse centres around the ethics of electronic systems (typically called artificial intelligence). In drawing from science fiction work, we often ascribe to them human traits and aspirations. This further leads to treating them as not too different from working-class people. This includes a fanciful discussion of revolt in the spirit of worker-led revolutions. But this lens is less helpful when the robots stop being allegorical.

In our own societies assigning agency to these machines actually has the effect of taking away agency from the flesh and blood people that not only use these machines but also create them. If a toaster malfunctions we

consider the manufacturer liable. Yet when a self-driving car malfunctions there is a temptation to blame the car. This isn't even a fanciful scenario: when in 2018 the autopilot of a Tesla malfunctioned killing the driver, Tesla chose to blame the driver.

But even if we choose not to assign agency to these systems, automation itself has a peculiar way of shaping ethics. Whenever we use a tool to complete a task we must conform to what the tool allows. A hammer cannot make cuts into a piece of wood, and a saw cannot flatten nails. More poetically, we cannot make the same sounds with a piano as with a fiddle. When we use technology to solve problems with an ethical

component these constraints can come to hurt us. So how are these ethical components manifest? The field of ethics in AI mostly concentrates on such situations. We have software to predict who is likely to commit a crime, we have software to predict prison sentences, along with what neighbourhoods police should patrol. Sadly, this software has been shown to have been trained on the decisions of people and unfortunately people in racist institutions. The net result being that we once had a person shaped by racism, but we now have software shaped by racism. The added danger is that while we expect such people to be flawed in some way, we often portray machines as impartial. This gets worse again as machines can make decisions millions of times faster than any person. It thus becomes essential to understand what it means to be ethical. But just as it becomes so crucial to understand the ethical component of these decisions, we are being divorced from making them directly, depriving us of necessary context.

But what about automation itself? We sometimes muse that machines could eventually do everything. A machine might not just be a tool that helps cut the doors for a car but one that makes the entire car all by itself. One of the ironies of automation is that the more sophisticated the automation the more sophisticated the technician is needed to repair the machine when it breaks. This leads to this bizarre situation where the ideal amount of automation is just enough to make speed up the work, but that the machines are themselves simple to repair. So we are likely to stay indefinitely in this situation where humans co-exist with technology. We need to then think carefully about who is responsible for bad behaviour. This must be considered in combination with the reality that automation in the workplace is often dictated by the needs of the company owners than the workers themselves. With technology imposed on workers implicitly making moral decisions, we have deprived these workers of the ability to behave ethically in the workplace. They can blame the machine, but that does nothing or worse be themselves blamed.

Work that required skill and had notions of seniority and expertise is upturned as the technology changes the nature of the work. This has serious implications for the dignity we derive from our work. Daniel Pink talks about three things being essential for work to be deeply satisfying. It must allow a sense of autonomy, we must provide a sense of mastery, and it must be in service to a higher purpose. How much mastery can one feel over a process introduced yesterday and which will be replaced tomorrow? Every new piece of technology in the workplace has the potential to change how meaningful a job feels. Even something fundamental like the division of labour can take a job that felt meaningful as we saw the final product of our hard work into one that feels meaningless or even bullshit. The choices we have in terms of how we use the technology has substantial moral consequences as work that was once engaging becomes mundane and deadening.

Machines though also have a final thing that they do and have done from the beginning. They change the person who operates the machine. From the sewing looms to the assembly line, technology has also had a history of de-skilling work.

What used to be the unique province of craftsmen and their guilds becomes a thing many more people can do. While this often makes the goods cheaper it also changes the nature of the work. In settings where the machines have an administrative function, they can often set the pace of the work. With the ability of emails to be sent at any time comes an expectation of replying to them at any time. Going further, the technology - if centrally positioned in the workplace - encourages us to treat human beings as machines.

I bring all these up as threads to ponder and to encourage discussion over how much moral weight we give technology and automation, and how much this moral weight is pushed back on us.

Can We Ascribe Agency To Machines?

The group met as usual in the basement of the Opera Cafe, Walton Street, Oxford between 16:00 and 18:00 on a Wednesday. On 11th March Rob Zinkov presented on the topic of how much moral weight we give technology and automation.

CHRIS SEDDON

Concerns to reduce viral transmission and other matters meant only five attended this week, but we all felt that the topic was so productive of significant philosophical issues that we hoped Rob would soon present on this topic again.

At the outset Rob asked us not to focus on science fiction, but on technology with which we are already engaged. He highlighted several related issues: assigning agency to machines can reduce the agency we ascribe to ourselves; our ethical decision-making can be constrained by the automated tools we employ in making those decisions; the difficulty of questioning the authority of an automated decision can make us reluctant to do so; de-skilling work through automation can have moral consequences for our sense of our own and others' worth; these issues combined can impair our skill to make ethical decisions.

Rob opened with the example of a self-driving car malfunctioning and killing the driver, suggesting that this was a scenario in which we were tempted to assign agency to an automated machine by blaming the car. Others agreed that this was a clear example of why agency should never be ascribed to machines, however it was also suggested that instead of talking about agency as if it expressed a clear concept regardless of context, closer inspection of the actual example revealed why those concerned did not in fact ascribe agency to the machine in this particular context. The most urgent question following the fatal accident was posed by the legal representatives of the driver and the manufacturer. Each claimed that

the other was to blame. In the event, it was determined by technical experts that the driver had operated the car as instructed and that the car had malfunctioned. In that sense, the car was to blame. As an obvious consequence the manufacturer was legally liable. This reveals a useful distinction between the etymology of *blame* and *responsibility*, although in normal usage the distinction is not always so clear. The design of the car was to blame for the accident and the driver was not - in the sense that in this case it was the design of the car which *needed to change*, not the driver. Since the car was to blame, the manufacturer was responsible, not the driver - in the sense that it was the manufacturer which *needed to make changes*, not the driver. Everybody eventually agreed that the car needed to be changed, but nobody even thought of making it responsible for making those changes - only the manufacturer could do that. This does not represent a deep philosophical divide between man and machine, it represents the inability of this particular machine to improve the way it responded in this particular situation, or to pay for the damage caused. Responsibility, or in Rob's terms, agency, requires ability. As machines are designed with new abilities, so we justifiably ascribe to them new responsibilities. The limits to these abilities and responsibilities are matters of complexity, not principle.

This discussion revealed divided opinion as to the extent to which mental terms such as *knowledge*, *beliefs*, *objectives*, *behaviour*, and *ability* could be ascribed to machines. At one extreme one speaker seemed to feel that any such ascription was a quaint and amusing metaphor, whilst at the other extreme another



Automation

speaker claimed that they were literally true. Although Rob wanted to focus on automation rather than artificial intelligence, this gave rise to the question of what role the term *Artificial* played in the phrase *Artificial Intelligence*. A distinction based on the relative roles of carbon or silicon was not seriously considered. A distinction based on the role of Darwinian evolution versus design was considered more telling, despite the theological stances of some present. The design of a machine is by definition the specification of its function and operation. In that sense a machine cannot change its own design, so it cannot be responsible for such a change. However as information technology becomes more sophisticated, the design can become more general in the sense of specifying very general goals and very general learning strategies, which, as in the case of modern chess programmes, can make the actions of the machine and their ability to change those actions in the light of the results of previous actions, more efficient at meeting the specified goals in ways that the designers could not have predicted. Just as organic life has evolved simply as a consequence of surviving under prevailing conditions, so too could machines, given enough time and sufficiently flexible design.

In the face of such extreme views, discussion moved on to the effect of the ethical affordance provided by automated decision-making. Rob provided illustrations from military drone targeting, police use of facial recognition, and loan approvals.

Legislation is still only just being discussed to oblige manufactures to build into software used to make important decisions the ability to provide an explanation of the factors and protocols relied upon in a specific recommendation. There seems to be a general tendency in the face of the increasing complexity and ability of machines to trust their recommendations without question. This can affect our ability to trust human authorities, based not, as would be rational, on the credibility of their explanations for their decisions, but instead on who they are. Rob provided the example of a study on gambling which was rejected by many simply because it had been funded by a casino.

The racism inherent in automated facial recognition is not built in by the designers of the software, but by its trainers - that is, such software is more likely to classify those belonging to a racial minority incorrectly as known criminals

Follow Up

than those belonging to the racial majority, simply because the recognition software has been exposed to fewer people in the racial minority. This means that, although the design is adequate for developing ways of differentiating images of faces based on information as to which of several thousand images are of the same person, the rules it has developed are not so good at distinguishing people in the racial minority. This is exactly the same phenomenon as occurs in human police officers - with the best will in the world, it can be harder to differentiate individuals of an unfamiliar race.

Rob provided further examples of the unintended automation of racism, showing how software 'beautifying' his Japanese wife insisted on making her skin paler, and, with perhaps more serious consequences, how sentencing software trained on known judgements still reflected the alleged racial bias of the original judgement based on characteristics correlated with race even though the convict's race was not explicitly input.

Military drones provided an unexpected example of abdicating moral responsibility - unlike staff driving out on the wrong side of the road, none

of the staff at the US Air Force base in nearby Croughton has ever had to claim diplomatic immunity to avoid prosecution for negligent killing by the drones operated via that site.

In response to the question as to whether a machine could ever have an immoral intention, Rob referred to algorithms used in websites to display options to individual users based on options they had previously selected. In the case of pornographic websites, the designers know that those with certain fetishes are more likely to select options compulsively - and thus generate more revenue - and as a consequence weight the algorithms to display more options likely to attract those with such fetishes. In this case the moral responsibility for the preponderance of such options seems to lie with the designers and arguably the users, however it is also possible that a more generic algorithm could create the same preponderance simply as a result of being designed to maximise the number of visits, without any design intended to benefit specifically from such fetishes. In this case we might say that the designers were morally responsible only in that they neglected to predict this unintended consequence. This was seen by some as an example of the beginning of a partial



Military drone



Work in the future

reduction in moral responsibility, which could continue in line with increases in the complexity and power of the automating algorithms.

It was suggested that the temptation to automate complex decision-making by means of a system comparing the disparate outcomes of disparate decisions in terms of comparable units of utility has given rise to an unjustified reliance on moral utilitarianism. It was felt that this simplistic utilitarianism, through its dominance in the automated decision-making to which we are increasingly becoming accustomed, might unduly and unconsciously influence our personal or corporate decision-making, reducing our ability to make more appropriate nuanced moral decisions.

Another moral aspect of automation is the complexity of identifying legal responsibility. Modern manufacturers do not seek to ship software which is completely free of errors - or *bugs*. This is simply because modern software is so complex that this is an unrealistic expectation. Instead the proportion of errors is reduced to a level supposed to be consistent with the purpose of the software. Even so, interdependence between software from different manufacturers makes it in practice impossible to allocate definite responsibility for all significant errors to any one manufacturer - for example, an error in the inter-operation of the operating system, the device drivers, and the application software



Chess master

could cause a fatal accident in some applications, with no one of the manufacturers being solely responsible. However, it was pointed out that this complexity has long been recognised in purely human decision-making - with more liberal thinkers, for example, recognising that society must in many ways share responsibility with the criminal.

As well as depriving skilled workers of justifiable pride when their jobs are either automated or split into the repetitive actions of a single unsatisfying stage of a production line, Rob pointed out how automating jobs can create other moral issues. Sometimes the broader human functions of a role can be lost - for example, if a general medical practitioner were to be replaced by a diagnostic machine. In other situations, if we start treating machines like people, then we are in danger of treating people like machines. Rob gave the example of a more mature user instinctively saying 'thank you' to an automated personal assistant, and the converse example of a younger user brought up on such assistants not realising the value of saying 'thank you' to a human. It is not hard to imagine that as such machines develop more complexity and abilities, that more significant aspects of social interaction might be degraded.

Virus

After dinner we sip Baileys.
You tell me the latest news
about the corona virus infection worldwide
and our lost paradise planet earth, each new week
in an even tighter grip.

I talk about coming spring
and soon warmer days.
Your eyes half-mast, you reach
me another glass.
I tell you to put on your rose-tinted specs,
forget the outside world
and watch our cat doing somersaults,
my heart shining gold.

Let us dance you say, put
your shocking-pink lipstick on,
your silk underwear and high heels.

Cut -
In my head the new pictures. Someone
enters the house, screaming: for each one
you ignore... you will be judged!





Poem and Artwork by Scharlie Meeuws

Adorn



Negatives



CHRIS NORRIS

He who wishes to know the truth about life in its immediacy must scrutinize its estranged form, the objective powers that determine individual existence even in its most hidden recesses.

T.W. Adorno, 'Dedication', in *Minima Moralia*

He who is not malign does not live serenely but with a peculiarly chaste hardness and intolerance. Lacking appropriate objects, his love can scarcely express itself except by hatred for the inappropriate, in which admittedly he comes to resemble what he hates. The bourgeois, however, is tolerant. His love of people as they are stems from his hatred of what they might be.

Adorno, 'Final Serenity', in *Minima Moralia*

Why think – you good-news bearers – why on earth
Think everything's predestined to transpire
As reason has it, plus your tuppence-worth
Of bright-side twittering bred of heart's desire,
When all the signs instruct us there's a dearth
Of negatives, a stink of minds for hire,
About your always heralding the birth
Of brave new worlds although the times are dire.

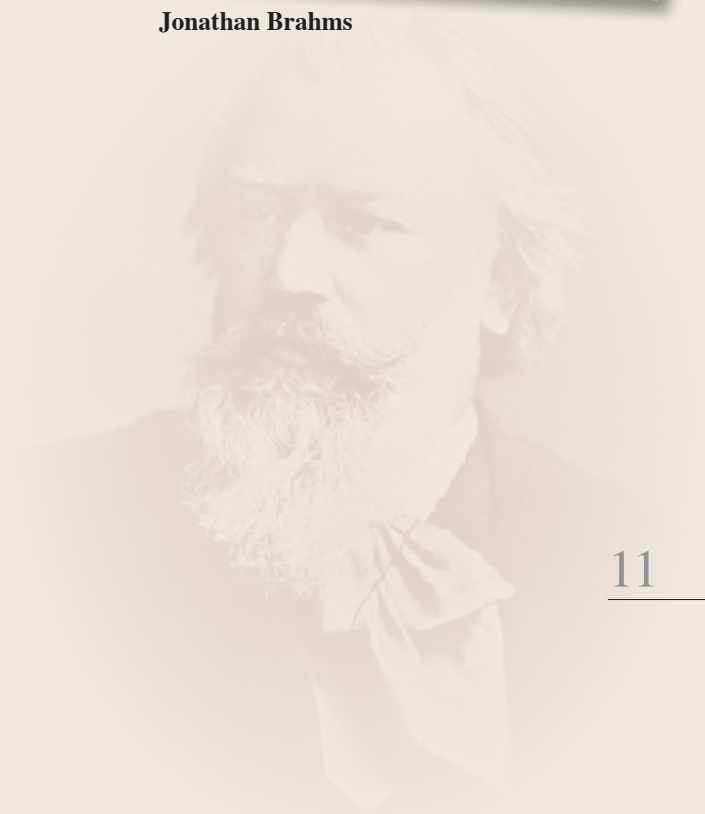
O progeny of Hegel, think before
You hitch your case to that delusive star,
That dialectic primed to fix the score,
Ensure that positives prevail and bar
All thought of negatives, of what's in store
Unless we face the worst and force ajar
The rusty-hinged, the barely outlined door
To what-might-be, shut tight by how-things-are.

It's hopes misplaced that send our lives awry,
False optimism put about by those
With most to gain that has us magnify
The signs of progress, thinking they disclose
The kind of future bliss long peddled by
The culture-industry to all who chose,
Or swallowed outright, its pie-in-the-sky
Fake substitute that led them by the nose.

Try telling them 'It kills your every chance
Of genuine happiness, that mindless pap',
And they'll have answers worked out in advance
By corporate taste-purveyors keen to slap
Down any challenge to their high finance,
Their hefty stake in any means to sap
Our mental powers by having cheap romance
And wish-projection stuff the thinking-gap.



Jonathan Brahm



Poetry

‘Relentless negativity’: they bring
The same old charge against me, but it flies
Back like a boomerang since everything
They say confirms my negative surmise,
My case that it’s those sanguine types who cling
Most tightly to their positives who’ll rise
To any bait the culture-mongers fling
Before them as the losers’ booby-prize.

For only if we think against the lure
Of those false tidings can we hope to gain
Some leverage, some fulcrum-point secure
Enough to take the torsion-bearing strain,
Unleash the negative, and so ensure
Truths glimpsed aslant, perceived against the grain,
Not offered falsely as a total cure
For false totality, the thinker’s bane.



Aldous Huxley



Proust

No prizes for those constantly impelled,
Like me, to seek out flaws, to tap the wheels
Of thought's great locomotive, find what held
The thing on track though every test reveals
Another crack, another system felled
By bad fits, poor materials, sloppy seals,
Or some years-long unnoticed faulty weld
Ignored as thought pursued its high ideals.

Some special positives there are that get
Beneath my guard, disarm the critic, take
Me decades back – like favourite poems set
By Brahms, each one a Proustian keepsake
Of childhood scenes and memories that let
Me, bourgeois as I am, enjoy the ache
Of guilty longing for a past as yet
Untouched by culture's all-corrupting stake.



Brave New World

But then I think: how keep it up, that last-
Ditch alibi of every bourgeois soul,
That spirit-sanctum barricaded fast,
Like bourgeois households, lest its social role
As private-property enforcer cast
An undeceiving spotlight on the whole
Domestic scene that Kierkegaard once passed
Off, comically, as spirit's aureole.

Yet there's a point where dialectic hits
The limit of negation's power to strip
Soul's vestiges away, where thinking quits
That endlessly renewed attempt to clip
Hope's flimsy wings, and where the force of its
Unstinting pessimism may just tip
The scales so thought's long sojourn in the pits
Can find new strength to give despair the slip.

You'll say: 'it's just a cop-out, that retreat
 From your fixed rule so zealously pursued,
 If thereby bound to end in self-defeat,
 To the idea that thinking should conclude
 Its *via negativa* on Hope Street,
 Relying on a momentary mood
 Of chance nostalgic uplift to unseat
 Negation for some fake beatitude'.

I say: without that *promesse de bonheur*,
 That fleeting glimpse of what life yet might be
 In what our own life-memories confer,
 We'd have no means at hand for thought to free
 Itself from hope's unrelenting saboteur,
 The negative that stifles every plea
 For thinking-space by dint of *force majeure*
 And locks it down with hindsight's master-key.

No contradiction here despite what they,
 My critics, deem a case of double-think,
 With 'negative' used mostly to convey
 How thought dismantles falsehoods, link by link,
 But then – just now! – with reference to the way
 That capital and culture join to sink,
 By serially negating, every stray
 Redemptive impulse as life-prospects shrink.

It's bad faith pure and simple on their part,
 That motivated failure to conceive
 How taking thought may lead to taking heart,
 How untruth's negative inversions leave
 A lasting mark, and how life-changes start
 With some snagged thread in error's subtle weave
 Which, as it yields to nay's yet subtler art,
 Shows truths no good-news bearer could retrieve.

The Wednesday

Editor: Dr. Rahim Hassan

Contact Us:

rahimhassan@hotmail.co.uk

Copyright © Rahim Hassan

Website:

www.thewednesdayoxford.com

Published by:

The Wednesday Press, Oxford

Editorial Board

Barbara Vellacott

Paul Cockburn

Chris Seddon

Correspondences & buying The Wednesday books:

c/o The Secretary,
 12, Yarnells Hill,
 Oxford, OX2 9BD

*We have published eight
 cumulative volumes of the
 weekly issues. To obtain
 your copy of anyone of the
 cumulative volumes, please
 send a signed cheque with your
 name and address on the back
 £15 for each volume
 inside the UK*

*or £18 for readers
 outside the UK:*

*Please make your cheque out to
 'The Wednesday Magazine'*

or pay online

Account Number:

24042417

Sort Code:

09-01-29

So Little Trace



After the years of war,
The bombers came no more.
After the blackout's cease,
Willowherb blossomed in the square
The bombing had laid bare.

With sixty years of peace,
The willowherb and empty sky
No longer meet the eye.
Now bankers have no time to spare
For danger from the air.

That was a time when people used to race
To escape the roar
Of bombers drawing near,
Inspiring fear,
And long for the all-clear.

Now traders merely hope to find a place
To have a lunchtime beer
And be refreshed from selling dear.
How could the past so disappear
And leave so little trace?

Edward Greenwood