



**TRANSPARENT MINDS
IN SCIENCE FICTION**

**AN INTRODUCTION TO ALIEN,
AI AND POST-HUMAN
CONSCIOUSNESS**

**PAUL
MATTHEWS**



<https://www.openbookpublishers.com>

©2023 Paul Matthews



This work is licensed under an Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text for non-commercial purposes of the text providing attribution is made to the authors (but not in any way that suggests that they endorse you or your use of the work).

Attribution should include the following information:

Paul Matthews, *Transparent Minds in Science Fiction: An Introduction to Accounts of Alien, AI and Post-Human Consciousness*. Cambridge, UK: Open Book Publishers, 2023.

<https://doi.org/10.11647/OBP.0348>

Further details about CC BY-NC licenses are available at

<http://creativecommons.org/licenses/by-nc/4.0/>

All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at <https://archive.org/web>

Digital material and resources associated with this volume are available at

<https://doi.org/10.11647/OBP.0348#resources>

ISBN Paperback: 978-1-80511-046-0

ISBN Hardback: 978-1-80511-047-7

ISBN Digital (PDF): 978-1-80511-048-4

ISBN Digital ebook (epub): 978-1-80511-049-1

ISBN XML: 978-1-80511-051-4

ISBN HTML: 978-1-80511-052-1

DOI: 10.11647/OBP.0348

Cover image: NASA, Nebula, May 4, 2016. <https://unsplash.com/photos/rTZW4f02zY8>

Cover design: Jeevanjot Kaur Nagpal

6. Supercedure: Into the Posthuman

No individual exists for ever; why should we expect our species to be immortal? Man, said Nietzsche, is a rope stretched between the animal and the superhuman—a rope across the abyss. That will be a noble purpose to have served.¹

How might it be for consciousness to transcend the limits of bodies and minds? A quest for transcendence has been seen as a main force in the creation of much ‘golden age’ science fiction, though this pure and optimistic aspiration was later nuanced and challenged by authors and critics.

While many futurists have hypothesised that human supercedure might come suddenly and irrevocably, others have seen it as a more gradual, subtle process of transformation to a posthuman state. Indeed, authors have noted the futility of demarcating exactly when a recognisable posthuman will emerge. Similarly, while on the one hand authors have foreseen the complete extinction of the human species and its replacement by cyborg or AI intelligences, another approach is to see technology as enabling an acceleration in human mental abilities and a step change in the nature of consciousness.

Stories of future experiential supercedure have several themes. A strong theme has been the decoupling of mind and body—software and hardware—allowing mind uploading into new corporeal form. The enhancement of evolutionarily-bound senses and cognitive power is another, with the accompanying impact on how and where societies come to live, and what they can achieve.

1 Arthur C Clarke, *Profiles of the Future* (London: Pan Books, 1964), 216.

In humans, the net effect of supercedure has been seen to result in either a grand unification or a possibly fatal fragmentation in the species, indeed the risk of crashing through a human ceiling in the ability to operate.

Backup and restore

The software backup-restore metaphor has proven irresistible to a large cohort of SF authors, with the 'natural' logical step of application to encoded human personalities. In a recent example, Dennis Taylor's Bobiverse trilogy shows a human who has his brain frozen after a car crash converted 100 years later into a software program to be uploaded as the controlling AI into a self-replicating interstellar probe. On becoming aware, he agonises over his continuing humanity:

Was I conscious? Could I actually consider myself to be alive? And, was I still Bob?... I thought back to old arguments about Turing tests and thinking machines. Was I nothing more than a Chinese room? Could my entire behaviour be explained as a set of scripted responses to given inputs?²

Bob decides that his inner dialogue invalidates the possibility of being a kind of Chinese room. Also, his inner doubt and its implied concern for the future leads him to conclude that he is a conscious entity.

According to the inventor of the Chinese room idea,³ philosopher John Searle, if a person can have the illusion that they are conscious, then they are conscious. In other words, there is no way that their subjective experience can be denied, being uniquely their own. The normal appearance/reality distinction, whereby we may be mistaken and our view may not reflect reality, simply does not apply to this conscious experience.

2 Dennis Taylor, *We are Legion (We are Bob)*, 2nd ed. (New York: Ethan Ellenberg Literary Agency 2017), chap. 10.

3 The Chinese Room is a thought experiment designed to illustrate how an AI can simply manipulate symbols to appear intelligent, without any awareness of what the symbols mean. A man in a room uses a code book to convert input symbols to the appropriate output. David Cole 'The Chinese Room Argument', in *The Stanford Encyclopedia of Philosophy*, eds Edward N. Zalta and Uri Nodelman, Summer 2023. Metaphysics Research Lab, Stanford University, 2023. <https://plato.stanford.edu/archives/sum2023/entries/chinese-room/>.

Bob's mission in *We are Legion (We are Bob)* includes replicating his probe and installing new versions of himself into them, which in turn leads to existential confusion: 'The process of creating new Bobs would ignite that whole internal debate about who, or what, I was. I would load backups of myself into the new vessels. Would they be me, or would they be someone else?'⁴ Bob reflects on some of the advantages of being a software thing:

One of the advantages of being a software emulation was that I never got tired, never needed rest, never needed to eat or go to the bathroom. My ability to concentrate on a problem had been legendary when I was alive. Now I felt all but invincible.⁵

Another interesting difference between the simulated Bob and his former self is the suppression of extreme emotion through an 'endocrine control switch'. Originally intended to maintain his mental stability, and to cope with his finding himself disembodied, Bob later steels himself to re-engage it:

I flipped the switch to *off*. You know that sinking feeling you get when you suddenly realise you've forgotten something important? Like a combination of a fast elevator and an urge to hurl? It hit me without any warning or build up...my thoughts swirled with a all the things that had been bugging me since I woke up.⁶

In time Bob finds it useful to add or remove this function depending on the circumstances. In some cases, without it he is better at dispassionate decision making. In others, he needs the emotion to ground himself and rediscover his drive and moral compass.

Whilst, as we will see later, there are some cogent objections, mind uploading could be a (far) future technical feasibility,⁷ perhaps occurring in the absence of a fully-formed explanation of consciousness. Such uploading would bring with it considerable ethical challenges, particularly in relation to having multiple, perhaps unlimited versions

4 Ibid., chap. 17.

5 Ibid., chap. 13.

6 Ibid., chap. 13.

7 Though of course all stages of scanning, representing and emulation of mental states still remain unknown challenges.

of a person running at once, along with the kind of use that emulated people might be put to.

One method of brain copying and emulation is presented in Greg Egan's short story 'Learning to be Me'. The brain is eventually replaced by a 'jewel' or embedded computer that has gradually learned the personality of its host. The narrative device is an excellent one for exploring ideas of free will and whether the person after the switch—where the brain is decommissioned—is the same as that before. The narrator has a suspicion that he won't be:

Living neurons, I argued, had for more internal structure than the crude optical switches that served the same function in the jewel's so-called 'neural net'... who knew what the subtleties of biochemistry—the quantum mechanics of the specific molecules involved—contributed to the nature of human consciousness?... Sure, the jewel could pass the fatuous Turing test—but that didn't prove that *being* a jewel felt the same as *being* human.⁸

The protagonist is full of doubt before his final switchover, a doubt also heightened by the possibility of divergence from the human imprint not being detected. But, of course, after the switchover, he is content with his immortality, despite not knowing exactly what it felt like to be his former self.

The Pinocchioesque theme of the inanimate becoming human is found in both Ishiguro's *Klara and the Sun*—where Josie's mother wants to replace her after her death with a synthetic facsimile—and the haunting short story 'Grand Jeté' by Rachel Swirsky. In both cases the motivation is love and the impending loss of a daughter. In 'Grand Jeté' the terminally ill girl Mara faces her replacement with an embodied AI (which her father claims he has created 'for her'). Creating a cultural echo, the story references the ballet *Coppélia*⁹ and Mara wonders at the characters' failing to tell apart the automaton from Swanhilda, despite the dancers' dissimilarity. With similar jealousy to Swanhilda, Mara

8 Greg Egan, 'Learning to be Me', in *Axiomatic* (London: Millennium, 1995), 206.

9 Where the male character Franz falls in love with an automated doll and is fooled into believing that the female lead, Swanhilda, is the doll come to life. The story is influenced by the popularity of automata featured in travelling fairs in the eighteenth and nineteenth centuries.

faces her future replacement and the prospect of having her brain scanned to provide the doll's mind:

She'd read enough biology and psychology to know that, whatever else she was, she was also an epiphenomenon that arose from chemicals and meat and electricity. It was sideways immortality. She would be gone, and she would remain. There and not there. A quantum mechanical soul.¹⁰

Out of love, Mara consents to the scans. The new Mara ('Ruth') is set up and works perfectly, feeling itself to be more real than the old one, having a healthy (as opposed to deteriorating) body: 'But no, her experiences were diverging. Mara wanted the false daughter to vanish. Mara thought Ruth was the false daughter, but Ruth knew she wasn't false at all. She *was* Mara. Or had been.'¹¹

The divergence is subtle at first, but believable as the new Mara faces challenges the other didn't or lacks experiences the other has lived. At the end, we see how this extends to cultural preferences, with the new Mara preferring tragic ballet to comedy—needing to 'feel the ache of grace and sorrow'.

Sensory and cognitive superpowers

The brain's apparent limitations are perhaps why authors have seen a need to leave it behind as the organ of perception and control at some future stage. One such author who attempted such cognitive future gazing was Arthur C Clarke. Inspired in the 1960s by contemporary advances in computing and electronics, Clarke saw no difference between short nerve signals from the fleshy extremities and mechanical or long distance electronic transmission as enhancement to our power to manipulate and to sense:

One can imagine a time when men who still inhabit organic bodies are regarded with pity by those who have passed on to an infinitely richer model of existence, capable of throwing their consciousness or sphere of

¹⁰ Rachel Swirsky, 'Grand Jeté' (The Great Leap), in *Subterranean*, Summer 2014.

¹¹ Ibid.

attention instantaneously to any point on land, sea or sky where there was a suitable sensing organ.¹²

Both Clarke and Stapledon, quoted below, to some extent foresaw what is becoming increasingly clear about human minds: that we can accommodate new sensory input in a flexible, adaptive way, and that the new input can be experienced subjectively not as a superficial excitation, but as representing external objects in the same way that we perceive the visual scene as things rather than patterns of light on the retina.

Clarke's vision chimes with William Gibson's description of the computationally enhanced protagonist of *Neuromancer*, when he 'flips' and neurally links to cyberspace:

Case's consciousness divided like beads of mercury arcing above an endless beach the color of dark silver clouds. His vision was spherical, as though a single retina lined the inner surface of a globe that contained all things, if all things could be counted.¹³

If Gibson's character could flip in and out of an enhanced mental life, still others have passed a point of no return, replacing and reaching far beyond the grey matter medium. Charles Stross' *Accelerando* is a superb speculative projection of human enhancement and cognitive succession, starting from the almost-real—internet-connected implants—and projecting forward in time, to digital minds connected and meshed to aliens. The early stages illustrate a single generational change to cyborg brain complexes:

She [the posthuman Amber] doesn't have a posterior parietal cortex hacked for extra short-term memory, or an anterior superior temporal gyrus tweaked for superior verbal insight, but she's grown up with neural implants that feel as natural as her lungs or fingers. Half her wetware is running outside her skull on an array of processor nodes hooked into her brain by quantum-entangled communication channels—her own personal metacortex. These kids are mutant youth, burning bright.¹⁴

This is followed by mental uploading, where earthly flesh substrate is finally shed:

12 Clarke, *Profiles*, 199.

13 William Gibson, *Neuromancer* (New York: Ace, 1984), 162.

14 Charles Stross, *Accelerando* (London: Orbit, 2005), 122.

We've [the posthuman civilization] been migrated—while still awake—right out of our own heads using an amazing combination of nanotechnology and electron spin resonance mapping, and we're now running as software in an operating system designed to virtualise multiple physics models and provide a simulation of reality that doesn't let us go mad from sensory deprivation!¹⁵

The software by which humans are superpowered contains the ability to spawn and run millions of software agents to carry out cognitive tasks. This enormous shift in processing power leads to a step-change in society:

The ten billion inhabitants of this radically changed star system remember being human... Some of them still *are* human, untouched by the drive of metaevolution that has replaced blind Darwinian change with a goal-directed teleological progress... But eight out of every ten living humans are included in the phase-change.¹⁶

Similar to those of Stross's enhanced characters, N.K. Jemisin's 'orogenes' of *Broken Earth* series have developed additional brain structures that allow them to sense and trigger seismic events. They can also—given the required skills—manipulate the mysterious obelisks that have been created by a previous, forgotten generation. The obelisks enable a huge amplification of their power, though can be deadly to the operator. The lead character, Essun, makes a connection to Topaz obelisk but is distracted by the sight of an enemy army ('something much closer'):

Without waiting to see if they understand, you plunge into the obelisk... Then you're in the topaz and through it and stretching yourself across the world in a breath. No need to be in the ground when the topaz is in air, is the air; it exists in states of being that transcend solidity, and thus you are capable of transcending, too; you become air. You drift amid the ash clouds and see the Stillness track beneath you in humps of topography and patches of dying forest and threads of roads, all of it grayed over after the long months of the Season... But you are committed; you have connected; the resonance is complete. You launch yourself northward anyhow. And then you stutter to a halt. Because there is something much closer than the equator that draws your attention. It is so shocking that you fall out of alignment with the topaz at once, and you are very lucky.

15 Ibid., 191.

16 Ibid., 206.

There is a struck-glass instant in which you feel the shivering immensity of the obelisk's power and know that you survive only because of fortunate resonances... and then you are gasping and back within yourself and babbling before you quite remember what words mean.¹⁷

In Jemisin's alternative Earth, the orogenes are persecuted for being different despite the powers which might enable them to fix an apocalyptic world where humans face extinction.

While Stross and Jemisin describes new cognitive capabilities or hardware, an alternative take is to imagine a more natural augmentation of mental processing power. In Ted Chiang's short story 'Understand', cognitive superpowers are enabled purely through enhanced brain growth. After an accident, the protagonist Leon Greco receives hormone K therapy to regenerate neurons, but to his doctor's surprise, he eventually surpasses common intelligence tests. Eventually absconding from further trials, he steals his own supply of the agent, and he develops his powers further to a 'supercritical' state.

Leon begins to see patterns, connections and gaps across realms of human knowledge. He leaves behind what he sees to be the constraints of human language, developing pictograms and other techniques for more powerful thought. Crucially, his metacognition, through recursive and self-monitoring sensitivity, is boosted to a hugely more powerful level:

I understand the mechanism of my own thinking. I know precisely how I know, and my understanding is recursive... I know my mind in terms of a language more expressive than any I'd previously imagined... What I can do is perceive the gestalts; I see the mental structures forming, interacting. I see myself thinking.¹⁸

Leon finds that another supercritical person (Reynolds) exists and, through sophisticated clue-leaving, they find each other. Able to share their learning at a deep and profound level, they are nevertheless fundamentally in conflict and different in their conclusion over the purpose of their newfound power. Reynolds is motivated by saving humanity and pragmatism, hoping to start a new, dictatorial movement

17 N.K. Jemisin, *The Obelisk Gate* (London: Orbit, 2016).

18 Ted Chiang, 'Understand', in *Stories of Your Life and Others* (London: Picador, 2014), 65.

for change. Leon, on the other hand, sees aestheticism and the unification of knowledge as the goal.

Echoing much posthuman projection, Chiang highlights the potential for the emergence of a two-tier society, in which the cognitively enhanced have the potential to rule over or manipulate those without access to the technology or who explicitly opt out.

If Chiang's mental superpowers come through recursive intensification and enhanced mental interoception, the mental revolution depicted by Toh Enjoe in his short story 'Overdrive' is framed as a breaking of the speed of thought. Here, thrust created by the 'counting of wild ideas'—a kind of creative/randomlike mathematical reasoning—enables the protagonists to reach, and then transcend, the limits of 'thought space'.¹⁹ This is described as a bridging of space and time and the new capability to travel to other possible worlds:

In that instant, a kind of understanding came to him, and he could sense the next surge of acceleration. He could see forms appearing on the curtain spread out before them... the actual piercing of the curtain was... the hyperthought navigation itself, and its fruits, and all the possibilities that emanated from it, tearing themselves apart to shreds.²⁰

Enjoe's prose makes heavy use of point and counterpoint, of imaginative and metaphoric possibilities and limits, to embody the difficult ideas he attempts to weave. And it seems to work, particularly the idea that at the mental limits, inner concept may become outer travel.

Set against these fictional visions, current cognitive enhancement approaches are somewhat variable in their overall effect and ability to provide lasting improvement in abilities. Whether biochemical, behavioural or physical, the research shows that interventions that boost some areas of cognition have little effect on, or may inhibit, others. The brain's competitive nature means that gains in some abilities or focus may always need to be balanced by losses elsewhere.²¹ Cognitive enhancement effects also vary person to person—as we all vary in

19 Toh Enjoe, 'Overdrive', in *Self-Reference ENGINE*, trans. Terry Gallagher (San Francisco: Haikasoru/ VIZ Media, 2016), 20.

20 Ibid.

21 Lorenza Colzato, Bernhard Hommel, and Christian Beste, 'The Downsides of Cognitive Enhancement', *The Neuroscientist* 27, no. 4 (1 August 2021): 322–30. <https://doi.org/10.1177/1073858420945971>.

hormone and neurotransmitter levels, what works for one person will not work for another. As a society, we also seem more accepting of 'natural' approaches that aid performance within normal limits, but suspicious of technological interventions that may go beyond these.²²

A joining

While enhanced function of single brains could be one cause of a phase change in cognition, as we have seen from the power of hive mind, collective brain power could be another. Stapledon's *Star Maker* is a unique work in providing ever greater visions of cosmic connection and melding. Consciousnesses evolve independently in planets all over the universe, then develop the capability to connect across vast distances. The vision is largely one of expansion of empathy and optimism and the enlargement of the sphere of awareness from the individual to the physics of space:

In these conditions, to be a conscious individual was to enjoy immediately the united sensory impressions of all the races inhabiting a system of worlds. And as the sense-organs of the worlds apprehended not only 'nakedly' but also through artificial instruments of great range and subtlety, the conscious individual perceived not only the structure of hundreds of planets, but also the configuration of the whole system of planets clustered about its sun.²³

Arthur C Clarke's *Childhood's End* has a similar vision of a melded and multiracial consciousness. Human children are purposefully nurtured by The Overlords, an invading alien race, developing their mental powers in order to be ready to join this Overmind:

The Overmind is trying to grow, to extend its powers and its awareness of the universe. By now it must be the sum of many races, and long ago it left the tyranny of matter behind. It is conscious of intelligence, everywhere... this is a transformation of the mind, not of the body. By the standards of evolution it will be cataclysmic—instantaneous. It has

-
- 22 Martin Dresler, Anders Sandberg, Christoph Bublitz, Kathrin Ohla, Carlos Trenado, Aleksandra Mroczko-Wąsowicz, Simone Kühn, and Dimitris Repantis, 'Hacking the Brain: Dimensions of Cognitive Enhancement', *ACS Chemical Neuroscience* 10, no. 3 (20 March 2019): 1137–148. <https://doi.org/10.1021/acscchemneuro.8b00571>.
- 23 Olaf Stapledon, *Star Maker* (London: Penguin Books, 1937), chap. 9.6.

already begun. You must face the fact that yours is the last generation of *Homo sapiens*.²⁴

The last human witnesses the children's ascension to the Overmind via a biblical column of fire. Clarke describes the Overmind itself as having a tangible structure:

A hazy network of lines and bands that keep changing their positions. It's almost as if the stars are tangled in a ghostly spider's web. The whole network is beginning to glow, to pulse with light, exactly as if it were alive. And I suppose it is: or is it something as much beyond life as *that* is above the organic world?²⁵

Clarke hints here at both the unknowability of higher dimensions of being and the loss of our known and trusted operating medium.

Beyond capacity

Could human hardware actually cope with sudden artificial upgrades? It is worth remembering that our knowledge and experience is built up over time and that behaviours such as sleep are sometimes hypothesised to allow the gradual assimilation of new information and selective forgetting.

Towards the end of *The Black Cloud*, Fred Hoyle portrays the human brain's inability to cope with a massive influx of galactic knowledge from the alien cloud, who has proposed a visual, light-based interface to which people can be connected. The first guinea pig to connect to this stream is killed rapidly by brain inflammation. The second lasts a little longer, having tried desperately to develop the ability to assimilate the glut of new information and reconcile it with his current beliefs:

Probably he decided to accept as rule that the new should always supercede the old whenever there was some trouble between them. I watched him for a whole hour systematically going through his ideas along such lines. As the minutes ticked on I thought the battle was won. Then it happened. Perhaps it was some small unexpected conjunction of

24 Arthur C Clarke, *Childhood's End* (London: Pan Books, 1953), chap. 20.

25 *Ibid.*, chap. 24.

thought patterns that took him unaware... He tried desperately to fight it down. But evidently it gained the upper hand—and that was the end.²⁶

Due to these tragic results, the humans are forced to abandon their attempts to learn what the cloud knows. This information brainstorm described by Hoyle has few analogues in humans, though we can draw a few parallels with the—obviously less extreme—psychological effects of religious or cult estrangement, which causes a kind of ontological insecurity or ‘world collapse’ accompanied by depression, dissociation and paranoia.²⁷ The difference here is the absence of a belief network to replace the old, rather than the rapid imposition of mutually incompatible knowledge.

Ted Chiang’s character Leon has similar challenges with his new super cognitive growth, and realises the need for checks and balances in progression:

I must keep a tighter rein over myself. When I’m in control at the metaprogramming level, my mind is perfectly self-repairing; I could restore myself from states that resemble delusion or amnesia. But if I drift too far on the metaprogramming level, my mind might become an unstable structure, and then I would slide into a state beyond mere insanity.²⁸

Despite the success in self-control, Leon does not fully protect himself against a highly sophisticated external attack, a ‘self-destruct trigger’ based on memories implanted by his nemesis Reynolds, and activated simply by associating previously disparate events:

Against my will, a lethal realisation is suggesting itself to me. I’m trying to halt the associations, but these memories can’t be suppressed. The process occurs inexorably, as a consequence of my awareness, and like a man falling from a height, I’m forced to watch.²⁹

Leon’s imposed destruction is not dissimilar to the decline of Charlie in Daniel Keyes’ *Flowers for Algernon*, except that Charlie’s is a natural

26 Fred Hoyle, *The Black Cloud* (London: Heinemann, 1958), ‘News of Departure’,

27 Brooks Marshall, ‘The Disenchanted Self: Anthropological Notes on Existential Distress and Ontological Insecurity Among Ex-Mormons in Utah’, *Culture, Medicine and Psychiatry* 44, no. 2 (June 2020): 193–213. <http://dx.doi.org/10.1007/s11013-019-09646-5>.

28 Chiang, ‘Understand’, 71.

29 *Ibid.*, 83.

conclusion of his momentary boost of cognitive power, a reversion he himself has predicted. The experience is described as similar to the onset of dementia:

June 15: Dr. Strauss came to see me again. I wouldn't open the door and I told him to go away. I want to be left to myself. I have become touchy and irritable. I feel the darkness closing in. It's hard to throw off thoughts of suicide. I keep telling myself how important this introspective journal will be. It's a strange sensation to pick up a book that you've read and enjoyed just a few months ago and discover that you don't remember it.³⁰

Leon's decline is referenced in Richard Powers's *Bewilderment*, where Robbie's brain training gradually wears off, and he again becomes his single personality, unable to deal with the shocks and realities of the world: 'I woke from a nightmare with a tiny hand clamped around my wrist. Robin was standing by my bed. In the dark, I couldn't read him. *Dad. I'm going backwards. I can feel it.*'³¹

In another of Powers' novels, *Galatea 2.2*, he depicts an AI being trained on literature by an academic. The AI, Helen, becomes more and more curious and sophisticated in understanding and interpreting the work it is exposed to. Helen builds representational maps and shows inference and understanding that convinces its trainer that it is conscious. But when he switches its training to news stories—to expose it to the harsher realities of human life—the AI pulls the plug on itself: 'You are the ones who can hear airs. Who can be frightened or encouraged. You can hold things and break them and fix them. I never felt at home here. This is an awful place to be dropped down halfway.'³²

The question raised by tales such as *Algernon* and *Bewilderment* is whether the protagonist is worse off having once experienced the power of a heightened human mind. The question of *Galatea* is whether a possibly conscious AI can fully reconcile the contradictions of the human world. 'Graceful degradation' is how Helen's developer coldly refers to its suicide.

Physical and mental drift

30 Daniel Keyes, *Flowers for Algernon* (London: Gateway, 1966).

31 Richard Powers, *Bewilderment* (New York: W.W. Norton, 2021), 87.

32 Richard Powers, *Galatea 2.2* (New York: Picador USA, 2004).

How far might future humans or human replicants diverge from one another? Is there a risk that burgeoning artificial enhancements might lead to incompatibilities that threaten the coherence of the species?

In Greg Egan's *Diaspora*, the human body has become relatively obsolete, confined to a extinction-bound population of 'fleshers'. Fleshers have experimented with artificial genetic modification, splintering into a number of disconnected subgroups. Exuberants had diverged to extent that they can no longer easily communicate with one another, leading to the emergence of a group dedicated to establishing intermediate forms, dubbed the 'bridgers':

Most exuberants have tried more constructive changes: developing new ways of mapping the physical world into their minds, and adding specialized neural structures to handle the new categories. There are exuberants who can manipulate the most sophisticated, abstract concepts in genetics, meteorology, biochemistry or ecology as intuitively as any static can think about a rock or a plant or an animal with the 'common sense' about those things which comes from a few million years of evolution. And there are others who've simply modified ancestral neural structures to find out how that changes their thinking—who've headed out in search of new possibilities, with no specific goals in mind³³

Citizens (synthetic posthumans) can specialise in their focus and motivation by choosing an 'outlook', a kind of computational locus of control. This helped stave off inevitable degradation due to them having artificial minds:

Most other miners used outlooks to keep themselves focused on their work, gigatau after gigatau [1 gigatau = 11.5 days]. Any citizen with a mind broadly modeled on a flesher's was vulnerable to drift: the decay over time of even the most cherished goals and values. Flexibility was an essential part of the flesher legacy, but... even the most robust personality was liable to unwind into an entropic mess... It was judged far safer for each citizen to be free to choose from a wide variety of outlooks: software that could run inside your exoself and reinforce the qualities you valued most, if and when you felt the need for such an anchor.³⁴

Applying the outlook gives rise to a subtle change in Yatima, the citizen character's mind, conferring a new stability:

³³ Greg Egan, *Diaspora* (London: Millennium, 1997), chap. 3.

³⁴ *Ibid.*, chap. 2.

It was like switching from one gestalt color map to another, and seeing some objects leap out because they'd changed more than the rest. After a moment the effect died down, but Yatima still felt distinctly modified; the equilibrium had shifted in the tug-of-war between all the symbols in vis mind, and the ordinary buzz of consciousness had a slightly different tone to it.³⁵

Drift and divergence have been portrayed in uploaded minds too. In the Bobiverse trilogy, every clone of the original human Bob has different personality characteristics, some more extreme than others. They don't always get on with each other either. This leads to debate around which is the original personality—of course, they claim they each are—and this leads to the divergence:

'Each of us is a bit different. Differences in hardware, quantum effects...' Marvin waved his hand dismissively. 'Invoking quantum effects is just hand-waving. Just means we don't know. I wonder if, as we get older and accumulate memories, we're getting too complex for a backup to contain everything. The backup is a digital attempt to save an analog phenomenon. It may simply be too granular.'³⁶

While Taylor explores the question about mental viability in the initial development and vetting of mind uploads, and also portrays other uploads with psychoses developed while in service, he does not really explore the risk of a new Bob drifting just too far and losing its original personality completely.

Intentional downgrade

Margaret Atwood's *Oryx and Crake* projects a future where humans are all but eliminated in an engineered viral pandemic, but their genetically engineered creations live on in bizarre hybrid forms. The human-replacement race, the 'Crakers', have been designed with limited cognitive powers, deliberately to avoid reintroducing the suffering and introspection to which humans are vulnerable. Reminiscent of the Gammas and Deltas in *Brave New World*, Crakers at first appear to be happy with their lot and their limited range of action and emotion.

35 Ibid.

36 Dennis Taylor, *We are Legion (We are Bob)*, 2nd ed. (New York: Ethan Ellenberg Literary Agency 2017), chap. 39.

But one of the post-pandemic world's last remaining humans, tasked with looking after the Crakers, inadvertently triggers the stirrings of symbolic thinking again in them, by inventing a mythology and creating a mutual dependence:

'Snowman! Snowman!' They touch him gently with their fingertips. 'You are back with us! We made a picture of you, to help us send out our voices to you.' *Watch out for art*, Crake used to say. *As soon as they start doing art, we're in trouble*. Symbolic thinking of any kind would signal downfall, in Crake's view. Next they'd be inventing idols, and funerals, and grave goods, and the afterlife, and sin, and Linear B, and kings, and then slavery and war.³⁷

The novel ends on an ambivalent note regarding hope vs despair. Would the new engineered race continue to develop and thrive, or would they be left to be picked off by the smart feral creatures stalking the forests?

Conclusion—What kind of transformation is possible and desirable?

Alongside Arthur C Clarke, Brian Aldiss was another renowned and prolific author who expounded an essentially optimistic if flexible view of the future. In his short story 'Cognitive Ability and the Lightbulb', long periods of space travel and colonisation are predicted to be a spur to cognitive evolution. The stages of human growth are likened to the power of a lightbulb to illuminate its surroundings:

Early consciousness could be likened to a forty-watt bulb... The First Renaissance marks a shift in brightness to sixty watts. We have now reached the stage of, in Kunzel's terms, the thousand-watt brain. Our offspring are born with an understanding of fractals. This great expansion of cognitive ability led to the new perception of the Universe as a series of contiguities.³⁸

Aldiss was optimistic that this new revolution would lead to an end of war and a period of joy and fulfilment.

37 Margaret Atwood, *Oryx and Crake* (Toronto: McClelland & Stewart, 2003), 'Idol',

38 Brian Aldiss, 'Cognitive Ability and the Lightbulb', *Nature* 403 (January 2000): 253. <https://doi.org/10.1038/35002217>

Of course, not all outlooks on the future are as positive as those of Aldiss and Clarke. For every progression scenario is an imagined regression or evolutionary dead-end. The turn toward pessimism being spurred in part by contemporary challenges to human progression including authoritarianism, climate change, pandemics and ecological breakdown. And even those developments seen by many as part of an optimistic future—mind uploading or backup, artificial general intelligence—are questionable in just how far they might diverge from those cognitive and ethical qualities we might still associate with being human. As Katherine Hayles has pointed out, a view of the posthuman that ignores our embodiment might be fundamentally flawed. She has long emphasised a need to ‘put back into the picture the flesh that continues to be erased is contemporary discussions about cybernetic subjects.’³⁹

Hayles worried that seeing information as a fundamental essence was mistaken when it is purely imaginary and cannot be instantiated without a medium. This finds agreement with the embodied cognitive stance that we saw earlier.

A further concern that is often rightly pointed out by feminist theorists is the domination of visions of posthumanism by existing societal power structures. The concept of posthumanism is often associated with a crass liberal individualism and with mastery over nature, or with long term utilitarianism over current human wellbeing. Furthermore, in fictional imaginings it is often something engineered and purposefully controlled, not allowing for emergence, discovery and nuance. Hayles saw a need to therefore re-route the discussion of posthumanism toward something more ecological.⁴⁰

And transcendence also carries a moral weight as heavy as any psychological one. Too often, it has been seen as a moment of division between those who have the upgrade and those unable or unwilling to get it. The human underclass thus produced is left behind, if often given a noble savage aura. It is therefore easy to understand the caution of some critics about this new path to inequality and elitism.

39 N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago: University of Chicago Press, 1999), 5

40 *Ibid.*

But despite the above caveats and concerns, authors often have their narrators describe the experience of supercedure with joy and euphoria. At its best, we hear a kind of union with the nonhuman, a closer kinship with both the animate and inanimate. New kinds of senses, extending over a large range, giving more empathic insight. Enormously extended and distributed cognitive power conferring wisdom and understanding. These visions continue to provide a possible future in which humanity might save both itself and other species.