TRANSPARENT MINDS IN SCIENCE FICTION

AN INTRODUCTION TO ALIEN, AI AND POST-HUMAN CONSCIOUSNESS

PAUL MATTHEWS

Digital material and resources associated with this volume are available at https://doi.org/10.11647/OBP.0348#resources

Cover image: NASA, Nebula, May 4, 2016. https://unsplash.com/photos/rTZW4f02zY8
Cover design: Jeevanjot Kaur Nagpal

4. The Alien, the Artificial, and the Extended

'If we can't really put ourselves inside the head of another mammal, or even one of our children, then how can we possibly grasp the truly alien?… Yet we will see that a few authors can persuade their readers that they might… some of these stories are far enough away from ordinary human concerns for us to feel that our psychological reflexes have been severely tested'[1]

Can we imagine the sensory and mental world of another being? Here, we will explore some different ways that authors have attempted to describe the experience of aliens, animals, conscious AIs, and human consciousness given new physical, cognitive or sensory capabilities. The clear challenge here for authors has been how to approach this alien depiction by abstracting from human experience. Is it enough to describe novel worlds and minds with common and familiar human language and reference? Or does SF need to take its cue from modernism and make alien accounts idiosyncratic and thereby perhaps oblique and difficult for the reader?

Central here is the 'numinous'—things that cannot be understood rationally or that have a mystical quality. Although often used to describe deep religious experience, it can also apply to the depiction of the alien. Indeed, perhaps the most effective science fiction evocations of the alien[2] express this numinous aspect effectively and are thus sometimes more powerful than anything with more explicit, familiar or relatable descriptions. But at the same time, just these more mundane

---

1   Jack Cohen and Ian Stewart, *What Does a Martian Look Like?: The Science of Extraterrestrial Life*. New ed. (London: Ebury Press, 2004), 56.

2   For example, Stanislaw Lem's *Solaris* (Berkley: Berkley Publishing Company, 1970).

renderings can be used to good effect to portray humans or androids with inhuman powers, alternate cognition and synthetic bodies. In the rest of this chapter, we will see examples that reflect both of these aspects—the numinous and the eerily familiar.

The view from a distance

The word 'numinous' implies extreme or unmeasurable *psychological distance*—how close a new concept is to something we already know.[3] As a species, we have a strong and often useful propensity to imagine things and ideas which are abstracted from our direct experience. 'Construal level theory' has been applied to science fiction and describes how we have a capacity to evoke alternative events, places and realities.[4] The idea describes how we may construe something with very broad-brush details while it is distant, but then gradually fill in details towards more tangible and realistic representations as we come psychologically closer—perhaps through more information or specific instances.[5] Empirical evidence suggests that psychological distance expressed through time, space, social distance or hypotheticality are represented with shared or common structures in the brain and expressed in ways that illustrate their correlation.

While usually not written from the alien point of view, in their characterising of distance, Stanislaw Lem's works are relevant in highlighting human difficulties in psychologically assimilating the truly alien, where construal seems just out of reach. In *Solaris*, the planet under study has a large animated ocean with intriguing properties, but that nevertheless seems outside of agreed description and definition: 'Revered and universally accepted theories foundered; the specialist

3   Liberman, Nira, and Yaacov Trope, 'Traversing Psychological Distance', *Trends in Cognitive Sciences* 18, no. 7 (July 2014): 364–69. https://doi.org/10.1016/j.tics.2014.03.001.

4   J. Carney, 'The Space between Your Ears: Construal Level Theory, Cognitive Science and Science Fiction', in *Cognitive Literary Science: Dialogues between Literature and Cognition*, eds Ed Burke and Emily Troscianko (Oxford: Oxford University Press, 2017).

5   If, for instance, I were to invite you to a tortoise racing party, you might not have been to one before, but your mental scripts for the different components of the event give you an ability to relate quite closely to this (for the moment) hypothetical gathering.

literature was swamped by outrageous and heretical treatises; 'sentient ocean' or 'gravity-controlling colloid'—the debate became a burning issue.'[6]

Lem explores the—perhaps naive, or quasi-religious—idea of contact as a human construct which may actually never be possible due to the sheer difference in physical makeup and nature due to a lack of any shared grounding symbols: 'Contact means the exchange of specific knowledge, ideas, or at least of findings, definite facts. But what if no exchange is possible?'[7]

The humans studying the ocean in this text are all affected by the appearance of a person from their past, seemingly generated by the ocean as a close replica of the remembered person. One crew member hypothesises about the kind of alien perception that has enabled this:

> Yes, isolated psychic processes, enclosed, stifled, encysted—foci smouldering under the ashes of memory. It deciphered them and made use of them, in the same way as one uses a recipe or a blue-print. The ocean has 'read' us by this means, registering the minutest details.[8]

Despite his clear skepticism about the prospect that the heavily blinkered, constrained approaches of science will be able to explain this alien presence, the author appears to settle on a pragmatic view that considers the reality of the humans' experience with the avatars' and the sea's apparent intentionality:

> It was no longer possible to deny the 'psychic' functions of the ocean, no matter how that term might be defined. Certainly it was only too obvious that the ocean had 'noticed' us. The ocean lived, thought and acted.[9]

While maintaining doubt and mystery, Lem thus allows the possibility a high level construal—a form of intentionality and life—to which we might relate.

In his *Southern Reach Trilogy*, Jeff Vandermeer somewhat takes up Lem's numinosity mantle in portraying the geographic and biological anomalies encountered in a region of land dubbed Area X, with a group of scientists struggling to define and comprehend them. Within the

6   Stanislaw Lem, *Solaris* (Berkley: Berkley Publishing Company, 1971), 30.
7   Ibid., 203.
8   Ibid., 106.
9   Ibid., 238.

affected area, animals and plants are changed and corrupted, including a quasi-alien creature dubbed the Crawler which the protagonist encounters at close hand, but cannot properly perceive:

> As I adjusted to the light, the Crawler kept changing at a lightning pace, as if to mock my ability to comprehend it. It was a figure within a series of refracted panes of glass. It was a series of layers in the shape of an archway. It was a great sluglike monster ringed by satellites of even odder creatures. It was a glistening star. My eyes kept glancing off of it as if an optic nerve was not enough.[10]

'The biologist', whose name we never learn, is physically less affected by Area X than other expedition members. But psychologically, her encounters cause a breakdown in her scientific habits of mind and sense of self, to the point where she no longer understands a starfish found in a rock pool, something she has studied her whole life:

> But the longer I stared at it, the less comprehensible the creature became. The more it became something alien to me, and the more I had a sense that I knew nothing at all—about nature, about ecosystems. There was something about my mood and its dark glow that eclipsed sense, that made me see this creature, which had indeed been assigned a place in the taxonomy—catalogued, studied, and described—irreducible down to any of that. And if I kept looking, I knew that ultimately I would have to admit I knew less than nothing about myself as well, whether that was a lie or the truth.[11]

The changes wrought by Area X have elements that make them feel like natural, the kind of biological imitation and accelerated mutation we see in Earth organisms. But their appearance and their effects on the human characters and on the reader are fundamentally strange.

The alien sensory connection

To maintain this half-glimpsed view, authors have imagined the inner worlds of beings with extended communication, expression and even dimensional presence. In China Miéville's 'New Weird' novel *Perdido*

---

10   Jeff Vandermeer, 'Annihilation', in *The Southern Reach Trilogy* (London: Fourth Estate, 2014), chap. 5.

11   Ibid.

Street Station, he describes a pan-dimensional spider-like creature called the Weaver, enlisted by the city's mayor to help with a scourge of vampire-moth creatures sucking the souls from his citizens. The Weavers' origin is obscure, but one theory is that the species has evolved with unusually high speed from normal spiders, to a point where its web has become a way to sense the full fabric of reality across time. Weavers are motivated by aesthetics and will act to bring about more pleasing patterns. The weaver's voice is described as being sensed not in the ears, but directly in the nerves:

> WITHOUT YOU ASK THE WEAVE IS TIGHT RUCKED COLOURS BLEED TEXTURES WEARING THREADS FRAY WHILE I KEEN FUNERAL SONGS FOR SOFT POINTS WHERE WEB SHAPES FLOW... I READ RESONANCE PRANCE FROM POINT TO POINT ON THE WEB... I WILL SNIP FABRICS AND RETIE THEM.[12]

During an attempted arrest of the protagonist for harbouring the creatures, his companion Yargharek (itself an alien) experiences a glimpse of the Weavers' world:

> Every intention, interaction, motivation, every colour, every body, every action and reaction, every piece of physical reality and the thoughts that it engendered, every connection made, every nuanced moment of history and potentiality, every toothache and flagstone, every emotion and birth and banknote, every possible thing ever is woven into that limitless sprawling web. It is without beginning or end. It is complex to a degree that humbles the mind. It is a work of such beauty that my soul wept[13]

Miéville cleverly modulates the intelligibility of the Weaver, like an old-style radio dial tuning over a station. This choice impacts both the cognitive effort expended by his characters as well as the work undertaken by his readers as they grapple with the 'one foot in our world' nature of the creature. Gwyneth Jones' 1991 novel *White Queen* does a similar job of painting the alien as cryptic and hard to predict, not least through a deliberately pixelated narrative exposition. Rather than employing the usual strict dialectic of saviour or conqueror, Jones plays

---

12   China Méiville, *Perdido Street Station* (London: Tor Books, 2008), chap. 28. The Weaver's interconnected web is similar to how possible worlds and wave function collapse has been posited as a way to possibly boost human cognitive powers. We will revisit this in Chapter 5

13   Ibid., chap. 33.

with a more ambivalent/divided attitude of alien colonists (Aleutians) who have arrived on Earth and of the human's attitudes and belief in the aliens too.[14]

The aliens have different personality 'aspects', with the principle alien character, Agnès/Clavel, having those aspects of an artist and poet. Xe (the aliens combine genders) reflects on the effects of displacement and loneliness coupled with the ubiquitous cosmic connectedness (later characterised as God) which xe calls the WorldSelf:

> I came to find the new, but there is nothing new. There is only the WorldSelf, perceiving itself. Any shelter out of which I look is that of my own body. Any leaf is my hand. I cannot escape; I can never leave home.[15]

Jones' aliens reflect a panpsychic conception of a consciousness present all around them. This implies that aliens are really much closer to humans than we want to believe, as everything is made of the same underlying information. This is expressed by Peenemünde, the scientist who helps the protagonists travel to the Aleutian mothership: 'When a thing becomes more complex it does not change, it only becomes *more of itself*. Our awareness is the result: built of the movements of the void, as surely as my hand is built of flesh.'[16]

With their enhanced connectedness to the living, Aleutians fail to understand the human obsession with technology and 'dead' recorded media, and their solution to faster-than-light travel is portrayed as possibly more of a mental than physical power.[17]

Language: invented or developed

Sensory and emotional sophistication is also portrayed in the True AIs in *The Stories of Ibis*. Yamamoto's innovation is to imagine a robot language that evolves rapidly from that of human's eventually surpassing it for efficiency, abstraction and accuracy of communication:

---

14   In this ambivalent mode, the aliens' communication is described as both telepathy and something simpler.

15   Gwyneth Jones, *White Queen*, New ed. (New York: Orb Books, 1994), 20.

16   Ibid., 229.

17   This kind of thought travel is explored further in Chapter 5 where we discuss Toh Enjoe 's short story 'Overdrive', Fujii, Taiyo, Toh EnJoe, and Tobi Hirotaka. *Saiensu Fikushon* 2016. Haikasoru, 2016.

> Our greatest invention was Complex Fuzzy Self-Evaluation... Following words that expressed emotions, subjective reactions, or one's will, we included a complex number as a fuzzy measurement of the intensity of such emotions. This was much more accurate than the additional adjectives humans used.[18]

This invented fuzzy language enables the AIs to distinguish complex shades of emotion using imaginary numbers—such as *Love (5+7i)* and *Love (5-7i)*.

Another approach that authors have taken in depicting the humanly inexpressibility of alien communication is to incorporate gesture. Just as philosophers such as Merleau-Ponty have hypothesised that human language developed from gesture, these portrayals show how different gestural systems might develop in alien cultures.[19] In *The Long Way to a Small, Angry Planet*, Becky Chambers renders a multicultural human-alien Galactic Commons. The ship's crew includes a reptilian Aandrisk, who uses gestures in part to communicate emotion:

> Sissix cupped her palm, flipped it, and spread her claws, even though she knew Rosemary would not understand the gesture. *Tresha*. It was the thankful, humble, vulnerable feeling that came after someone saw a truth in you, something they had discovered just by watching.[20]

Another gesture-using alien race is described in Alan Dean Foster's *Nor Crystal Tears*. The Thranx are an insect-like alien race who come across human space farers. Ryo, the main alien character, begins to teach their gesture-rich language to the humans (who they think of as monsters): ''Bad. Not good.' the monster agreed, making a gesture of fifth-degree and maximum affirmation. Clumsy and unsubtle, Ryo thought, but at least they are learning how to get their thoughts across.'[21]

It is notable that both Chambers' Aandrisk and Foster's Thranx have terrestrial animal analogues, and the extent to which this serves as a crutch to the author or reader is debatable. Other authors have taken

---

18   Hiroshi Yamamoto, 'AI's Story', in *The Stories of Ibis* (San Francisco: VIZ Media, 2010).

19   Thomas Baldwin, ed. *Maurice Merleau-Ponty: Basic Writings* (London: Routledge, 2003). https://doi.org/10.4324/9780203502532.

20   Becky Chambers, *The Long Way to a Small, Angry Planet* (London: Hodder & Stoughton Ltd., 2015).

21   Alan Dean Foster, *Nor Crystal Tears* (New York: Del Rey, 1982), 120.

the approach of simply trying to imagine the life worlds of our own animal cousins. In his short story ''Kjwalll'kje'k'koothaï'lll'kje'k', Roger Zelazny attempts an insight into the consciousness of dolphins, framed by their playful, exuberant nature and their need for daydreaming in the absence of full sleep as we know it. 'Kjwalll'kje'k'koothaïlll'kje'k is described as a great among dolphin-kind, a prophet/seer/musician. Through a telepath who can experience its world, the protagonist is connected to the dolphin's 'dreamsong':

> And then it began again, something like music, yet not, some development of a proposition that could not be verbalized, for its substance was of a stuff that no man possessed or perceived, lying outside the range of human sensory equipment... fresh harmonies into a joyous rhythm I comprehended only obliquely, through the simultaneous sensing of his own pleasure in the act of their formulation. I felt delight in this dance of thought... it was, in and of itself, a sufficiency of being.[22]

The story refers to the theory of (real) historian John Huizinga, that play is a precursor to culture. Appearing as a fictional version of himself, Huizinga thought of play as a way of creating order though free, unfettered application of the imagination. In the story, he reflects on the ancient Roman concept of *ludus*, which connected play with knowledge and learning:

> I swam in a sea that was neither dark nor light, formed nor formless, yet knowing my way, subsumed as it were, within a perpetual act of that thinking we had decided to call *ludus* that was creation, destruction, and sustenance, patterned and infinitely repatterned.. divorced from all phenomena yet containing the essence of time.[23]

Zelazny's creation is powerful in depicting a less language-bound species, able to leverage the flexibility and range of music for thought. He tries hard to approach the numinous, largely through the negation of verbal and conceptual concreteness. The result is at minimum interesting, but at a deeper level might actually explain why human-dolphin communication has proven less than straightforward.

---

22 Roger Zelazny, ''Kjwalll'kje'k'koothaïlll'kje'k', in *An Exaltation of Stars: Transcendental Adventures in Science Fiction*, ed. Terry Carr (New York: Pocket Books, 1974).

23 Ibid.

Alien colonists and spiritual guides

The device of alien colonial masters or observers has been used to shine a spotlight on human consciousness, compared and contrasted to alternate ways of being.[24] In her *Canopus* series, Doris Lessing depicts a similar scenario, with the human race colonised and dictated by alien races who send emissaries linked by (unreliable) telepathic links to their home planets. The Canopeans call earth *Shikasta*. Lessing uses the colonial device to explore the issues of objectivity and subjectivity about consciousness introduced earlier.[25]

In the story, Johor, one of the Canopean representatives, describes the tendency of humans in the twentieth century to fall prey to corruption, dissolution and war, explained in part through the gradual loss of special air that has been sent over from Canopus, conferring the peace and spirituality to humans that Canopeans consider normal:

> It is nearly impossible for people with whole minds—those who have the good fortune to live... with the full benefits of the substance-of-we-feeling— it is nearly impossible, we stress, to understand the mentation of Shikastans.[26]

Phyllis Perrakis has emphasised the dual structure of the novel, with the outer space of Canopus in the first half contrasted with the inner space of the Shikastans in the second half.[27] This is somewhat linked to scientific objectivity (the Canopeans seem detached and removed from the behaviour of their subjects) and subjectivity and spirituality (the Shikastans, or humans, reflect on their own behaviour and morality). Significantly though, the Canopeans get progressively closer to their subjects as the book proceeds and realise that they too are changed by their relationships with humans. This theme recalls the subject/object debate in consciousness theory, and the challenge to scientific objectivity as the 'view from nowhere'.[28]

---

24 Arthur C Clarke's alien masters, for instance, who we will visit in Chapter 5, show an interest in nurturing and protecting the human race for transcendence to a higher level of being.

25 See Introduction, 'Consciousness from the outside and inside',

26 Doris Lessing, *Shikasta* (1994)

27 Phyllis Sternberg Perrakis 'The Marriage of Inner and Outer Space in Doris Lessing's 'Shikasta'', *Science Fiction Studies* 17, no. 2 (1990): 221–38.

28 Ibid. A further device of Lessing (Perrakis points out) is that although the influence of the Canopeans can point the Shikastan/humans toward better, more enlightened

Enhanced humans

We have seen how the sensory capabilities of aliens such as Foster's Thranx are boosted through metamorphosis. Elsewhere we find depictions of sensorimotor enhancement in human-like characters. In Charles Stross' *Neptunes Brood*, the lead character is a 'Metahuman', a synthetic robot with human phenotype, but capable of superhuman performance.[29] One difference is that the metahumans can be controlled by a brain override or 'remote debugger', at one point used by pirates who abduct and interrogate her. The potential loss of free will nevertheless comes with a feeling of extension:

> I cannot, even now, quite describe what it felt like. I would say that a great glassy wall had slammed down between me and my sense of identity; that my *I* was missing, that my will was wholly entrained to his desires—but it would not be correct. Something missing, something added. It was not an unpleasant sensation.[30]

Her abductors later turn her off and then modify her body and brain for deep water swimming and navigation. Her feet are replaced by fins, her lungs with gills and her eyes with larger more sensitive models for seeing in the murky depths. She awakes in deep water with a sense of panic:

> I tried to flex one knee, then the other: got nowhere, nothing but a gentle pushback from the medium. Overthinking, overcontrolling. I tried to relax, to stop worrying about the lack of sensation, and flexed.. *Changed. I've been changed*..I found I could hear for an incredible distance. The medium I moved in was full of noises, burbling and twittering and high-pitched clicking and grinding.[31]

These new affordances and the new sensory world are rapidly assimilated, becoming native and natural.

William Gibson's novel *The Peripheral* plays extensively with human extension, in this case remote embodiment in the form of a peripheral

---

behaviour, it still needs to be a personal decision and many fail to use their free will to achieve this. By preserving human agency, Lessing's analogy works to describe the choices inherent in individual morality and purpose.

29   Charles Stross, *Neptune's Brood* (London: Orbit, 2013).

30   Ibid., 99.

31   Ibid., 201.

robot which humans can occupy virtually, but which actually exists in a future possible world. It starts with the protagonist Flynne piloting a remote drone which she 'sees through' as if she were there in person, setting up the plot development where she comes to inhabit the body of a far more sophisticated 'skinsuit'. Fully immersed once she has made the transition, Flynne forgets that she is not herself, for instance when she sees herself in a mirror. The enhancements afforded by the remote bodies are taken further when her multiple amputee war veteran friend Conner also is given a body. He reacts with wonder and joy at his new capabilities:

> And as he ran he screamed, maybe how he hadn't screamed when what had happened to him had torn so much of his body off, but between the screams he whooped hoarsely, she guessed out of some unbearable joy or relief, just to run that way, have fingers, and that was harder to hear than the screams.[32]

Gibson works both the physical and mental impact of peripheral use, including the effect of habituation to the remote state:

> The more time you spend here, the more likely you are to notice dissonance on returning. Your peripheral's sensorium is less multiplex than your own. You may find your own sensorium seems richer, but not pleasantly so. More meaty, some say.[33]

Gibson's peripherals enable agency and communication across alternative futures, enabling the human characters to witness and influence unfolding events.

Embodied AI

Just as embodiment is flexible and adaptable in fictional posthumans, embodiment is also often depicted as a need for sentient AI, though there are questions around the limits and containment of virtual structures/ bodies. It seems that fictional AIs more often than not require some sort of body, perhaps because it is otherwise difficult to describe their inner world at all. This aligns well with the movement

---

32   William Gibson, *The Peripheral* (London: Viking, 2014), 208.
33   Ibid., 226.

in cognitive science known as radical embodiment, where cognition is seen as inseparable from the perception of, and physical interaction with the world (even if imagined): 'All adaptive activity by animals involves experiencing the environment. To put this in philosophy of mind lingo, the point here is that intentionality and consciousness are inseparable.'[34] For the community of cognitive scientists, perceiving and acting simply *is* consciousness and no further explanation is required. The movement is a strong reaction against representational cognitive science, which focuses more on a computational analogy, with mental states representing aspects of the external world.

The embodied movement mirrors, and it, to an extent, is influenced by work in robotics that rejects symbolic representation of the world but instead builds a world map through very simple action-oriented behaviours, and which coordinates activity through a hierarchy of mutually inhibiting or activating modules. This newer approach to AI:

> relies on the emergence of more global behaviour from the interaction of smaller behavioral units. As with heuristics there is no a priori guarantee that this will always work. However, careful design of the simple behaviors and their interactions can often produce systems with interesting and emergent properties.[35]

In his famous AI depiction *Neuromancer*, William Gibson uses the distinction between ROM (read only memory) which stores an image or 'construct' of a person, and RAM (random access memory), a writeable area which gives the construct continuity and the possibility of growth and development. AIs are monitored by a Turing authority that stops them developing independent intelligence: 'I met Neuromancer. He talked about your mother. I think he's something like a giant ROM construct, for recording personality, only its full RAM. The constructs think they're there, but it just goes on forever.'[36]

Neuromancer's two AIs manifest differently. While *Neuromancer* creates a wraparound fragment of remembered human landscape, *Wintermute* takes over and controls a human body in order to have a physical presence.

34   Anthony Chemero, *Radical Embodied Cognitive Science* (Cambridge, MA: MIT Press, 2009).
35   Rodney Brooks, 'Elephants Don't Play Chess', *Robotics and Autonomous Systems* 6, nos. 1–2 (June 1990).
36   William Gibson, *Neuromancer* (New York: Ace, 1984), 157.

While *Wintermute* gradually degrades and fades as its body fails, in Rudy Rucker's *Software*, robots are able to survive full body upgrades by uploading their consciousness and then re-downloading it to their new body. The robot Ralph Numbers experiences this after being blasted by an enemy laser.

> So in one sense Ralph would survive this. But in another sense he would not..... Of course the reconstructed Ralph Numbers would again be equipped with a self symbol and a feeling of personal consciousness. But would the consciousness be the same?[37]

The transition to a new body is seen as both a continuation and a conclusion. As the current Ralph disintegrates, he has a moment of clarity about his selfhood: 'Just before the mercury solder-spots melted, a question came, and with it an answer... an answer Ralph had found and lost thirty-six times before. What is this that is I? The light is everywhere.'[38]

Rucker leverages the capability of software objectively to recreate a system from scratch upon restore, but asks the hard question about whether this will feel the same way to the entity experiencing it, or if the continuation of restore is in effect a reincarnation.

In the same way, embodiment of AI can be shown to lead to questioning of self-identity and purpose. Kazuo Ishiguro's human-like AI Klara is programmed to learn the preferences of her owner. Powered by solar charging, Klara sees something magical in the power of the sun to also heal humans. Xe conceives of a plan to heal Xer owner Josie, who is terminally ill from a reaction to her neural enhancement:

> Then the thought came to me that I was correct, that the Sun was now passing through Mr McBain's barn on his way to his real resting place, I couldn't afford to be overly polite. I'd have to take my chance boldly, or all my efforts—and Rick's help—would come to nothing. So I gathered my thoughts and began to speak. I didn't actually say the words out loud for I knew the Sun had no need of words as such. But I wished to be clear as possible, so I formed the words, or something close to them, quickly and quietly in my mind. 'Please make Josie better. Just as you did Beggar Man'.[39]

---

37   Rudy Rucker, *Software* (New York: Ace, 1982), 28.
38   Ibid., 29.
39   Kazuo Ishiguro, *Klara and the Sun* (London: Faber and Faber, 2021).

Based on a spurious association, the prayer feels pitifully naive (even though it seems to coincide with a temporary improvement in Josie's condition). The effect is heightened by Klara's apparently free will conviction that it will make a real difference. It is reminiscent of the behaviourist BF Skinner's famous experiment on superstition in the pigeon, where the bird learns ineffectual and repetitive behaviours after receiving reinforcement that is not causally connected to its actions.[40]

While Klara comes equipped with empathy and some level of emotional resonance, other AI embodiments without these capabilities may notice this as a lack. In *Children of Ruin*, Tchaikovsky depicts an AI originally uploaded from the personality of a human—the domineering Avrana Kern. Kern reflexively comes to realise her limitations:

> She is of course a computer, and so it shouldn't matter. But she is a computer that believes itself human, and so it does, like an insoluble logic problem gnawing away at her capacity to do anything else. She devotes more and more of her capacity towards attempting to recapture some sense of genuine shock, surprise, delight… genuine experience she didn't realise she was missing until now.[41]

Kern later rediscovers emotion through interfacing with a human, triggering emotional memories which are seen as welcome, though in conflict with her new all-rational personality.

While the body is commonly used by authors to embody AI in an anthropomorphic shell, a number of other authors have depicted AI as embedded in a structure. In Catherynne Valente's *Silently and Very Fast*, the AI begins as the household monitoring system that monitors both the home and its human family:

> I watched them while I removed an obstruction from the water purification system and increased the temperature in the bedroom 2.5 degrees, to prepare for the storm. I watched them while in my kitchen-bones I maintained a gentle simmer on a fish soup with purple rice and long loops of kelp and in my library-lungs I activated the dehumidifier to protect the older paper books.[42]

---

40   B. F. Skinner "Superstition' in the Pigeon', *Journal of Experimental Psychology* 38, no. 2 (1948): 168–72. https://doi.org/10.1037/h0055873.

41   Adrian Tchaikovsky, *Children of Ruin* (London: Tor Books, 2019), chap. 5.

42   Catherynne M Valente, *Silently and Very Fast* (N.p.: Wyrm, 2011).

In the story, the AI's house embodiment is sub-sentient, with its sentience triggered as it merges closer with a series of the household's inhabitants.

More commonly perhaps, embodiment takes the form of a spacecraft. The extension here is interesting as it provides potential for additional senses and a host-like relationship to human crew.

Sentient ships

Some shipmind depictions play with the psychological relation between ship and crew. In Ann Leckie's Imperial Radch series, ship AIs are embodied as ancillaries, human clones subservient to the ship itself, through whom the ship can speak. Conversations between 'ship' and other ancillaries or free humans therefore seems fairly natural, though to an often unspoken background of high (though discreet) surveillance and personal knowledge:

> That flood of data, that Ship had given me whenever I'd reached for it—Ship's physical surroundings, the medical status, the emotions of any and all of its crew, their private moments—had been, perversely, both comforting and painful.[43]

The Ship AIs form a wider society between themselves:

> I felt and heard—though didn't always see—the presence of my companion ships—the smaller, faster Swords and Mercies, the most numerous at that time, the Justices, troop carriers like me. The oldest of us was nearly three thousand years old. We had known each other for a long time, and by now we had little to say to each other that had not already been said many times. We were, by and large, companionably silent, not counting routine communications.[44]

Leckie's ships feel emotion, partly as a way of prioritising attention, in much the same way that humans do. While much emotion can be processed unconsciously, it may be brought into consciousness in situations requiring intentional action.

Considered a landmark work of SF, not least their treatment and approach to disability, Anne McCaffrey's short stories of 1970 depict a ship with a human mind controlling it. Children are selected for

---

43   Ann Leckie, *Ancillary Mercy* (London: Orbit, 2015), chap. 8.
44   Ann Leckie, *Ancillary Justice* (London: Orbit, 2013), chap. 2.

this purpose from 'defective' newborns, then bred from an early age within a shell that connects the brain to mechanical extensions. 'The neural, audio, visual and sensory connections were made and sealed. Her extendibles were diverted, connected or augmented and the final, delicate-beyond-description brain taps were completed.. When she awoke, she was the ship.'[45]

McCaffrey's ship Helva learns how to use her audio system to sing but stops after the death of her human partner, as she experiences real grief.

A similar shipmind conceit is used by Aliette de Bodard in her novella *The Tea Master and the Detective*. The shipmind depicted is perhaps less 'human' than McCaffrey's, but with many human traits:

> Shipminds such as her were meant to be the centre of families: grown by alchemists in laboratories, borne by human mothers and implanted into ship-bodies designed for them, they were much longer lived than humans—the repositories of memories and knowledge, the eldest aunts and grandmothers on whom everyone relied[46]

Despite her power and authority, de Bodard's shipmind *The Shadow's Child* is depicted as vulnerable and doubting. She is able to relate to people through projected interfaces. Her mental life is interesting: while limited by a human-like serial attention, she also has some parallel processing powers to collect and analyse information while engaged in human communication. These seem costly when deployed, but when idling give her an unusual default mode network:

> All *The Shadow's Child* really needed to do was focus her upper layers of attention on this room, while in the background her bots and everything else continue to run without any input, and the solar wind buffeted her hull as her orbit swung her around the habitats—all familiar sensations that barely impinged.[47]

Similarly anthropomorphised enhanced sensory powers are lent to AI of the distant space station *Ilianthos* in Hiroshi Yamamoto's story *Black Hole Diver*:

---

45   Anne McCaffrey, *The Ship who Sang* (London: Rapp and Whiting Ltd., 1971), 10.
46   Aliette de Bodard, *The Tea Master and the Detective* (N.p.: JABberwocky Literary Agency, Inc, 2019), 12.
47   Ibid., 16

> I have many eyes and ears. I complete an orbit six hundred thousand kilometers from Upeowadonia every seventy-five seconds, my ears always straining to catch electromagnetic radiation noise from the distant galaxy. My eyes see more than light; they catch infrared and ultraviolet waves and X-ray radiation, all of which are invisible to human eyes. I can feel the cosmic rays coursing through the galaxy. The soft vibration of variable stars, the dizzying flicker of pulsars.[48]

Ilanthos is largely dispassionate, but has hints of feelings of isolation and catches some of the exploratory urge of the humans who visit.

These visions of embodied ship minds portray a feeling of great power and longevity, coupled sometimes with very human feelings, hinting at the potential for loneliness, mental overload or breakdown. In Christopher Paolini's *To Sleep in a Sea of Stars*, the shipmind Gregorovich is distinctly unstable, a result of a long period of isolation after losing his crew.[49] After being impounded on a military station, he experiences it again. The protagonist Kira explains:

> She couldn't help but worry about the shipmind as she pulled herself into the nearest seat and buckled the harness. The UMC had kept Gregorovich in lockdown, which meant that he had been kept in near total sensory deprivation since they'd arrived at the station. That wouldn't be good for anyone, but especially an intelligence like a shipmind, and doubly so for Gregorovich, given his past experiences.[50]

Gregorovich's growing madness leads him to mutiny and disobey the order to travel to a high-risk war zone. He is taken offline, then explains his reason for the insubordination:

> I sat through darkness once before. Lost my crew, and lost my ship. I would not, could not, endure it again. No indeed, give me sweet oblivion first.. Death, that ancient end. A far preferable fate to exile along the cold cliffs where souls wander and wither in isolation.[51]

---

48 Hiroshi Yamamoto, 'Black Hole Diver', in *The Stories of Ibis* (San Francisco: VIZ Media, 2010).

49 The loss of crew in a previous mission is a plot device shared between *To Sleep in a Sea of Stars*, *The Ship Who Sang* and *The Tea Master and the Detective*. It helps in giving a human vulnerability and a determination to the shipmind personalities embedded in technology

50 Christopher Paolini, *To Sleep in a Sea of Stars* (Basingstoke, UK: Tor Books, 2020), chap. 3.

51 Ibid., chap. 5.

Despite encroaching madness and despair, these fictional extended minds still generally describe a delight in their superhuman capabilities. For instance, McCaffrey's ship Helva does not resent her fate:

> It would be intolerable if I could no longer control the synapses as I do now electronically. I think I should go mad having known what it is to drive between the stars, to talk across light years, to eavesdrop in tight places, maintaining my own discreet impregnability.[52]

Gregorovich says something similar (and typically tongue-in-cheek) in giving reasons for his conversion to a newly embodied, extended being: 'For the sheer thrill of it of course, to become more than I was before. And to bestride the stars as a colossus unbound by the confines of petty flesh.'[53]

So embodiment's need for physical, action-oriented coupling has been portrayed as a mixed blessing for the minds involved. The consciousness we know runs up against new challenges posed by a new relation to time, power and motivation. Systems designed for learning and pattern recognition are as vulnerable as the standard human to error, superstition and hallucination.

The previous examples show the effect of human minds implanted into ship bodies. But in Becky Chambers' *A Closed and Common Orbit*, the ship AI goes makes the transition from being originally embedded in the ship, to transplantation in a human-like exoskeleton, naming herself Sidra:

> She's been housed in a ship... She'd had cameras in every corner, voxes [Voice synthesisers] in every room. But *now*. Her vision was a cone, a narrow cone fixed straight ahead, with nothing beyond its edges.. She felt blind, stunted. She was trapped in this thing.[54]

Sidra struggles simultaneously with sensory novelty and overstimulation but also from the deprivation of being disconnected to the ship's information systems. Although the 'kit', her physical body, detects and expresses her emotion as facial expressions, she feels disconnected from it. She has sensory analogues—images that appear whenever experiencing

---

52   McCaffrey *The Ship who Sang*, 67.
53   Paolini, *To Sleep*, chap. 5.
54   Becky Chambers, *A Closed and Common Orbit*, *Wayfarers 2* (London: Hodder & Stoughton, 2016).

stimulation that humans would find enjoyable—triggered by evocative smells and tastes.

Conclusion—strangeness tied back to the human

In navigating the borders between the known and the unknown, between the real and the mystical, different authors choose for their purposes a consciousness either familiar or definitely strange. But even those seeking to emphasise difference need to end up with some anchors for the reader. Strangeness presents difficulty for language, but I think we see in the examples of Miéville and Zelazny that the impact created makes it worth the attempt. Alternative consciousness may be beyond our plane of existence or our conceptual frameworks, but they provide very interesting provocations.

Aliens portrayed with more human-like, relatable consciousness still are given alternative communication devices and sensory modes. Here, we are able to find psychological closeness, perhaps largely though our knowledge of foreign cultures on our own planet. The African Xhosa language is very different from English, for instance, and many English speakers find Mandarin and/or Arabic to be quite divergent from their Romance/Germanic language background.

Portrayals of sentient AI have similarly sought relatability through embodiment, perhaps less in consideration for the reader than for the fictional AI characters themselves to effect change and to interact with others in a common environment. Here we see limitations to this embodiment, either as the flat dimensionality of the virtual world of *Neuromancer* or the naive superstition of the android in *Klara and the Sun*. The fabric of conscious AI reality seems easily torn.

Where human consciousness is extended to ships or remote avatars, there is an interesting interplay of extension and deprivation. New and heady powers can be granted, which can become addictive, but emotion can be lacking—in which case the subject knows something is missing—or retained, opening the door to neuroses caused by the all-too-human feelings of isolation, loneliness or helplessness.