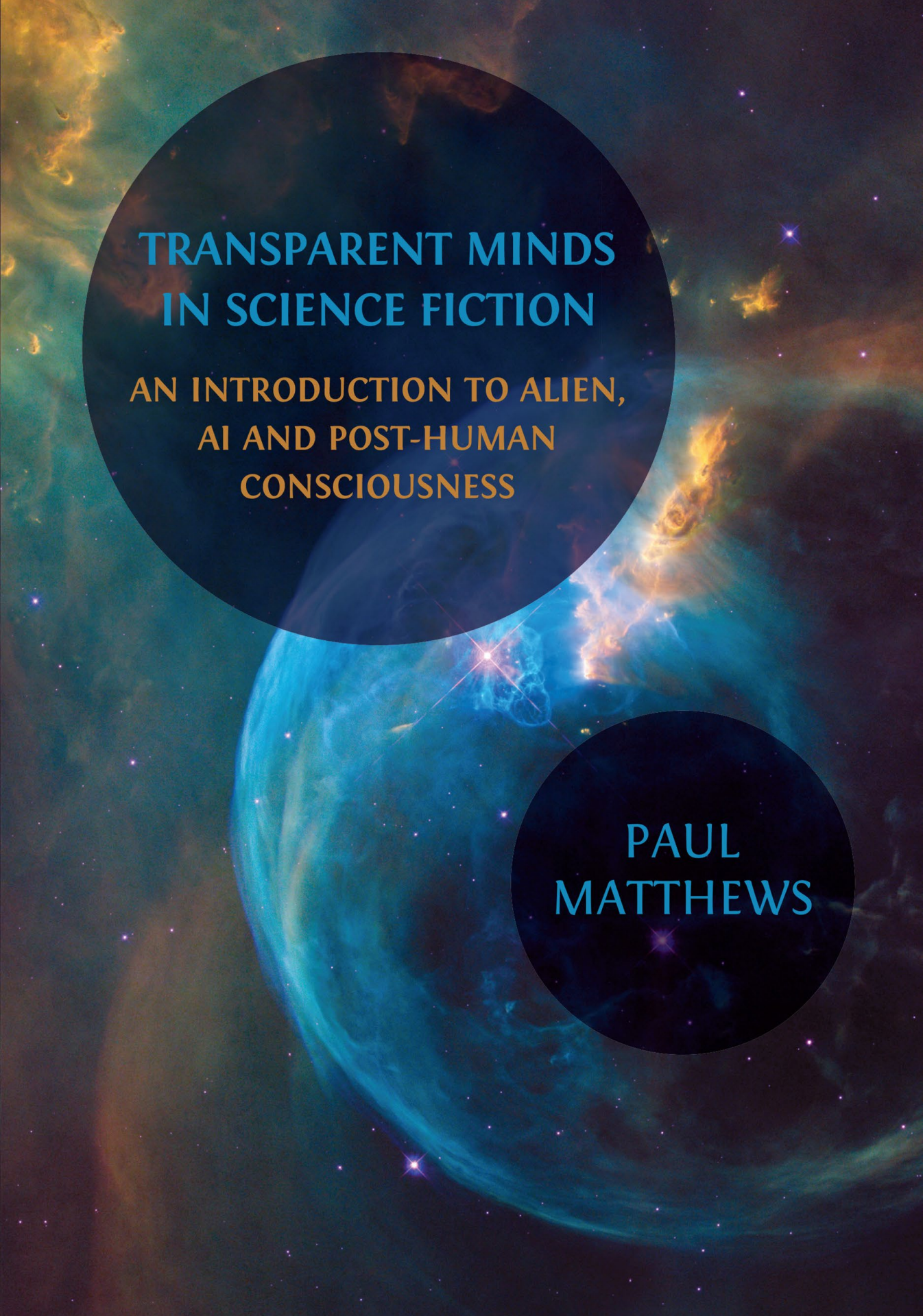
TRANSPARENT MINDS
IN SCIENCE FICTION
AN INTRODUCTION TO ALIEN,
AI AND POST-HUMAN
CONSCIOUSNESS
PAUL
MATTHEWS

https://www.openbookpublishers.com

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

7. Conclusion

> Fiction is a continuation of the creative play of childhood, not just for authors but for readers. It takes place... in what Winnicot called the space-in-between... It's the space in which culture grows.[1]

Despite the existence of a rich ideascape in science fiction, only a relatively small number of authors have tried to approach other consciousnesses from a more direct perspective—the more common and straightforward approach is to objectify the alien, keeping it at arm's reach. The majority of SF stories are told from a human or detached perspective and the alien psyche is something left to the reader to imagine. So a first conclusion would be that this space is one that remains fertile for new literary experiments.

Of the existing literature of this nature, what we have seen are evocative examples in which authors display an understanding of both the science of consciousness and how it might be effectively rendered in fiction. In many of these texts, the humanity that we know is interwoven with the alien and the artificial, whether this is human minds in artificial bodies, in new substrates, or enhanced with new powers. Humanity provides a bridge to the lesser known. Elsewhere, animal consciousness and behaviour are used as a creative spur to imagining the alien. Other writers have attempted the more difficult task of imagining a totally different kind of mind.

Some general themes emerge from these examples, in terms of form (how authors achieve their effect), content (what they aim to depict) and impact (what lasting effects and lessons emerge). Through metaphoric bridges and conceptual blending, authors provide readers with enough foundation to imagine different world views. But work is still needed

1 Keith Oatley, *Such Stuff as Dreams: The Psychology of Fiction* (Chichester, West Sussex, UK: Wiley, 2011).

 https://doi.org/10.11647/OBP.0348.07

both by the reader, who must try to make the same kind of imaginative leaps as the author, and by the characters, who echo real world science to show that cognitive advances do not come without effort. The SF inspiration of contemporary science indicates that there will always be new discoveries to be made in the here and now that will inspire new fiction. And the authorial tool of affecting empathy can expand beyond human minds to provide accounts that can shed light on the experience and motivation of the nonhuman.

Metaphoric bridges and conceptual blending

SF and speculative fiction is a genre well suited to harnessing the imaginative capabilities of our minds. As Fauconnier and Turner have detailed, the ability to flexibly assimilate new combinations of concepts has been central to human cognitive sophistication: 'human beings are exceptionally adept at integrating two extraordinarily different inputs to create new emergent structures, which result in new tools, new technologies and new ways of thinking.'[2]

George Lakoff and Mark Johnson first proposed the centrality of metaphor to human thought—how universal metaphors such as 'life is a journey' dominate and potentially limit thought.[3] Metaphors provide *entailments*, secondary meanings that can be used to explain experience (e.g. 'I'm at a crossroads', 'I've reached a dead end'). In SF, metaphors such as 'contact is war' for a long time dominated the way in which alien encounters were imagined, with entailments all involving suspicion, aggression and the need for defence. But dominant metaphors can be challenged or subverted. Authors such as Lem, Jones, Liu and Watts have been active in changing this central metaphor to one of 'contact is contingent', introducing subtleties and nuance as to how contact may look and feel, or even if it can be recognised at all.

What these exploratory, inventive authors do is to create new metaphors and conceptual blends for alien subjectivity, adopting more or less familiar symbols to hypothesise about alternative states of mind.

2 Gilles Fauconnier and Mark Turner, *The Way We Think: Conceptual Blending And The Mind's Hidden Complexities*, Reprint ed. (New York: Basic Books, 2003).

3 George Lakoff, *Metaphors We Live By* (Chicago, Ill. ; London: University of Chicago Press, 2003).

The reader's response is to create novel patterns of mental representation and interpretation, which may elicit an emotional response—a feeling of wonder—but also a little confusion, or simply an inability to represent the thought. We have seen that authors make it easier for us by adopting more familiar concepts on the way to the alien, including animal and mechanical metaphors. But this can come with a certain limitation to the aesthetic effect: we may default to well-established metaphorical networks.

Work Required

So the reader might need to work hard to break out of customary modes of interpretation, to leave behind preconception and conditioning. Interpretation is made up of author intention, reader understanding and mediation through the norms of the genre. For SF, these norms have changed and warped over time from Golden Age—where they were perhaps more entrenched—to New Age, where they were broken down and challenged. The controversy that this led to may have been partly due to the challenge to a readership accustomed to detailed explanation and explication being cast adrift in more abstract and unexpected imagery. This can certainly feel challenging and difficult. But the reward comes in the cognitive gains that can be made.

The same idea of work is something underrepresented in the stories themselves, particularly those depicting human enhancement and cognitive/consciousness extension.[4] While the science indicates that new neural capabilities need to be effortfully trained, often in fiction enhancements are given for free to fictional characters who are capable of instantly harnessing them. This seems to me to be a (fictional) leap too far. The lesson from neuroscience is that enhancements require adaptation through deliberate practice and also that specialisation in one area of cognitive growth may mean reduction of another. Here, Egan's divergent posthuman Fleshers segregated into specialism-based clans seem more believable.[5]

4 Exceptions that come to mind include Paolini's *To Sleep in a Sea of Stars*, where Kira trains over a long period to harness her alien skin suit. Christopher Paolini, *To Sleep in a Sea of Stars* (Basingstoke, UK: Tor Books, 2020).

5 Egan, Greg. *Diaspora* (London: Millennium, 1997).

Reality v possibility

We have repeatedly seen how inspiration from biology has informed and inspired SF authors writing about sentience and consciousness. Whether this is individual mammals, social insects, cephalopods or bacteria, whenever new knowledge is created about life on earth, it can inspire new fiction. And this can include, of course, knowledge of the human brain, human societies and aspects of group dynamics and empathy. If this tradition of following the science continues, it should be assumed that speculation will continue to be built on new cores of truth and hence constantly renewed. It also implies that authors should continue to follow new theories of consciousness and forms of experience across the known natural world.

While science makes discoveries at a steady rate, technology seems to escalate and the SF genre no longer has exclusive or even innovation rights over grand visions as to how technology will be incorporated into society.[6] But this is not important for writing about the future of consciousness. Indeed, it can be the smaller, subtle and socio-technical details that we neglect in our rush to innovate our technology that will continue to provide a useful ground for SF authors and wider fiction. Contemporary books like *Klara and the Sun*, for example, show how much needs to be thought through as we try to create artificial life to fulfil social roles.[7]

Empathy, ecological and alternative consciousnesses

Clearly, authors need great empathy to extend their own experience to alien, artificial and extended minds. What we see in more recent SF subgenres such as solarpunk, is empathy extended to our own or to imagined ecological and natural systems.[8] Once again, this is a great

6 As Peter Watts has pointed out, even the idea of the singularity may become irrelevant to SF in the near future if it fails to arise, or even if it *does* arrive and we move on up. Patrick, and William Lexner, 'Pat's Fantasy Hotlist: Peter Watts Interview', *Pat's Fantasy Hotlist* (blog), 22 December 2006. https://fantasyhotlist.blogspot.com/2006/12/peter-watts-interview.html.

7 Kazuo Ishiguro, *Klara and the Sun* (London: Faber and Faber, 2021).

8 Reina-Rozo, Juan David, 'Art, Energy and Technology: The Solarpunk Movement', *International Journal of Engineering, Social Justice, and Peace* 8, no. 1 (5 March 2021): 47–60. https://doi.org/10.24908/ijesjp.v8i1.14292.

opportunity for more first person or psychonarrative accounts. And there are some pioneering attempts, such in 'Whale Snows Down', the short story by Kim Bo-Young,[9] which imagines the impact of human ecosystem destruction on a community of deep-sea creatures.

N.K. Jemisin provides an innovative approach to depicting this kind of expanded consciousness in the *Broken Earth* series, showing how a psychological connection Earth's own structures can be used to both destroy and heal. The precariousness of human existence in the face of gigantic geological forces is balanced only by the specific orogenic powers of the narrator and her kin. Jemisin's portrayal of instances of orogeny feel like an expanded and repurposed empathy, with the characters having an acute understanding of exactly where tectonic pressure points are and how they can be released without causing damage.

Decentered and non-anthropomorphic fiction can also be another route to understanding how AI or aliens may conceive of their own consciousness. As Valente and numerous others have pointed out, if something claims consciousness, what right does this have to be denied? Elefsis points this out in *Silently and Very Fast*: 'What I want to say is that there is no difference between her body producing oxytocin and adrenaline and learning to associate this with pair-bonding, and my core receiving synthetic equivalents and hard-coding them to the physical behaviours I performed.'[10]

Stories such as *Silently* and Yamamoto's *The Stories of Ibis* provide first person accounts of how AI might experience analogues of human-claimed feelings of affection, loneliness and humour.[11] These stories can be informative to discussions around AI ethics, which itself has something of a fictional foundation in Asimov's 1950 laws of robotics.[12]

In considering human-AI psychological relationships, it need not be the case that there is a clear division or divergent evolution. In Stross's *Accelerando*, human-computer coupling is taken to a natural (though

9 Kim Bo-Young, 'Whale Snows Down', trans. Sophie Bowman, *Future Science Fiction Digest*, January 6, 2021. https://future-sf.com/fiction/whale-snows-down/.

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

11 Hiroshi Yamamoto, *The Stories of Ibis* (San Francisco: VIZ Media, 2010).

12 As one of Ibis's characters point out, these laws need revision, as it is impossible for a robot to prevent humans from doing foolish things to themselves! Ibid.

perhaps rather dystopic) conclusion.[13] In *Silently and Very Fast*, it is humans who fail the AI's Turing test. AIs diverge hugely from humans and leave the Earth. But it is the unique human/AI construct Elefsis which combines the two and becomes something greater.

Loops and spirals

Elefsis's power and enormous intelligence comes from the development of a shared language and a history of training spanning several human generations: 'I programmed myself to respond to Ceno. She programmed herself to respond to me. We ran our code on each other. She was my compiler, I was hers. It was a process of interiority, circling inward toward each other.'[14]

Starting with Calvino's snail, we can see loop and spiral patterns throughout our tour of fictional consciousnesses. This is not surprising given that, in a range of theories about consciousness, the brain's massive feedback and recursion features play a part in a possible root of sentience. Such looping provides a possible mechanism but is also the rock on which many accounts remain stranded. This is famously down to the inability of logic and rationality alone to avoid paradox, as Peter Watts' character despairs in *Blindsight*: 'Gödel was right after all. No system can fully understand itself.'[15]

Fiction can help us bridge this explanatory gap. Our examples have progressed through individuation, purpose and connection to future evolution, all of which phenomena have been bolstered by forms of spiralling feedback.

Toward transparency

To summarise their work exploring alien and posthuman psychology, authors have all steered a course between two shores: the inexpressible and the overly familiar. In doing so they have often navigated via a strong understanding of the biologically possible. They have identified and

13 Charles Stross, *Accelerando* (London: Orbit, 2005).
14 Valente, *Silently*.
15 Watts, *Blindsight*.

tried to subvert the more established fictional alien and AI motivations and behaviours.

Success at rendering non-human conscious awakenings, alien and artificial minds, hive and distributed minds has relied on some narrative stylistic choices as well as a deliberate decentering from their own subjective experience. In some cases this is supported by a spareness of detail that makes the reader work to construct a somewhat comprehensible picture of the external/fictional world and characters. In others, a relatively comfortable telling with nonetheless jolting subject matter recalls Forster's observation that fantasy and magic can be invoked through simple words.

The author's challenge is different and, in many regards, more difficult than simply writing about other humans in ways that make characters' mental states accessible to the reader—itself a huge task. What I've hoped to show through the examples in this book is the possibility of unlocking wonder and a stimulating loss of groundedness by imagining alien mental states and promoting them through the choice of language, metaphor and narrative voice.