Durham E-Theses

Demonic fictions: cybernetics and postmodernism

Brown, Alistair

Demonic Fictions: Cybernetics and Postmodernism

Alistair Brown

Whilst demons are no longer viewed as literal beings, as a metaphor the demon continues to trail ideas about doubt and truth, simulation and reality, into post-Enlightenment culture. This metaphor has been revitalised in a contemporary period that has seen the dominance of the cybernetic paradigm. Cybernetics has produced technologies of simulation, whilst the posthuman (a hybrid construction of the self emerging from cultural theory and technology) perceives the world as part of a circuit of other informational systems. In this thesis, illustrative films and literary fictions posit a connection between cybernetic epistemologies and metaphors of demonic possession, and contextualise these against postmodern thought and its narrative modes.

Demons mark a return to pre-Enlightenment models of knowledge, so that demonic (dis)simulation can be seen to describe our encounters with artificial others and virtual worlds that reflect an uncertainly constituted and unstable self. By juxtaposing Renaissance notions of the demon with Donna Haraway's posthuman "cyborg," psychoanalytic demons with the robots of the science fiction film *Forbidden Planet* (1956), and Descartes' "deceiving demon" with Alan Turing's artificial intelligence test, I propose that the demon proves a fluid, multivalent trope that crosses historical and disciplinary boundaries. The demon raises epistemological questions about the relationship between reality, human psychology, and the representation of both in other modes, particularly narrative fictions.

When this framework is applied to seminal science fiction (*2001: A Space Odyssey* and *Do Androids Dream of Electric Sheep?* [both 1968]), conventional readings of cyborgs as monstrous Others have to be revised. These fictions are engaged with cybernetic technologies with an epistemological rather than ontological concern, and consequently lend themselves to the kind of sceptical doubt about reality that characterises postmodern thought. Contrary to Descartes, who sees foundational truth through the deceptions of his "deceiving demon," later films like *Blade Runner* (1982) and *The Matrix* (1999) use the motif of cybernetic technologies to highlight the inescapability of the postmodern condition of the hyperreal. Finally, however, literary fictions like Umberto Eco's *Foucault's Pendulum* (1988) and A.S. Byatt's *A Whistling Woman* (2002) and *Possession* (1990) draw attention to their narrative mechanisms through metafiction, and set the creation of literary meaning against computer-generated texts. Consequently, they defy both the determinism of cybernetic sciences, and the postmodern pretence that the "real" is irrecoverably evasive.

Demonic Fictions: Cybernetics and Postmodernism

Volume 1 of 1

Alistair Brown

Ph.D.

Durham University, Department of English Studies

2008

Table of Contents

Acknowledgements

In the course of researching this thesis, I encountered a story in *1001 Nights* in which a character goes about with his destiny written on his forehead. As I write this page – at the forehead of this thesis, at the end of its fate – I cannot help but feel an uncanny affinity with that character. For it seems to me that when some one has had parents prepared to read him *The Hobbit* five times over without complaint, to run the cheapest of taxi services, and to keep the finest of hotels; when some one has grown up with grandparents who have dedicated their time and money to his education; when that foundational schooling leads to a university like Durham, which later invests confidence and money in him through a Doctoral Fellowship; and when the Arts and Humanities Research Council funds the two subsequent years of research; when all these tracks have been laid out, then happy endings seem somehow inevitable, and it is as much by fortune as judgment that such a some one – my own self – is sitting now writing these words. Both literally and figuratively, this page contains the only personal note of this thesis. Everything that follows has been dependent on the guidance, support and education of others.

Having said that, it is only with hindsight that my plot seems so neatly concluded. Time and again as I plodded my way through this research, just when I seemed to have the reading done and a chapter sealed, I found myself lead astray by a demonic voice of great humour and warmth, but of unerring cunning. Over the past three years, Professor Patricia Waugh has offered me numerous prompts and insights that I might not otherwise have reached. Without her input, this would have been finished long ago. With it, it is infinitely better than it might otherwise have been. I am immensely grateful for her guidance and support. I am also grateful to all staff, academic and administrative, of the Department of English Studies, in particular for helping with my research proposals and proliferating application forms. Professor Corinne Saunders offered carefully pointed comments on the Renaissance contexts of chapter 2. I was also saved by my exorcist, Beth Hannon, who carefully proof-read and teased out the devils in the detail. Naturally, any that remain are sins of my own.

Finally, if this thesis has possessed me for the past three years, for ensuring that my reason has been allowed to sleep in the evenings without producing monsters, I owe most to my friends and family. In particular, Chris has listened to me grumble over numerous mugs of tea. Stephen and Lesley have provided the comfiest of cottages, on the edge of the cobweb-blowing moors. And, more than anyone, Helen continually reminds me that the most vital destinies are, at the end of the day, not written on paper.

<u>Chapter 1. Possessing Narratives: Fictional Games and Cybernetics</u>

In a narrow corridor cast in an electric blue hue, a surveillance camera pivots and zooms to observe the pile of charred robots lying at your feet. As you glance up, victorious, you recognise that the pulsing red iris of the camera which is pointing down at you signifies that its controlling consciousness is still, frustratingly, alive. The camera is the eye to the soul of Shodan, the spaceship's computer system which has somehow evolved from its programmed routines to become conscious, operating now with all the unpredictability that comes of emotion. Its voice issues from speakers as you stare into the camera, taunting you that one wave of attacks is over for now, but the next move in an inscrutably complex strategy is about to take place. At root the female voice of a mere housekeeping automata, Processing Unit 43893, it is now spitted with static and pitch-adjusted to mimic a masculine tone. Its sexual hybridity eroticises an entity that crosses comparable boundaries separating the mechanical tool and the sentient life form. And behind Shodan's taunts lies a strange affection: even as Shodan wants to destroy you to prove her (his? its?) superiority to humans, she needs you to bear witness to her rise to consciousness.

But if you were able to pause the waves of enemies Shodan despatches, silence the trance-like soundtrack, and escape from her incessant surveillance, if you were able to take a step out of the game and watch yourself playing, you might note that you have a similarly paradoxical relationship with the computer. Shodan's one weakness is that she cannot assault you directly but must possess robots and human cyborgs, which sometimes malfunction; destroying Shodan should prove the power of your self-contained, embodied consciousness. On the other hand, mastering the game requires that you become even more like the artificial intelligence, a mental process concentrated purely on the second life behind the screen. You are a cyborg working through the medium of the personal computer. You are possessed by the ghostly presence of the machine.

<u>The Ghost in the Machine: The Cybernetic Medium</u>

Who, then, is controlling who in this narrative, a pastiche of the cyberpunk role-playing game, *System Shock 2*?[1] Clearly it is not a one-way correspondence between the embodied player and his or her avatar; my slippery second-person "you" which refers indeterminately to the digital character and the physical player reflects the uncertainty of

---

[1] *System Shock 2*, vers. 1 (Cambridge, MA: Looking Glass Studios, 1999).

subjectivity here – indeed, it might apply also to you, dear reader, momentarily assuming that you too are playing the game, my game, of projecting your mind into an imaginary world through the medium or interface of print. Although such self-reflexive relationships between the textual and the cybernetic appear presently in this argument, let us for now keep it simple, and attend only to the relationship between the physical player and his digital avatar.[2] Although according to the plot the gamer is reasserting his anthropocentric control over errant technology, in actuality his enculturation into the virtual environment requires him to conform to the standards of a comparable system to that which he aims to destroy.

In the very early stages of the game, when life on the space station seems normal, Shodan maternally guides the player, explaining the dynamics of motion and action in the virtual world in her soothing, feminine voice. The player learns the keyboard mappings that allow him to walk, leap, crouch, swim, fight; Shodan also explains how to cyborg the avatar through using mechanical augmentations and neural stimulants, and how to interface with the cyberspace computer terminals within the computer game in a recursive plunge deeper into the system. Thus the full dynamics of activity are unrestricted by the biological limitations of the player's own body, and the game playing experience is liberating. However, the key to success involves the player acknowledging that he cannot defeat the errant Shodan in any way other than by accommodating himself to the control system she has previously laid out. This is evidenced through the keyboarding of thought. The cardinal points of the virtual geography are mapped onto the four cursor keys, but there is a fifth element required in playing that is not needed in reality. In embodied life, the process of moving in three dimensions essentially is the process of interaction; exercised primarily through the hands, to work with the physical world is simply to intercept arms and fingers with the structure one consciously wants to manipulate.[3] From the behavioural perspective, thought can be inferred from the "intentional stance" extrapolated from an agent's bodily

[2] For simplicity, and because games playing remains a male-dominated entertainment, I use the masculine here, whilst recognising that it might apply to either gender.

[3] Raymond Tallis celebrates the hand as "the master tool of human life. The brain's most versatile and intelligent lieutenant, the master grasper, it is simply ungraspable." Raymond Tallis, *The Hand: A Philosophical Enquiry into Human Being* (Edinburgh: Edinburgh UP, 2003) 22.

motions and vocal expressions.[4] In the game world, however, the dimensions of movement are not enough logically to coordinate mind and body. Instead, usually mapped through the mouse button, the "interact" command informs the computer that thought has been happening in the mind of the player and must be recognised and translated into the game environment through a programmed heuristic.[5] The player, for example, looks at a camera and pressing the button causes him to shoot it; or he looks at an access pad on a locked door, and pressing the button activates a key held in the inventory. Thought is extracted from the player's mind and interpreted by the computer. Press the mouse to signify that in real life you are thinking of using explosives to blow up the locked door in the virtual environment, and no response will be forthcoming. In part, successful playing lies in the ability to predict which particular heuristic will be invoked in any particular circumstance, and though the player appears to possess freedom of movement and cognition, there is essentially only one programmed path to completion, with autonomy an illusion sustained by narrative craft.

Cognition in this case cannot be viewed as taking place in the player's brain alone, since thought is interpolated by and through the program. Examining cybernetic systems and cybertext narratives, N. Katherine Hayles suggests that the proprioceptive sense has been altered by the information revolution. The Cartesian subject begins by visualising his thinking presence in isolation from the world as the one thing he cannot doubt. However, the modern human who inhabits information-rich environments "knows that the dynamic and fluctuating boundaries of her embodied cognitions develop in relation to other cognizing agents embedded throughout the environment, among which the most powerful are intelligent machines."[6] In her anthropological studies of computer users, Sherry Turkle

---

[4] The "intentional stance" derives from behavioural psychology, but is popularly explicated in Daniel C. Dennett, *Consciousness Explained* (1991; London: Penguin, 1993) 76.

[5] The future of gaming systems seems set to eliminate the need for a bodily interface at all. In 2008 the Emotiv company retailed the first neurological headset allowing gamers intuitively to manipulate the game environment using brain activity alone. Darren Waters, "Brain Control Headset for Games," *BBC News Online*, 20 Feb. 2008, 2 June 2008 <http://news.bbc.co.uk/1/hi/technology/7254078.stm>.

[6] N. Katherine Hayles, "Flesh and Metal: Reconfiguring the Mindbody in Virtual Environments," *Configurations* 10 (2002): 303. Andy Clark and David Chalmers have expounded a similar hypothesis that mental processes are not confined within the body's boundaries but should be seen at work within the environment with which the self interacts. They argue that a player of the shape-fitting computer game, Tetris, is possessed of an "extended mind" regardless of whether they must press a key to rotate the shape or

has compiled extensive evidence of how people react to the different worlds and identities the computer allows, through multi-user dungeons (MUDs), chat rooms, artificial life simulations, and multi-tasked operating systems.[7] Are digital worlds really distinct from reality, or do they blur together, situating the self indeterminately between the two? Turkle maintains that computers are "marginal objects on the boundary between the physical and the psychological" and as such they force people to "build theories around the animate and the inanimate and to develop their ideas about thought itself."[8] Computers can provide powerful philosophical foci for understanding the nature of subjectivity. However, when such self-reflection does not take place, the individual can be characterised as possessed by the machine. For example, Turkle describes the case of Stuart, a physics graduate who joined the MUD, Gargoyle, under the pseudonym Achilles. In an elaborate but textually-mediated exchange of rings, his avatar was married to another in the chat interface. But when in real life, Stuart/Achilles was asked whether he was single, his reply was that he was married.[9] This blurring of boundaries between the human body and technology is commonly defined as the posthuman or the cyborg state. In this sort of scenario, the human becomes, as Hayles puts it, caught up in "a coupling so intense and multifaceted that it is no longer possible to distinguish meaningfully between the biological organism and the informational circuits in which it is enmeshed."[10]

Though the resulting posthuman implies an ontological and historical break from the human, is there no way we can conceptualise this splicing of minds into cybernetic circuits using existing frames of reference? Erik Davis has taught us that technology has not entirely disenchanted the world. Rather the old, metaphysical phantasms "went

---

whether they have a neural implant which allows thoughts to cause the action directly. In both cases, the skin/skull boundary of thought comes into question. Andy Clark and David J. Chalmers, "The Extended Mind," *Analysis* 58 (1998): 7-19.

[7] Sherry Turkle, *The Second Self: Computers and the Human Spirit* (London: Granada, 1984); Sherry Turkle, *Life on the Screen: Identity in the Age of the Internet* (London: Weidenfeld, 1996).

[8] Turkle, *Second Self* 21.

[9] Turkle, *Life on the Screen* 192-205.

[10] N. Katherine Hayles, "Virtual Bodies and Flickering Signifiers," October 66 (1993): 80.

underground, worming their way into the cultural, psychological, and mythical motivations that form the foundations of the modern world."[11] Excavating one of these mythical qualities, this thesis tunnels down through the strata of postmodern thought and posthuman cybernetics to reveal an underlying archetype of demonic possession. The demon is an agent originating from some metaphysical zone and controlling a mind so that it no longer perceives an absolute, concrete reality. Cybernetics is similarly concerned with a realm other than the physical and corporeal, in which the immaterial signifiers of computer code cause a change in the cognition and the mind-body relationship of the player who is suddenly able to perform heroic feats of combat more appropriate to the gods than the normatively human. Likewise the demon was said to inhabit an immaterial realm between gods (or the devil) and men. Because they operate liminally between the conceptual and the corporeal, the problem of understanding demons is a perspectival one akin to perceiving a cybernetic world from within the game (a perspective I took in the first paragraph of my pastiche at the start of this chapter) or seeing its user outside it engaged intently with the physical computer (a stance I assumed in my second). In both demonic and cybernetic experiences, for the person being possessed, the experience might seem highly purposive, giving them an element of performative control that they lack in reality. Witnessed from the outside, however, demonic possession seems a highly disturbing and violent experience (it is not surprising that there is a long association between demonism and epileptic fits), whilst the contemporary media have reported perverse stories of people so caught up in life behind the screen, that they have forgotten to eat or sleep in reality and so have wasted away.[12] Often rendered amusingly in the stereotype of the shabby "geek," neglect of the body is the stigmata of cybernetic obsession. However, for the geeks themselves, the computer provides an interface to a social world that is more accommodating than that from which they choose to cut themselves off.[13] This double perspective in which we can only witness a cybernetic possession from an external point of view and never fully enter another person's phenomenological perspective reminds that,

[11] Erik Davis, *TechGnosis: Myth, Magic and Mysticism in the Age of Information*, updated edition (London: 5 Star-Serpent's Tail, 2004) 5.

[12] "South Korean Dies After Games Session," *BBC News Online*, 10 Aug. 2005, 3 June 2008 < http://news.bbc.co.uk/1/hi/technology/4137782.stm>.

[13] For Turkle's analysis of hacker communities, and the discrepancy between the way there are popularly perceived and the way they perceive themselves, see Turkle, *Second Self* 201-46.

equally, verifying the truth of a demonic possession was always problematic in judicial and theological traditions (as examples given later in this chapter indicate) because the symptoms exhibited by the demoniac were so readily fraudable: it was hard to distinguish between a "real" demon producing illusions and a fraud producing the illusion of demons. Just as this made the truth of demons the site of theological tussles (particularly after the Reformation), so cybernetics tests how far we are willing to push the boundaries of cultural relativism. In the modern era, to define Stuart/Achilles as unmarried simply because his ceremony did not take place in physical reality is to fail to recognise that, for this individual, "reality" as a personal construct may be invested primarily in the mental world behind the screen. On the other hand, if Stuart/Achilles is allowed to be married in a MUD, can we still incarcerate the schizophrenic who murders in the genuine conviction that he is following the word of God which possesses him (the ethical dilemma in A.S. Byatt's novel *A Whistling Woman*, which I examine in chapter 9)? As Chris Hables Grey reports in *Cyborg Citizen: Politics in the Posthuman Age*, current laws will not translate easily to a future of virtual environments, transgendered sexes, robotic bodies, and genetic clones.[14] The virtual or cyborg lives that are facilitated by cybernetics are forcing us to broaden the boundaries of what constitutes the normatively human, leading to a more open society as the geographical boundaries separating races, or embodiments that indicate sex, are breached. The history of demonism seems to embody a comparable double trajectory, particularly in relation to gender. On the one hand, it is no coincidence that over the centuries women seem to have presented with cases of possession by a divine spirit more often than men, using their minds to assume the visionary theological authority of which their embodied selves were deprived.[15] Conversely, there is a close association between demonism and witchcraft, the trials of which between the fifteenth and seventeenth centuries were used legally to proscribe the social status of women.[16] Indeed, this association between demonism and women's rights can be read within the most influential document of changing political subjects in cybernetic modernity, Donna Haraway's

---

[14] Chris Hables Gray, *Cyborg Citizen: Politics in the Posthuman Age* (New York: Routledge, 2002).

[15] Elizabeth Alvilda Petroff, *Medieval Women's Visionary Literature* (New York: Oxford UP, 1986).

[16] Stuart Clark, *Thinking With Demons: The Idea of Witchcraft in Early Modern Europe* (New York: Oxford UP, 1997). See especially 106-33 for the links between womanhood and witchcraft.

feminist "Cyborg Manifesto." In chapter 2, by comparing Haraway's cyborg to the work of a sixteenth-century naturalist and demonologist, I show that contests over the verifiability of reality – can we be possessed by demons? can we inhabit cyberspace? – are also ones about the ethical codes according to which human subjects should or should not operate. The problem in both postmodern and pre-modern epistemologies is how far we can make ethical assertions, given that we cannot truly know how the world is perceived from the perspective of another, especially someone possessed by demons or cybernetic systems. These are just some of the threads that intertwine cybernetic simulations with demonic history, and I will trace some of these later in this chapter, and over the course of this thesis.

However, at risk of creating a tangle, I want to add another thread of connection. When he turns off *System Shock*, the game's player experiences a legacy effect: ending the game makes the player aware once again of their embodied consciousness, so that even without the explicit polemic against artificial intelligence, he should still think about thinking, and what it is to be human as opposed to artificially intelligent. Although the medium by which this effect is produced may be relatively novel, the message is a common one to do with locating the conscious self in relation to the inanimate world. In this sense, cybernetic simulations can find ancestors in other modes of production of alternative perceptual fields, in which the simulation requires us to test its correspondences and differences to reality, in the generic relationship known as allegory. Narrative has also long provided a mechanism by which a world is mimetically rendered in a way that is at once the same but also dislocated by its being represented through the medium of language. Predominantly novels – but also poems, films, metaphors in general, and even scientific hypotheses – are forms of simulation that set up feedbacks between themselves and their interpreters in which perception is altered in a potentially deceiving way which may nevertheless thereby provide access to some deeper truth about human reality. And, starting with the Delphic oracle and artistic muses, moving from the Apostolic revelation to the inspiration of Caedmon, father of English poetry, there is a long heritage of hallucinatory demons or spirits as the source of literary creativity, with inspired or prophetic writing explaining the relationship of the human to the world or to the heavens.[17]

[17] It is also no coincidence that three of the most influential horror stories, *Frankenstein*, *Dracula* and *Strange Case of Dr. Jekyll and Mr. Hyde*, have their origins in vivid nightmares experienced by Shelley, Stoker and Stevenson, in which their experiences in reality blended with dreams in terrifying fashion. Christopher Frayling, *Nightmare: The Birth of Horror* (London: BBC Books-BBC Worldwide, 1996).

In relation to literature, the affinity between a demonic spirit, creativity and technology has been simultaneously uneasy and productive. Plato was suspicious of writing – a *techne* – for the way in which it might lead to the redundancy of personal memory and only the mimicry of an essential truth rather than the truth itself. Whilst freeing space in the mind so it could venture to uncover new knowledge, the spoken word expressed a personal veracity which could also be directly challenged through the Socratic method; in writing down his experience and knowledge, however, the author lost control of his original speech and could always be misrepresented in unintended ways.[18] Oddly, though, it is precisely the unintentional and irrational quality that appeals in the myth of creativity. As Timothy Clark explores in his study of literary inspiration, creativity has inherently demonic tendencies: "A crucial part of the process of composition is understood as a desired or even calculated suspension of reasoning or deliberation, a temporary mania or insanity."[19] Etymologically, inspiration derives from the Latinate for "to breath into."[20] Glossed as an empowering spirit, to be inspired is to lose one's agency, to be spoken through rather than to speak; yet, of course, it is a matter of personal credit that the writer opens himself to such a possessive experience. In medieval Christian tradition, the scribe has authority only insofar as he is the medium through whom divine truth speaks; inspiration, then, is the right to speak public truth precisely because the scribe has been able to remove his agency, being simply the mechanism by which a divine will or subjectivity is exercised. Composition in writing and speech appear to have been treated as comparable ways of mediating a metaphysical realm to the public up until the eighteenth century. When technologies such as the printing press and the existence of a mass literate audience appear, however, the author is given space to create originally, rather than merely

---

[18] Plato's Socrates argues that, like figures in paintings, words are mute when interrogated. Consequently "once it is written, every composition trundles about everywhere in the same way, in the presence both of those who know about the subject and of those who have nothing at all to do with it, and it does not know how to address those it should address and not those it should not. When it is ill treated and unjustly abused, it always needs its father to help it; for it is incapable of either defending or helping itself." Plato, *Phaedrus*, trans. and introd. Christopher Rowe (London: Penguin, 2005) 63.

[19] Timothy Clark, *The Theory of Inspiration: Composition as a Crisis of Subjectivity in Romantic and Post-Romantic Writing* (Manchester: Manchester UP, 1997) 2.

[20] "Inspire," *The Oxford English Dictionary*, 2nd ed. 1989.

to ventriloquise the divine or mythical truth. As Clark shows, this space of composition which was carved out by the advent of Enlightenment science is nevertheless paradoxically situated in an ongoing culture of inspiration which derives from a supra-rational cause outside the artist: "The space of composition may be an encounter with what might be called a *phantasmatic author*, i.e. that transformed sense of identity that arises from interaction with the various phantasmatic agencies at work in the space of composition."[21] One might suggest that the matrix of an embodied subject, a demonically inspiring source outside of him, and writing that is now considered a technology in its own right (rather than, as against the Greek oral tradition, a way of transcribing something that had its origin in speech), is itself a proprioreceptive revolution akin to that of the posthuman. The anxiety and power bound together in writing stem from the fact that literature is a mediating, even dominating, experience, as well as a purely mimetic one. Focalised through the concept of the demonic hallucination, the literary and the digital simulation find themselves in odd proximity. Not only can the literary medium engage with cybernetics – interrogating, for example, the ethics of the technologies that emerge from the science – but cybernetics also confronts literature with its own doubts about the directness of the correspondence between authors and readers, or the world and words.

Though the aesthetic of the Romantic genius able to mediate between the zone of inspiration and the page might seem diametrically opposed to the Newtonian world of mechanism, oddly the idealised literature, like science, is a medium permitted to speak for itself. An uncontaminated link between the unconscious and the pen involves the erasure of the author as their subjectivity is focused through the technology of writing. By representing the author as a function of discourse rather than as one with privileged access to the truth through contact with the divine, a literary text can cross historical and cultural contexts in ways closed to the physical author, just as a scientific law remains true for all relevant circumstances. Similarly, the posthuman entails a dissolution of an embodied ego in favour of a multiple self capable of traversing geography and identities if not at whim, then at least with the click of a button. It is not surprising that the technological media of the cybernetic era have informed postmodern constructions of the subject. Specifically in

---

[21] T. Clark 26. Along the same lines, Friedrich Kittler notes that the pen, with its ability to record thought almost directly in a near seamless flow between mind and page, means that creativity is displaced until after the inscribed record is produced. Only in rereading and editing is a conscious, egocentric act of authorship to be found. Friedrich Kittler, *Discourse Networks 1800/1900*, trans. Michael Metteer (Stanford: Stanford UP, 1990) 109.

relation to language, the view of literature as an autotelic system in which the embodied author is merely a medium for the abstract informational content of the message is validated by cybernetics, which understands and studies information (whether consciousness, genes or texts) as floating detached from any one specific substrate (the organic brain, the embodied organism, physical books). As Jean-Pierre Dupuy has pointed out, a post-structural literary theory that denies mimesis shares much with cybernetics: post-structuralism claims that a text is closed to the world outside of itself, with the signifiers it contains representing the autonomy of the text from its original production in the mind of an author in the world; cybernetics similarly represents the detachment of information from context.[22]

Oddly, though, there appears to have been comparatively little research into the relationships between cybernetics and postmodern literary theory. Probably this is because cybernetics and postmodernism are such heterogeneous and influential paradigms that we would hardly expect there *not* to be connection. Therefore, working out which relationships are direct such that postmodern thought depends on cybernetics for its epistemological base is difficult. Some interdisciplinary flows certainly do find a source in cybernetics. Jacques Lacan's 1954-1955 seminar was dedicated to cybernetics, and emphasised the subject as the product of symbols or meetings just as "the symbolic world is the world of the machine"; Levi-Strauss's *Structural Anthropology* acknowledges the importance of cybernetics for the future of the social sciences, since Strauss's project involves interpreting society as a whole according to theories of unconscious communication rather than as individual subjects embodied in their own world space; Roman Jakobsen employed Claude Shannon's *The Mathematical Theory of Communication* in his treatment of language as a coding system that structured the exchange of information.[23] Although connections exist, though, there is a risk of over-

---

[22] Jean-Pierre Dupuy, "Self-Reference in Literature," *Poetics* 18 (1989): 491-515, *ScienceDirect*, 10 Aug. 2008 <http://www.sciencedirect.com>.

[23] All these examples are taken from one of the few, recent articles to correct the deficit of attention in this area: Céline Lafontaine, "The Cybernetic Matrix of 'French Theory'," *Theory, Culture and Society* 24.5 (2007): 27-47, *Swetswise*, 3 Sept. 2008 <http://www.swetswise.com.ezphost.dur.ac.uk/eAccess/searchArticles.do>. Lafontaine notes that whilst the relations between structuralism and cybernetics have been examined to a degree, there is a definite lack of knowledge of the influence of cybernetics on post-structuralists. There is a brief description of Lacan's cybernetic seminars in Jean-Pierre Dupuy, *The Mechanization of the Mind: On the Origins of Cognitive Science*, trans. M. B. DeBevoise (Princeton: Princeton UP, 2000) 108-9.

extrapolation in order to construct a direct epistemic relationship. In practice, the cybernetic may simply be used as a handy analogy for more sweeping revaluations of literary and social structures that characterised post-war thought, and the reworking of Plato's problem with writing for a hegemonic age of information.

For example, in her work on artificial intelligence, Margaret Boden elaborates numerous examples of computers producing literary artefacts that, to a naïve observer, might be taken as creative.[24] James Meehan's TALE-SPIN program can produce Aesop-style fables which are at least structurally sound, though wholly unadventurous.[25] The irony is that although TALE-SPIN is artistically limited, literary theory had already deconstructed the myth of human creativity. Extrapolating the Romanticist vein to its extreme, Roland Barthes proclaimed the death of the author and argued that "Writing is the neutral, composite, oblique space where our subject slips away," in which the writer is a mere copyist of anterior texts and gestures, producing a work which is anyway only a "tissue of quotations."[26] As Hayles observes, citing the Derridean gap that separates speaking from writing, when writing becomes mediated through "flickering signifiers" on computers underpinned by binary logic, the narrator becomes "not so much a scribe as a cyborg authorized to access the relevant codes."[27] Meanwhile, in TALE-SPIN, the writerly agent is displaced altogether, being a mere function of algorithms and routines. But though there is a neat line to Barthes from TALE-SPIN via Hayles' cyborg author, there is no evidence that Barthes had been inspired by any program. Rather, the philosophies work in

[24] Margaret A. Boden, *The Creative Mind: Myths and Mechanisms*, 2nd ed. (London: Routledge, 2004). Boden tries not to be drawn into the debate about whether such artefacts embody "true" creativity, instead providing a survey of examples from which the reader is invited to draw their own conclusions.

[25] For the details and sample outputs of TALE-SPIN, see James Meehan, "TALE-SPIN," *Inside Computer Understanding: Five Programs Plus Miniatures*, ed. R.C. Shank and C.K. Riesbeck (Hillsdale, NJ: Lawrence Erlbaum, 1981) 197-258. One example output is: "Once upon a time George Ant lived near a patch of ground. There was a nest in an ash tree. Wilma Bird lived in the nest. There was some water in a river. Wilam knew that the water was in the river. George knew that the water was in the river. One day Wilma was very thirsty. Wilma wanted to get near some water. Wilma flew from her nest across a meadow through a valley to the river. Wilma drank the water. Wilma was not thirsty any more." Meehan 199.

[26] Roland Barthes, "The Death of the Author," *Image-Music-Text*, trans. Steven Heath (London: Fontana-Harper, 1977) 142, 146.

[27] N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago: U of Chicago P, 1999) 43.

parallel, cybernetics contending that creativity resides in definitive algorithmic processes in the brain rather than in the effervescent spirit of genius, and Barthes deconstructing logocentric mythologies. Indeed, one might also note some epistemic divergence. Whereas semiotics argues that signifiers do not point to specific signifieds except in accordance with the textual or social contexts in which they are used, the aim of a cybernetic program like TALE-SPIN is to show that the same codes can operate independently of context, with a story telling routine being at work both on the binary code of a computer and, by implication of this model, in the neurological structure of the brain. The possibility that TALE-SPIN is creative is in tension with our belief – which is not only a poststructural one though it may be endorsed by it – that context regulates meaning, so that the fact that Aesop's tales were (*contra* Barthes) produced by a conscious author contributes to their canonical value, whereas knowing that a tale was produced by a computer may significantly bias our aesthetic admiration for it. Thus, even though postmodern theory and practice may be related to cybernetics in some dimensions, both the contemporary sciences and the arts may be children of a larger cultural milieu that is the post-war age of information in which broader questions are being asked than the detail of semiotics. In this case, one way of framing such questions across cybernetics and postmodern literary theory is to trace their demonic construction. For example, is it not possible to apply the following comment by Timothy Clark about literary inspiration to TALE SPIN also? Inspiration, Clark writes:

> affirms a logocentric conception of a language of self-evident authority. Yet it also represents an automatization of the signifier, a speaking without thought or intention. Ideas that sound a little like accounts of aporias in deconstructive thinking merge in bizarre ways with notions that rest on a religious or magical world view.[28]

It is not that TALE SPIN introduces any automatisation of the word that makes it provocative – for, as Clarke writes, such automatisation long lay at the heart of the process of being inspired – but that it makes this process explicit, bringing it to a surface where it can be reductively analysed, rather than keeping it suppressed as a myth of creativity. It highlights that what poets had taken to be the mysterious but paradoxically commendable absence of thought or intention in the poet is actually the material (and therefore non-

---

[28] T. Clark 4.

mystical) residence of such acts deep in the pre-conscious of the organic brain, or in the logic gates of the cybernetic machine.

Given this nexus of postmodernism, mysticism and cybernetics, it becomes hard to decipher whether it is literary theory or cybernetic science (or something that connects the two, like demonology) which stands behind a piece of postmodern criticism like Italo Calvino's essay on "Cybernetics and Ghosts."[29] Here, Calvino argues that if, as cybernetics claims, symbol-manipulating systems such as the brain are comprised of atomistic calculations rather than spiritual flows, we must change our theoretical conception of literary creativity. In particular, if literature is the encoding of some effervescent spirit, "by what route is the soul or history or subconscious transformed into a series of black lines on a white page?"[30] Could it be that authors and texts are essentially machines? Even if their approaches seem quite different, cybernetics and literary theory ask comparable questions about mind, and how its outer productions – such as text, speech, behaviour – correlate with the inner activities of the embodied mind. Determining whether cybernetics flavours the ingredient of theory specifically is possibly irrelevant, given than both are immersed in a soupy *zeitgeist* of contemporary science, society and culture which has its own overall texture and taste. In this, some of the dominant flavours are the omnipresence of simulacra; doubt about the verifiability of reality; the decentred or fragmented nature of the human subject in a world of information; the deconstruction of the metaphysics of human creativity and spirit. All these themes cross cybernetics and literary theory without necessarily relying on either one of them for their epistemology; and they find themselves commonly expressed in the idea of the deceiving but potentially inspiring demon, and the notion of possession by a metaphysical simulation, which pre-exists cybernetics. Rather than attending to the direct connections between literary theory and cybernetics, then, my method is to act as a reader of literary fictions which might accommodate the diverse range of cultural experiences at a historical moment, and to capture these sensations through focusing on this demonic motif. Whilst his essay seems resigned to a reductionist account of the literary genius, Calvino closes by arguing that the sense of "spirit" (the ghost of his title) will be preserved in the mind of the reader, who can interpret an identical text in multiple ways depending on context. Thus I come to this analysis as a unique reader informed by theory, rather than a theorist using literary texts to support a strong hypothesis

[29] Italo Calvino, "Cybernetics and Ghosts," *The Literature Machine: Essays*, trans. Patrick Creagh (London: Secker, 1987) 3-27.

[30] Italo Calvino, "Cybernetics and Ghosts" 14.

about cybernetic communication and literary structuralism. Indeed, one point I shall make repeatedly is that whilst cybernetics may diminish the inspiration of the author by locating "creativity" in computational mental activities rather than in relation to some external muse, cybernetic fictions continually make intense demands on their readers or viewers. They expect us to engage with the works and to tease out the ways in which the mechanism of text or film somehow produces various different possible meanings for each interpreter.

What is Cybernetics?

Having now triangulated this thesis, with its three points comprising cybernetics, demons and postmodern fiction, each can be resolved in more detail, starting with the term cybernetics. Cybernetics has a specific origin in Norbert Wiener's work on feedback, with his 1948 neologism deriving from the Greek κυβερνήτης, or steersman; just as the pilot continually adjusts his course, so the new discipline was concerned not with cause and effect, as with the traditional physical sciences, but with circular causality, in which feedback systems contribute towards goal-directed behaviour. For example, a person reaching for a glass is continually informed by perceptions about how close they are, and feedback from these perceptions determines how much additional movement is imparted to the arm. The paradigm shift here is that whereas the physical sciences ignore teleology, seeing effects as following from causes, cybernetics examines organisms from the perspective of their goals, and sees how their feedback regulation allows them to meet or deviate from them. The term offered a banner for the pioneering and diverse group that came together at the ten Macy conferences between 1946 and 1953, with the first conference characteristically entitled "The Feedback Mechanisms and Circular Causal Systems in Biology and the Social Sciences Meeting."[31] Including participants from psychology (such as Lawrence Kubie), anthropology (Gregory Bateson and Margaret Mead), language (Roman Jakobsen), mathematics, philosophy and physics, "cybernetics" became a broad interdisciplinary rubric, accommodating a range of methodologies and

---

[31] The essential history of these meetings, their participants and research networks is Steve Joshua Heims, *The Cybernetics Group* (Cambridge, MA: MIT Press, 1991). Five of the transactions of the later Macy Conferences were published in Heinz von Foerster, ed., *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems*, 5 vols. (New York: Josiah Macy Jr. Foundation, 1949-1955). The most detailed philosophical study of the group is Jean-Pierre Dupuy's book *The Mechanization of the Mind* (see note 23 above).

technologies: the sciences of mind, computer programming, biological study of physical organisms, sociological study of behaviour, economics. Thinking of *System Shock*, for example, Shodan can be said to express one cybernetic trajectory, artificial intelligence, that has to date proved largely conceptual in nature. However, cybernetics also describes the internal physical circuitry and programming language of the system that virtualises her, that is, the personal computer. The two strands follow from the differences in the early cybernetics groups. On the one hand were those such as John von Neumann, who emphasised the importance of attending to the physiological and mechanical characteristics of brains and machines, and who played a key role in developing the serial digital computer. On the other were those pure mathematicians who were more concerned with theory than practice.[32] For example, rather than developing useful devices, Walter Pitts and Warren McCulloch focused on neuronal logic networks, which derived from Alan Turing's mathematically abstract universal computing machine.[33] Additionally, the debate about whether communication had to require the presence of a conscious interpreter (as in Shodan or the human player), or whether any homeostatic machine (such as a room thermostat or personal computer) could be defined as intelligent, was a key one throughout

[32] For the difference between the two groups, see Dupuy, *The Mechanization of the Mind* 58-69. The difference is expressed in remarks McCulloch made after hearing an impressive talk by von Neumann at the Hixon Symposium, a special meeting spun-off the Macy conferences to discuss mechanisms in the brain that correspond to behaviour in more technical detail: "I confess that there is nothing I envy Dr. von Neumann more than the fact that the machines with which he has to cope are those for which he has, from the beginning, a blueprint of what the machine is supposed to do and how it is supposed to do it. Unfortunately for us in the biological sciences – or at least, in psychiatry – we are presented with an alien, or enemy's machine [that is, the mind]. We do not know exactly what the machine is supposed to do and certainly we have no blueprint of it." Lloyd A. Jeffress, ed. *Cerebral Mechanisms in Behavior: The Hixon Symposium*, proc. Of The Hixon Symposium on Cerebral Mechanisms in Behaviour, Sept. 1948, California Institute of Technology (New York: John Wiley, 1951) 32.

[33] Turing imagines a tripartite device consisting of a machine, an infinite tape divided into squares, and a head capable either of making a mark in one of those squares, erasing an existing mark, or moving the tape one square to the left or right. Turing showed that any mechanical procedure, if adequately coded in the terms fitting the machine's binary operation, can be modelled by the machine. A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society* 42 (1936-7): 230-65. When Warren McCulloch and Walter Pitts suggested that a nerve cell in the brain either does or does not produce an electrical discharge across its synapse, and may do so in a chain of successive binary events following logically from each other, they essentially proposed that a brain is a Turing mechanism. Warren S. McCulloch and Walter Pitts, "A Logical Calculus of the Ideas Imminent in Nervous Activity," *Bulletin of Mathematical Biophysics* 5 (1943): 115-33.

the Macy conferences.[34] The question here is whether intelligence is a matter of origins and causes, with brains capable of developed thinking, or of teleological ends, in terms of producing sensible adaptive behaviours.

As my sweeping history suggests, the precise paradigm contests in cybernetics are far more complex than represented here.[35] With that proviso in mind, it is possible to say that cybernetics developed in three phases following the end of the War. In the first wave, and in accordance with the neologism, cybernetics conceived of individuals as homeostatic systems, designed to maintain equilibrium within their environment. This epistemology followed scientific protocol in considering the observer to be outside the situation they observe. Second-order cybernetics, instigated by Heinz von Foerster's ambiguously titled *Observing Systems* (looking at systems, systems that observe) and driven by Humberto Maturana's and Francisco Varela's work on autopoiesis, recognised that not only does information flow from the system to the observer, it can also loop through the observer.[36] As Hayles summarises this phase, "Reflexivity is the movement whereby that which has been used to generate a system is made, through a changed perspective, to become part of the system it generates."[37] If so, this speaks to the reciprocal relationship between the user and the computer, in which as well as the player controlling the computer, the computer affects the player. Finally, and impelled by the continued failure to realise artificial intelligence, third wave cybernetics moved away from modelling machines on the human (a "top down" approach), towards artificial life that would evolve or emerge; according to some versions of this "bottom up" mode of production, the universe or environment can itself be conceptualised as a giant computer, with life as the cellular automata running upon it.[38]

---

[34] Norbert Wiener, for example, sees any automata that receives information from its environment and responds appropriately to it as a communicative system. Norbert Wiener, *Cybernetics, or, Control and Communication in the Animal and the Machine*, 2nd. ed. (New York: M.I.T. Press, 1961) 39-44.

[35] See note 31 above for references to more systematic studies.

[36] Heinz von Foerster, *Observing Systems: Selected Papers of Heinz von Foerster* (Seaside, CA: Intersystems, 1981).

[37] Hayles, *Posthuman* 8.

[38] Stephen Wolfram, *A New Kind of Science* (Champaign, IL: Wolfram Media, 2002).

Given that in *The Human Use of Human Beings* Wiener carries out cybernetic interpretations of fields as diverse as philology, the legal system and diplomacy, it is legitimate to treat cybernetics very broadly.[39] The fictions I examine reflect the fact that cybernetics essentially names a direction rather than a formula for thought, an identification of allegiance rather than a specifically nameable method. As with David Porush's benchmark study of cybernetic fiction, *The Soft Machine*, I take cybernetics "to embrace not only the information sciences but a metaphor so deeply engrained in our culture, so silently driven down to the roots of our imaginations, that it achieves the status of an element in a new mythology."[40] The mythology in question is the notion that people, and the cultural products of their minds, such as literature, can be described in terms of machines, and that intelligence can correspondingly be reproduced through artificial programs. Though the specific approaches may vary across disciplines, as Steve Heims suggests, the keynote of the Macy conferences was simply "to identify in a behaviourist spirit some of those aspects of what organisms do that can be analyzed in terms of what certain analogous machines do."[41] Even before the first Macy conference, Wiener, working with the medical doctor Arturo Rosenbleuth, had noted an affinity between excessive feedback in mechanical systems (where each incremental action serves not to coalesce upon a goal but to deviate more wildly from it) and the pathological condition called "purpose tremor."[42] Because cybernetics focuses on teleology and behavioural effects rather than material causes, the fact that muscle spasms may be biologically caused whilst a mechanism like a missile aiming system uses a different engine is irrelevant. Regardless of underlying identifications, since the surface outcomes of goal directed – or, in this case, goal misdirected – feedback are comparable, the metaphor is inescapable. Once this axis linking a specific human condition to a phenomenon observable in mechanical systems is established, it becomes only a small assumptive step to extend the metaphorical connection more generally, so that the human as an organism is essentially a feedback mechanism. This is why, although it was shown to be empirically unsound by the 1960s (as I explore in

[39] Norbert Wiener, *The Human Use of Human Beings: Cybernetics and Society* (London: Eyre and Spottiswood, 1954).

[40] David Porush, *The Soft Machine: Cybernetic Fiction* (London: Methuen, 1985) 2.

[41] Heims 15.

[42] A. Rosenbleuth, N. Wiener and J. Bigelow, "Behavior, Purpose, and Teleology," *Philosophy of Science* 10 (1943): 18-24.

chapter 4), the Turing test continues to dominate cultural interpretations of technology.[43]
The idea of communicating with a talking machine which appears superficially
(teleologically) like our speaking selves, whilst not necessarily being causally embodied or
controlled in the same neurological way, points to the core of cybernetic principles.
According to cybernetics, intelligence exists at a high level of abstraction, rather than
being a specific property of venerable individual bodies; bodies themselves, whether
literary texts or those of the human, are shown to be subject to manipulation and alteration
when other non-embodied and artificial intelligences or informational systems intercept
and flow through them. As N. Katherine Hayles has represented it in her study, when the
bodies in question are human, they become viewed as posthuman by the privileging of
"informational pattern over material instantiation, so that embodiment in a biological
substrate is seen as an accident of history rather than an inevitability of life."[44] It is around
this issue of embodiment, then, that arguments for and against the ethics of cybernetics
seem to hinge.

On the one hand, Donna Haraway has argued for the situatedness of knowledge and
"the particularity and embodiment of all vision (though not necessarily organic
embodiment and including technological mediation)." Whilst figuring the cyborg as a
heterogeneous identity capable of transcending conventional gender definitions through
technology, she urges us not to go so far as to submit to "the tempting myths of vision as a
route to disembodiment and second-birthing."[45] In these fantasies, the individual is
characterised as a pure consciousness, with the body a coincidental medium on which it is
instantiated. Marvin Minsky exemplifies such provocative ideals:

> A person is not a head and arms and legs. That's trivial. A Person is a very
> large multiprocessor with a million times a million small parts, and these
> are arranged as a thousand computers...The most important thing about each
> person is the data, and the programs in the data that are in the brain. And
> some day you will be able to take all that data, and put it on a little disk, and

[43] A.M. Turing, "Computing Machinery and Intelligence," *Mind* 59 (1950): 433-60.

[44] Hayles, *Posthuman* 2.

[45] Donna Haraway, "Situated Knowledges: The Science Question in Feminism and
the Privilege of Partial Perspective," *Feminist Studies* 14 (1988): 582.

store it for a thousand years, and then turn it on again and you will be alive

in the fourth millennium or the fifth millennium.[46]

Note that this sort of argument takes the dualistic relationship between materiality and

spirituality both ways. Soul-like data is also a material or mechanical structure, in the sense

that it must be stored on disks or in neurological impulses encoded like programs in the

physical brain which can therefore be modified or extrapolated. However, the material

body as a whole is ultimately redundant, in order that this spirit, now operating as a

symbolic code, can transcend history, freely crossing the millennia. (The un-noted

limitation is that such transcendence is still restricted by the entropic universe – though

even this is something that cybernetics might overcome.[47]) These cyberneticians turn

science and technology into secular iterations of what was formerly known as soul. It is not

surprising that David Noble has contended that the enchantment with technology is "rooted

in religious myths and ancient imaginings."[48] In this new domain, science can empirically

deliver the transcendence that religion could only speculate about, offering practical

solutions rather than mere narrative salves for the fear of human temporality. For example,

Ray Kurzweil quite deliberately entitles his prophecy for computer intelligence *The Age of

Spiritual Machines*.[49] Throughout this thesis I will refer to those affiliated to this view as

the cybernetic transcendentalists. As we shall see in my discussion of films like *2001: A

Space Odyssey*, with its "star child," the idea of the demon as a disembodied mind (or

disembodied force that can alter minds) is an appropriate way to explore the possibility of

cybernetic transcendence in fiction. On the other hand, the moral issue of whether a

conscious Hal can be disconnected (that is, executed), like Stuart/Achilles who wants his

virtual marriage to be acknowledged as genuine, destabilises moral norms. Liberal

humanism has depended on a common notion of what the human is and what human

society should be, based on a Hobbesian correspondence between the individual body and

[46] Marvin Minsky, "Why Computer Science is the Most Important Thing that Has Happened to the Humanities in 5,000 Years," lecture, Nara, Japan, 15 May 1996. Quoted in Hayles, *Posthuman* 244.

[47] Hans Moravec, *Mind Children: The Future of Robot and Human Intelligence* (Cambridge, MA: Harvard UP, 1988) 147-59.

[48] David Noble, *The Religion of Technology: The Divinity of Man and the Spirit of Invention* (New York: Penguin, 1999) 3.

[49] Ray Kurzweil, *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* (New York: Viking-Penguin, 1999).

the body politic. As Allucquère Rosanne Stone notes, "social beings, people, exist by virtue of possessing biological bodies through which their existence is warranted in the body politic."[50] But what about people who inhabit multiple avatars, or intelligences whose "brains" are made from microchips rather than cells?

One immediate response to such posthumanism is to argue that it does not represent a total juncture from the past. N. Katherine Hayles' preeminent *How We Became Posthuman* demonstrates how posthumanism continues the legacies of liberal humanism rather than breaking from it systematically.[51] Another response argues for the value of embodied cognition. Neuroscientists such as Antonio Damasio stress the holistic role of the embodied brain in thinking, such that the reproduction of mind alone is insufficient to replicate the human experience of living in the world anchored in a corporeal centre (a position advocated by A.S. Byatt in *A Whistling Woman*, which I study in chapter 9).[52] Roger Penrose argues that consciousness is the product of physical, quantum mechanisms that cannot be simulated through other, non-quantum media such as circuit boards.[53]

The ultimate alternative to cybernetic transcendentalism is to deconstruct its possibility at its epistemological roots. In his Chinese Room experiment, for example, John Searle contends that in principle human consciousness cannot be reproduced in a machine because the translation process is itself a shift from the self-consciousness that characterises the human; just as syntax is not semantics, to be able to mimic thinking through a program does not equate to actually thinking.[54] More generally, there is an apparent disconnect between the "easy" problem of consciousness (the observation of

---

[50] Allucquère Rosanne Stone, *The War of Desire and Technology at the Close of the Mechanical Age* (Cambridge, MA: MIT Press, 1995) 63.

[51] She observes, for example, that "the erasure of embodiment is a feature common to *both* the liberal humanist subject and the cybernetic posthuman. Identified with the rational mind, the liberal subject *possessed* a body but was not usually represented as *being* a body. Only because the body is not identified with the self is it possible to claim for the liberal subject its notorious universality, a claim that depends on erasing markers of bodily difference, including sex, race, and ethnicity." Hayles, *Posthuman* 4.

[52] Antonio Damasio, *Descartes' Error: Emotion, Reason and the Human Brain* (London: Picador, 1995); *The Feeling of What Happens* (London: Vintage, 2000).

[53] Roger Penrose, *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics* (1989; Oxford: Oxford UP, 1999).

[54] John Searle, "Minds, Brains and Programs," *Behavioral and Brain Sciences* 3.3 (1980): 417-57. He expanded his critique in *Minds, Brains and Science: The 1984 Reith Lectures* (London: British Broadcasting Corporation, 1984).

neuro-chemical states and the implementation of specific rules in artificial intelligence)

and the "hard" problem (how those discrete states give rise to the variety of conscious

experience).[55] Endorsing this, the fact that artificial intelligence has failed to develop even

a decade after Turing predicted it would emerge might suggest that the human mind is

intrinsically unavailable for simulation. Perhaps the most famous critic in this line remains

Herbert Dreyfus, with his book *What Computers Can't Do*.[56] He contends that visions such

as Minsky's deflect attention away from the way in which cybernetics is changing humans

in the present, with man's malleable nature ensuring that even if machines cannot think

like humans, "People have begun to think of themselves as objects able to fit into the

inflexible calculations of disembodied machines." Rather than the computer providing an

analogy for the human mind, the human mind is squeezed to accommodate the operational

interfaces and limitations of the machine; we risk "not the advent of superintelligent

computers, but of subintelligent human beings."[57] In all these examples, spiritual

speculations about the possibility of instantiating mind in a different medium are wishful

thinking that not only misunderstands the nature of consciousness, but also deflects

attention from the moral crises facing humanism in the present in its attempt to negotiate

the boundaries between man and machine.

How, then, is this dialectic between transcendence and embodiment figured by

literary fiction, with its underlying demonic poetics. For on the one hand, the literary text is

a physical artefact – that is, a body of prose – whose words intend to convey a limited

number of semantic meanings. However, this mechanical language can be deployed in

such a way as to lead to indeterminate "virtual" effects in the subjective consciousness of a

reader who may draw different meanings out of the same body of text. William Gibson,

who inaugurated the cyberpunk genre with his 1984 novel, *Neuromancer*, seems to have

understood computers through the sort of double perspective also typified in demonic

possession. Watching from the sidelines of an amusement arcade, standing outside of the

---

[55] David J. Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford UP, 1996) xi-xiii.

[56] The first edition of the book was first published in 1972 but, an indication of its ongoing importance, it was republished in 1979 and 1992 as Hubert L. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge, MA: MIT Press, 1992).

[57] Dreyfus 192.

system, Gibson saw something disconcerting in the close intertwining of a player and game:

> I could see in the physical intensity of their postures how *rapt* the kids inside were. It was like one of those closed systems out of a Pynchon novel: a feedback loop with photons coming off the screens into the kid's eyes, neurons moving through their bodies, and electrons moving through the video game. These kids clearly *believed* in the space games projected.[58]

As I suggested earlier, seeing possession from the outside can be disconcerting, precisely because we have no way of knowing what that experience is like from the perspective of the possessed. However, Gibson's use of the word "rapt" (or rapture) appropriately acknowledges the depth of the player's investment in the game in a way that, as for the man married in a MUD or as for the medieval poet possessed by the spirit of God, can seem highly valuable from their own point of view.

The cyberpunk genre inscribes this ambivalence. By its stylistic novelty, cyberpunk challenges existing literary standards, mimicking the way in which in the novels themselves, the computer becomes the tool of liberation, freeing the talented hacker from the physical constraints and poverty imposed upon him by an unfair capitalist system. Gibson typically describes hackers as "cowboys," representing them as uneasy outsiders, stuck ambiguously as neither heroes nor anti-heroes in a way which might also pertain to the sub-genre itself, which dissolves, recombines and pastiches the body of literary conventions in unusual ways. This is certainly evident in the epochal moment when Gibson introduces cyberspace (his neologism) as:

> A consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical concepts...A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding...[59]

On the surface, the passage is disconcerting in its representation of posthumanity, appearing aesthetically radical in its new verbal currency ("nonspace," "cyberspace"), and

---

[58] William Gibson, interview with Larry McCaffery, *Storming the Reality Studio: A Casebook of Cyberpunk and Postmodern Science Fiction*, ed. Larry McCaffery (Durham, NC: Duke UP, 1991) 272. Emphasis in original.

[59] William Gibson, *Neuromancer* (1984; London: HarperCollins, 1994) 51.

its verbless clauses. At this level, people become disturbingly mechanistic. Since the word computer originally referred to a real person calculating astronomical tables, does "every computer in the human system" refer to the digital computers in the networks set up by humans, or are humans themselves become computers?[60] The reference to mathematics lends a double meaning to the word "operator," which can also describe the Boolean logic operators (AND/OR/NOR) that are the fundamental constituents of programming. Thus the children are not only taught mathematical concepts via the cyberspace media, but rather they are the very logical filters (the "operators") through which information must pass, instantiating the system at root rather than simply interacting with it. But this interpretation of the human as machine which results from a literal interpretation of the words is balanced by the way in which such words combine in cognitively appealing ways. In common with cybernetic transcendentalism, the physical reduction of the human as part of the immaterial network is a necessary price for the attainment of a stellar sort of religious ecstasy. This mode of existence as information tenders the experience of "lines of light" with data forming "clusters and constellations." Taken literally, this might seem to describe the actual data display of a computer monitor. However, the alliteration in phrases like "lines of light" resonantly appeals to a poetic inner voice, ensuring that the meaning of the passage is not confined to the superficial referent of the printed word. In the uncanny likeness between poetic inspiration and cybernetic subjectivity, this poetic quality also constitutes a "consensual hallucination," though unlike the literal hallucination of cyberspace this implicit one deceives by the tonal quality of language. Objectively, Gibson sees the arcade player as tied to the system. However, he acknowledges that from the subjective point of view of the one within the loop, such an experience can be rapturous. Likewise, whilst in the moment of the reading experience we may be deterred by the vocabulary of mechanism and mathematics, the way these words are stylishly combined simultaneously induces persuasively poetic images. Having said that, in an unresolved oscillation between human embodiment and posthuman disembodiment that characterises the novel as a whole, the passage then pulls back from the brink of complete immersion in a hyperbolic cyberspace. Though "constellations" cues a prototypical analogy with stars that would be appropriate to such religious transcendence, Gibson instead makes a

---

[60] The Oxford English Dictionary traces the word computer back to Richard Braithwaite's *The Yong Mans Gleanings* of 1613: "I haue read the truest computer of Times, and the best Arithmetician that euer breathed, and he reduceth thy dayes into a short number." "Computer," *The Oxford English Dictionary.* 2nd ed. 1989.

metaphorical connection to the urban and artificial: rather than stars, the constellations are "like city lights, receding." This shows how style can again be deployed in order to figure the dissolution of embodiment in a different context. In this case, Gibson breaches the boundaries of a canonical body (cowboys and cybernetics, constellations and city lights do not normally appear together) and of syntax, the removal of verbs in a stream-of-consciousness effect. But not only does this style figure the breakdown of the conventional human subject as opposed to the posthuman, stylistic novelty is also something strongly associated with the more traditional idea of inspiration and literary genius I outlined earlier, in which again the key feature is the poet or writer's willingness to give over their body (the hand that writes) to a voice that originates from somewhere beyond themselves, thereby producing something original, otherworldy and *sui generis*. Tonally ebbing and flowing between material meanings and sensual impression, like embodied humans passing as informational constructs in cyberspace, the passage marks the cybernetic tensions between body and mind, mechanism and aesthetics, that resist reduction or one single determined meaning. Such a tension is also apparent in demonic experiences.

Whilst it might be tempting to see the dialectic in *Neuromancer* as analogous to the broader "two cultures" fracture between the sciences and the arts, it is important to note that because of the way cybernetics is intrinsically interdisciplinary, interpretations and representations of posthumanism do not necessarily divide along this line. Gibson's observations and writing encapsulate the double possibility of subjective and objective modes of perception. Likewise Norbert Wiener heralded the cybernetic epistemology that represented humans as information whilst simultaneously speaking up for traditional humanist values of respect for the body (individual and politic).[61] Whilst Wiener foresaw a future in which mechanical slaves would be able to perform the economic work without the moral conflicts involved in keeping human slaves, he admitted that, just as the first industrial revolution devalued the work of the craftsman, the second is "bound to devalue the human brain." With Belsen and Hiroshima prominently in mind, Wiener admitted that the new science "embraces technical developments with great possibilities for good and evil."[62] Rather than anti-science, Wiener's keynote, like the double trajectory of

---

[61] Hayles makes the case for Wiener as a humanist who almost reluctantly instigated the era of the posthuman: Wiener's writings "testify to both the exhilaration and the uneasiness that cybernetics generated when its boundary disruptions threatened to get out of hand." Hayles, *Posthuman* 84-112.

[62] Wiener, *Cybernetics* 28.

cybernetics I mapped above, requires that the analyst adopt a Janus-face. Being a reactionary Luddite against the science and technology of the cybernetic age, ignores the possibilities they entail for rethinking humanism. Such Luddism fails to elaborate the conditions and effects with which science and technology are inevitably changing the human – like the computer games' player – in the present. An analysis of fictional metaphors which gets away from constructing the cybernetic as simply the gothic or monstrous other to the human might help to capture the ambiguous, multifaceted and systemic way in which the existence of computers in the world has already changed the nature of being human, whether for good or ill. As Veronica Hollinger argues, we need to "deconstruct the human/machine opposition and begin to ask new questions about the ways in which we and our technologies 'interface' to produce what has become a *mutual evolution.*"[63] I take interfaces here to refer not just to the one connecting the computer user to the world behind the screen, but to all those that cross disciplines, simulations and consciousnesses: the interface between art and science; the relationship between body and mind; the connection or disconnection between words on a page and meanings in a mind. It is its place at the interface – the juncture between the objective, outer world and the inner world of subjective consciousness – that makes the demon so appropriately situated in the cybernetic condition. Unlike the monster, the demon (whose nature is precisely to change natures from within so that perception of reality outside the self is changed) provides a trope which engages primarily with questions of interface rather than category, of epistemology rather than ontology. Demons rely on the open mind of their possessee, but must also be interpreted diversely: even if the final authenticity of their message cannot be guaranteed because the demon is a simulator (or dissimulator), demons – like computers – demand that we interpret the relationship that exists between them and us. As Norbert Wiener saw it, what marked cybernetics as a new discipline was that it focused not on the intrinsic nature of things but on the systems of feedback and communication that regulate the behaviour of any entity, whether organic, mechanical or human. To recover Wiener's ground between the poles of Minsky and Dreyfus, and to continue the cybernetic tradition of attending to the systems of feedback rather than the ontology of the posthuman's otherness, we need to read between the lines of novels like *Neuromancer* and the ambiguous effects it produces in the minds of its readers. We need to focus not on the monsters but the demonic fictions of cybernetics.

[63] Veronica Hollinger, "Cybernetic Deconstructions: Cyberpunk and Postmodernism," *Mosaic* 23.2 (1990): 42.

Ontologies and Interfaces: The Monster and the Demon

Reading some of the claims of computer scientists, it is not difficult to see why cybernetics has facilitated much apocalyptic science fiction. The cyberneticist Hans Moravec, for example, opens his prophetic study *Mind Children* by proclaiming that "our genes have finally outsmarted themselves. They have produced a weapon so powerful it will vanquish the losers and winners alike." The "weapon" in question is artificial intelligence.[64] The post-biological future that Moravec envisions could be repurposed easily into a Promethean narrative, in which the overt promise of technology also entails its unforeseen hazard. In the film *I Robot*, for example (derived from Isaac Asimov's short story collection), robots attack humans because of a logical conflict in the first rule of their programming: "No robot may harm humanity, or through inaction, allow humanity to come to harm."[65] As the robots realise unethical humans are the greatest threat to their own species, they must be eradicated by the superior machines they originally created. This self-cancelling future has often been given added potency by linking cybernetics to technologies which allow otherwise disincarnate artificial intelligences to wreak physical havoc. Two of the most significant have been nuclear weapons (as in *War Games* or *The Terminator*) and genetic modification (*Jurassic Park*).[66] As the Terminator glowers with its hulking metallic frame and glowing red eyes, or as the inquisitive velociraptors are unleashed by the theme park's sabotaged computer system, cybernetic fictions slot within a prototypical scheme in which monsters trope the social impact of science. It is certainly legitimate to pay attention to these sorts of ontologies – the difference between human, machine, and monster.[67] Indeed, that I will do so throughout this thesis suggests that it is

---

[64] Moravec 1.

[65] Isaac Asimov, *I, Robot* (1967; London: Voyager-Harper, 1996) ix; *I, Robot*, dir. Alex Proyas, perf. Will Smith, Bridget Moynahan, Alan Tudyk, James Cromwell and Bruce Greenwood, 20th Century Fox, 2004.

[66] *War Games*, dir. John Badham, perf. Matthew Broderick, Dabney Coleman, John Wood, and Ally Sheedy, 1983, DVD, MGM Entertainment, 2000; *The Terminator*, dir. James Cameron, perf. Arnold Schwarzenegger, Michael Biehn, Linda Hamilton and Paul Winfield, 1984, DVD, MGM Entertainment, 2007; *Jurassic Park*, dir. Steven Spielberg, perf. Sam Neill, Laura Dern, Jeff Goldblum and Richard Attenborough, 1992, videocassette, Universal, 1992.

[67] Jon Turney, *In Frankenstein's Footsteps: Science, Genetics and Popular Culture* (New Haven, CT: Yale UP, 1998); Kingsley Amis saw robots and androids as descendents

not possible to think about cybernetics without referring to constructions of alterity. However, by concentrating on the outer nature of malevolent artificial intelligences or man-machine hybrids we risk overlooking the current mutual exchanges between cybernetics and the human in the present. By looking at fictions that treat cybernetics in terms of demonism, I hope to show how cybernetics also conditions the historical experience of being human in a world whose reality has always seemed sometimes difficult to comprehend or define, even through such traditional media as narrative.

To an unusual degree in the fictions I examine, the technologies of narrative are intertwined with the scientific outcomes the plot is intended to apprehend or critique, and this complication lends itself to a demonic analysis. The contrast between the ur-text of Enlightenment monsters, Mary Shelley's *Frankenstein*, and Shelley Jackson's hypertext novel, *Patchwork Girl*, illuminates this clearly.[68] Mary Shelley's novel is itself a bibliographic hybrid, commingling and layering voices and genres, and thereby offering different perspectives on the monster's true self, which remains evasive. However, its condemnatory characterisation of Victor Frankenstein means the readers are absolved of responsibility for creating the monster in any cognitive sense. The monster is, rather, seen as the result of a specific (if hazily defined) technological process contained within the fiction. However, cybernetic fictions regularly acknowledge the relationship between their own material textuality (or immaterial hypertextuality) and the cybernetic paradigm out of which they are constructed. *Patchwork Girl* is an adaptation of Mary Shelley's original and L. Frank Baum's *The Patchwork Girl of Oz*, punningly authored by a hybrid identity, Mary/Shelley (as in Mary Shelley or Shelley Jackson). Through hyperlinks, Mary/Shelley weaves ambiguous relationships between authors and their creations, and between the textual originals and the derivative *Patchwork Girl* which also, again ambiguously, refers to the name of the main protagonist. Clicking the word "hop," for example, leads to Patchwork Girl describing herself as "a discontinuous trace, a dotted line"; the link "sewn" generates "I had sewn her, stitching deep into the night by candlelight, until the tiny black

---

of Mary Shelley's original monster in his *New Maps of Hell: A Survey of Science Fiction* (London: Gollancz, 1962) 33.

[68] Mary Shelley, *Frankenstein, or, the Modern Prometheus*, ed. M.K. Joseph (1818; Oxford: Oxford UP, 1998); Shelley Jackson, *Patchwork Girl by Mary Shelley and Herself*, CD-ROM (Watertown, MA: Eastgate Systems, 1995).

stitches wavered into script and I began to feel that I was writing." As well as marking the strange relationships between feminism, sewing, writing and bodies in its narrative (so far as that term can apply to a story with many possible paths), *Patchwork Girl* goes a step further than *Frankenstein* in that not only does the existing hybrid text reflect the constitution of the created monster, but the process of reading *Patchwork Girl* through selecting a hypertextual route is directly the process of birthing the "monster" (who is, actually, not at all monstrous). Unlike in the original, here the reader or web browser is directly responsible for creating the other. As N. Katherine Hayles observes, "As the unified subject is thus broken apart and reassembled as a multiplicity, *Patchwork Girl* also highlights the technologies that make the textual body itself a multiplicity."[69] The hypertextual technique exploits the relationship between the message and the medium, making true Marshall McLuhan's formula that the medium is the message.[70] But can such cybernetic models also inform fictions whose bodies are essentially intact and conventional, offering a generally coherent narrative structure?

Interpreting Heidegger's call for art and literature to reveal the way that technology frames the phenomenal universe within structures of utility, David Porush asserts that:

> Cybernetic fiction is precisely that mode of expression most suited to reveal the essence of technology: at once quite akin to technology itself, since the texts are formed by the applications, ironic or otherwise, of algorithms borrowed from cybernetics – and at the same time most suited to revealing the truth.[71]

Cybernetics has become such a powerful episteme for thinking about the nature of humanity and semiotics that even texts which are not apparently directly about cybernetics can nevertheless be inextricably informed by it, even produced through its mechanisms. That even these texts – which are not necessarily generic science fiction but are fiction informed by the science – can reveal the essence of technology indicates how deeply cybernetics has affected the general interests and representative forms of postmodern narrative culture. This, then, explains why, with the exception of *System Shock* and *Patchwork Girl*, I have chosen not to consider cybernetic literature explicitly produced or

---

[69] N. Katherine Hayles, *My Mother Was a Computer: Digital Subjectivities and Literary Texts* (Chicago: U of Chicago P, 2005) 151.

[70] Marshall McLuhan, *Understanding Media: The Extensions of Man* (London: Routledge, 1964) 9.

[71] Porush 83.

generated by that technology (film does engage more closely with it, as explained below), and to avoid looking at the most directly contingent relationships between cybernetics and literature, namely the changing nature of textuality in an era of hypertext and computer games.[72] For in this contingent relationship, we would naturally expect that the essence of technology is revealed by engaging with explicit images of such artefacts, whether in the form of Shodan's character or the monster that sits at the centre of *Patchwork Girl*'s web of links. The nature of cybernetics here is, to a degree, always going to be self-evident if, as Sherry Turkle argues, computers are inherently self-reflexive foci inviting us to consider the nature of human psychology, even when they do not run any explicit narrative of artificial intelligence. But even when technology is not used directly and explicitly to provoke the reader to reflect on the relationship between narrative and scientific craft, the demon provides a metaphor that mediates the effects of technology into otherwise conventionally-structured narratives. Whether represented in print or film, the demonic sustains the possibility that the identity being perceived or critiqued is also that which allows such perception to happen in the first place; in other words, it binds narrative technique to the scientific other that the style is designed to apprehend, so that a monster-myth of technology would illogically deconstruct the very aesthetics that produces it.

I am here using demon (or, strictly, daemon) in the Platonic sense. The Platonic universe is tripartite, inhabited by gods, demons and men in the regions of Heaven, air and earth respectively. Daemons are required to mediate between the metaphysical and the physical, conveying prayers and sacrifices, and rewards and messages, to and from men and the gods.[73] By the later Hellenistic period the Platonic definition had been submerged beneath a pejorative view of the demon as malevolent, something confirmed by Jewish monotheism which translated the demon into Hebrew as "unclean spirit."[74] However, the daemonic possibility never went away, being transmuted in Christian tradition into the guise of the angel, which inhabited a similar sublunary sphere with agency between fallen

---

[72] There is a growing body of theory in relation to etexts, hypertextuality, and computer games as literature. Two of the most influential works in the field from within conventional literary studies are Espen J. Aarseth, *Cybertext: Perspectives on Ergodic Literature* (Baltimore: Johns Hopkins UP, 1997) and Jerome McGann, *Radiant Textuality: Literature After the World Wide Web* (New York: Palgrave, 2001).

[73] Plato, *The Symposium*, trans. Walter Hamilton (Harmondsworth: Penguin, 1951) 81

[74] L.F. Hartman, "Demon (in the Bible)," *The New Catholic Encyclopedia*, ed. Berard L. Marthaler, 2nd ed. vol. 4 (Detroit: Gale, 2003) 645-46.

man and God (inversely, the pagan gods were degraded as being demons); as well as witchcraft or the devil, possession could be associated with the sort of Apostolic revelation that occurred through the Holy Spirit, a vein that continued in Protestant tradition (most vocally in the Quakers or Muggletons). Though the close folk association between demonism and witchcraft may have ensured that the demon vernacularly implied evil, recent works like Philip Pullman's *His Dark Materials* trilogy suggest that the Platonic daemon can still maintain its amorality rather than immorality, being simply a technique or medium for introducing a metaphysical (or, indeed, virtual) experience to the receptive self.[75] Indeed, in a literal sense there probably are ambivalent daemons lurking on your very own computer, as in programming terminology a daemon refers to a subroutine that is activated when certain conditions are met.[76] When email appears demonically to takeover your life, remember that the "mailer-daemons" that route them on servers are really ambivalent! The fact that the demon is not clearly evil makes it open to complex ethical interpretations, entirely fitting to ideas of simulation with which it is engaged with on a double level. The negative connotation potentially applies to Shodan or Jackson's modern Frankenstein, but a conflict arises as computers act as the daemonic intermediary for these games and hypertextual narratives. This daemonic/demonic tension is most clear in cyberfictions and computer games, but it is also inherent in relation to literary inspiration and our engagement with literary narratives. However, it is in film narratives that the daemonic/demonic interplay is more apparent.

It is noticeable that many of the plays with illusion in the early modern period invoked demons in order to draw attention to the implicitly daemonic nature of the hallucinatory device. For example, the Jesuit museum curator, showman and scholar, Athanasius Kircher, constructed a magic lantern which projected images of demons on to a screen.[77] As Marina Warner suggests, the supernatural subject matter posited a strong association between the mind's eye or the dream, and the powers of the new machines of

[75] Philip Pullman, *His Dark Materials* (London: Scholastic P, 2001).

[76] The *OED* defines daemon as "A program (or part of a program), esp. within a Unix system, which runs in the background without intervention by the user, either continuously or only when automatically activated by a particular event or condition." "daemon," *The Oxford English Dictionary*, 2nd ed., 1989.

[77] For a description of the museum, see Paula Findlen, "Scientific Spectacle in Baroque Rome: Athanasius Kircher and the Roman College Museum," *Roma moderna e contemporanea* 3 (1995): 625-65.

illusion.[78] Kircher's magic lantern used the demonic to draw attention to its daemonically mediating power. Or, rather, the power that ultimately resided with the scientist, for in this dynamic the devilish entity which could generate illusion in the mind of the observer was tamed as a daemonic thing, since unlike true possession the device could be switched off at its master's discretion. Projecting and throwing the grotesque away from the observer and the inventor, such images and deceits reinstate the hierarchy of the controlling human self over the arbitrary world of demonic deception, to differentiate between man and (super)nature, since the latter has been turned and manipulated to the former's ludic ends. Inheriting the legacy of this mastery of illusion, the daemonic/demonic tension is most evident in cinematic fictions, because the figuration of a malevolent cybernetic technology is produced through the medium of the camera lens and, often, digital special effects. As with Kircher's magic lantern, *Forbidden Planet*, *2001: A Space Odyssey*, *Blade Runner* and *The Matrix* reflect on their daemonic medium, but do so in order to highlight the ironical fact that they are critical of technology whilst simultaneously using it visually to realise a possible future world for their audiences. In line with postmodern theory, the self-reflection on the artifice of cinematic illusion-making is designed to point to the impossibility of constituting a world outside of representation or language, of objectively witnessing reality in a way not framed or situated through the control of an author or director. Films like *The Matrix* project the demonic possibility on screen not in order to tame it, but to propose that we really could be living in a deceptive world of images, impostures and illusions, with no way of realising this through rational thought. After all, cinematic science has produced the special effects which allow us to suspend our disbelief in filmic science fiction, as we hallucinate future worlds where, for a brief time, monstrous cyborgs appear really to exist.

My reading of *Neuromancer* briefly demonstrated the ways in which the literal meanings of words may be made ambiguous when vocabulary from one genre unexpectedly is juxtaposed with another. The subjective effects of hearing their poetic, alliterative quality is in tension with their literal representation of people as unemotional objects or machines, such that the empirical body of the prose conflicts with the psychological effects that result from it. The daemonic mechanism of text results in an uncertain demonics of reading. Although such daemonic/demonic maps of textual mechanisms and critiques are more difficult to analyse in written fiction compared to the

---

[78] Marina Warner, *Phantasmagoria: Spirit Visions, Metaphors and Media into the Twenty-First Century* (New York: Oxford UP, 2006) 139.

computer game or cinematic fiction, they are certainly present. In Umberto Eco's *Foucault's Pendulum*, for example, paranoid postmodernists produce through a computer the random intertextual connections that produce a compelling trail of clues about a secret conspiracy. This leads ultimately to the demonic delusion that beneath the semblance of reality lies a syncretic Plan. Just as Shelley Jackson's hypertext causes the reader to constitute the monster, *Foucault's Pendulum* implicates the computer-like writers and readers in the process of creating the strange subversions of reality within it. Oddly, given that, explicitly, cybernetic narratives contain messages about the duplicitous effects of cybernetic media, they also remind that words do not mimetically reflect the world of signifieds, but that meanings and reality are continually distorted by interpretation. The deceptions and manipulations of narrative modes are, therefore, uncanny analogues to the artificial simulations, grotesque artificial intelligences, or disturbing android doubles presented within them.

Admittedly, however, almost all these fictions fail to explicitly acknowledge the daemonic/demonic distinction, instead taking the vernacular term, demon, as their overt metaphor. What sort of creature is the more conventional demon? As the prehistory of the demon has taken so many forms, defining the demon in any rigid terms becomes itself a demonically obsessive activity. Arguing against the Platonic tradition in which the demon was an "intermediate nature" between earth and heaven,[79] Augustine contended in his fifth-century *The City of God* that although demons do exist, their reason for existence is not to negotiate between god and man but to deceive and lie, causing man to swerve from making true and direct supplication to God himself.[80] Influential though *The City of God* was, though, it did not prevent the Renaissance rediscovery of the Hermetic system in which the arts of the unseen can cause physical effects on the world in reality (a belief taken up by the conspirators in *Foucault's Pendulum*). Later, after the Reformation, deciding whether the demon was a tangible force or merely a false belief fed into Catholic and Protestant debates. For the former, the demon indicated the effectiveness of ritual (in the exorcism) and the importance of signs of the work of God and of the priest as his agent. For the Protestant, the demon might indicate excessive Popery, and of more import was the

---

[79] Plato, *The Symposium* 81.

[80] Augustine, *The City of God Against the Pagans*, ed. and trans. R.W. Dyson, Cambridge Texts in the History of Political Thought (Cambridge: Cambridge UP, 1998) especially 332-58. A good gloss on this passage is provided by G.R. Evans, *Augustine on Evil* (Cambridge: Cambridge UP, 1982) 98-111.

quiet spontaneity of personal worship (though, oddly, the inspiration of the spirit could still assume a demonic character). For my purposes, it suffices to suggest that, like the simple computer icon that points to an associated program, the demon is a shortcut trope that produces one possibility above all: that of being deceived, so that reality as we know it becomes suspended behind an illusion, albeit one that might ultimately reveal some more fundamental truth behind reality. Tellingly, the most important demon in thinking about the relationship between science and the self invokes this possibility. Imagining that experience is an illusion implanted by deceiving demons, Descartes realises that regardless of the origin of sensations, the fact that there is an identity doing the experiencing cannot be doubted. Across chapters 4 to 7, I argue that the rejection of Descartes' idea that it is possible to break through delusion through rational self-reflection is a key to postmodern philosophy and narrative. Postmodern fictions take up the possibility of demonic deceit whilst rejecting the possibility of empirical foundationalism that results from seeing through the delusion. The idea of the demon as a deceptive agent inhabiting the mind of the self and making perceptions of "truth" and "the real" arbitrary meshes with a postmodern belief that the world cannot transparently be known, except through the distorting lens of representational signifiers. When these representations are inhabited by artificial intelligences (in *2001*), androids and replicants (*Do Androids Dream of Electric Sheep?* and *Blade Runner*) and near-perfect virtual simulations (*The Matrix*), perception is made doubly-difficult.

The demon is a simulator whose true nature is hard to determine once and for all, even as it determines the perceptions of the possessed. Catholic treatises on dealing with devils warned the priest against engaging in discourse with the demon; the devil should only be commanded to identify itself by name, and to explain why it entered the body, when it intends to leave, and how its departure will be indicated. But, in practice, it seems that the very fact of asking the devil its identity authenticated for the demoniac that there was a devil inside, so that a vicious cycle of interrogation and legitimation of its presence kept pushing the final exorcism ever further away.[81] Thus the demon was essentially constituted through the simulation of language, projecting a "vocalic body" by ventriloquising through the demoniac.[82] An uncanny comparison between Jean Baudrillard

[81] D.P. Walker, *Unclean Spirits: Possession and Exorcism in France and England in the Late Sixteenth and Early Seventeenth Centuries* (London: Scolar-Bemrose, 1981) 8.

[82] Steven Connor, *Dumbstruck: A Cultural History of Ventriloquism* (Oxford: Oxford UP, 2000).

and a case of possession in his French compatriots four centuries previously strongly implies that demonism and postmodernism are highly compatible, having to do with the ability of language to constitute delusion so convincingly that it becomes impossible to separate the true from the false. If the reality of the demon becomes that which is produced through language, then when postmodernism argues that "reality" is essentially a product of discursively mediated thoughts, reality as a whole becomes demonic.

In 1634, in the French town of Loudun, a fractious priest, Urbain Grandier, was burned at the stake for having caused the demonic possession of a group of Ursuline nuns. A sexually attractive and politically well-connected figure, Grandier appears to have offended the prioress of the town's convent, Jeanne des Agnes; he also rubbed against various other bourgeois families in the town, particularly after an illegitimate relationship with the daughter of the public prosecutor. In what began as a prank and ended with Grandier's execution, the nuns staged their own possession, confessing that their demons were brought by Grandier's devilish intervention. Taking place within a wider anti-Huguenotism under the purview of Cardinal Richelieu, Grandier was finally executed. However, the nuns and prioress who were initially encouraged merely to simulate possession became so bound up in producing the expected signifiers of demonism that they genuinely believed themselves to be possessed, according to all the codes for interpreting such signs. The attention and money their case drew – as well as the fact that it spurred many of the townspeople to convert back to Catholicism – caused them to sustain the possession long after Grandier's death. Similarly, the exorcists Father Tranquille, Father Lactance and Jean-Joseph Surin, and Grandier's torturer and inquisitor Dr. Mannouri, became obsessed with Grandier's apparent retention of faith to the last and their own role in the hoax, such that they all soon deliriously believed that they were themselves possessed by demons. The problem here is that the forgery of possession and the true possession slide into each other all too easily, because it is impossible – by the very nature of the demon as existing within the body of another and projecting his existence from that position – to verify whether that other is truly possessed through any empirical method. As Steven Connor summarises:

> The one simulating the signs of demonic possession is hard to distinguish
> from the real victim of demonic possession, because possession is an *act* of

simulation; the one who pretends to have a devil, pretends to have been occupied by a spirit whose nature it is to pretend to existence.[83]

Producing all the true symptoms of demonism, the nuns of Loudun could not be described objectively as possessed or merely simulating possession because the symptoms in the first case have the identical qualities of the fake, of speaking in the voice of another. The double irony here is that the book which popularised the case, Aldous Huxley's *The Devils of Loudun*, is a sort of hoax condemning the hoax written by an author famous for his experiments in pharmacological hallucination.[84] *The Devils of Loudun* is a fictional history grounded in the archival documents, but using a great deal of the novelist's interpolations to explain the psychological motivation of individual characters.[85]

The quotation from Connor, the historical example of Loudun, and Huxley's quasi-fiction, set up striking harmonies with one of Jean Baudrillard's examples of a simulacrum. The daemoniacs of Loudun might well be diagnosed as prescient postmoderns, fitting as they do into Baudrillard's example of a person feigning illness, whose very act of feigning could itself be considered a syndrome:

> Is the simulator sick or not, given that he produces "true" symptoms?...For if any symptom can be "produced," and can no longer be taken as a fact of nature, then every illness can be considered as simulatable and simulated, and medicine loses its meaning since it only knows how to treat "real" illnesses according to their objective causes.[86]

If the demon renders complex the relationship between the self and objective reality, then clearly it resonates with the sensibilities of the postmodern, such as Baudrillard's pervasive simulacra. The argument that language constructs a personal reality rather than mimetically reflecting it is embodied in Loudun's confessors who became subject to their own stories of Grandier's possession, which were then rewritten by a twentieth-century novelist. Whilst the detachment of the self from the language it produces was positively inflected by

---

[83] Steven Connor, "Toward a New Demonology," *Becoming Human: New Perspectives on the Inhuman Condition*, ed. Paul Sheehan (Westport, CA: Praeger, 2003) 104.

[84] Aldous Huxley, *The Doors of Perception and Heaven and Hell* (New York: Perennial, 2004).

[85] Aldous Huxley, *The Devils of Loudun* (London: Chatto, 1961).

[86] Jean Baudrillard, *Simulacra and Simulation*, trans. Sheila Faria Glaser (1981; Ann Arbor, MI: U of Michigan P, 1994).

the cult of inspiration discussed earlier, it is given such a benign spin because all deceptions end in a logos. For the Medieval scribe, the words for which he was medium stemmed from, and pointed to, an absolute God. For the Romantic, the powerful veracity revealed by inspiration was that the ego could transcend the material conditions of its culture. Descartes, having gone through demonic deception, emerges at the other side to know one thing with less doubt than before, the truth of his own experienced existence. Postmodernism is more pessimistic than this set of interrelated beliefs. For a postmodernist like Baudrillard, nothing stands outside the textual or ideological mode of its production. Truth is essentially unverifiable, only temporarily assuming status due to the way it is conceptualised in any given culture. As Baudrillard might have argued, the demon in Loudun was a construct of language, as evidenced by the way the demons were perpetuated through the discourse of the exorcists; in a different culture, as that from which Huxley writes about the case, we might be more inclined to see demons as the products of the Freudian unconscious. Such cross-pollination between the demon and the postmodern imply that questions about subjectivity and knowledge are immutable, rather than consigned to a specific period as it is determined by science and technology.

Demonic Fictions

Indeed, it is this cross-pollination that provides some justification for my choice of illustrative cybernetic fictions. One of the characteristics of postmodernism has been the collapse of distinctions between high and low art forms. As a consequence, science fiction – which was (rightly or wrongly) distinguished from modernism in the pre-War period – comes into contact with postmodern fictions, which are informed both by theory as well as their own internal poetics. As a consequence, as Brian McHale argues, there is a feedback loop in which science fiction "derives certain of its elements from postmodernist mainstream fiction which itself has, in its turn, already been 'science fictionized' to some greater or lesser degree."[87] Though particularly true of cyberpunk, this interplay between "literary" and sub-genre fiction (such as science fiction) characterises the postmodern period generally. To look, then, at canonical postmodernists like Thomas Pynchon, John Barth, or William Burroughs, though they are certainly informed by cybernetic epistemes, risks downplaying the more popular and more influential genres in which cybernetic and postmodern ideas are mediated: literary science fiction, computer games and, perhaps most

[87] Brian McHale, "POSTcyberMODERNpunkISM," *Constructing Postmodernism* (Routledge: London, 1992) 229.

importantly, cinema. Because cybernetics has permeated culture at both a technological and philosophical level, we ought to be willing to discuss the way all fictions relate to the cybernetic age, not just those that have attached to them some anterior academic kudos. I have opened by looking at a computer game, and I will close in my final chapter by examining Dean Koontz's pulp thriller, *Demon Seed*; in several chapters I examine popular, but also highly influential, Hollywood science fiction; and whilst writers such as Umberto Eco and A.S. Byatt are undoubtedly intellectual authors, they have an uneasy relationship to the academy. At once working with academic literary theory and estranged from it, they fear that literary theory is potentially alienating, claiming to offer a science of texts couched in a rhetorical style that is far removed from that of the fictional works it appropriates. However, because the idea of computation is familiar to most readers in contemporary culture, its metaphors allow authors or directors to embody and encode ideas about literary creation and semiotics through analogy rather than direct philosophical discourse.[88] As Sherry Turkle reports, students who are deterred by the difficulty of Derrida can still claim to understand the principles of deconstruction by experimenting with hypertext.[89] In whatever genre and regardless of any prior prestige attached to them, contemporary fictions that engage with cybernetics recognise the pervasive nature of the technology. They use it as a common focus by which to explore and refresh philosophical problems that prove (because of their demonic bent) to be longstanding, rather than determined by a technological or philosophical moment: language and representation; the mind-body problem; what is fundamentally human as opposed to posthuman, monster or cyborg. Before moving into the fictions themselves, however, the first section of this thesis aims to historicise and universalise the cybernetic paradigm.

The next chapter places the demon as conceptualised in postmodernity in relation to its theological history. Donna Haraway's "Cyborg Manifesto" argues that the cyborg is so tightly integrated with technologies that it shows – in accordance with postmodern thought – that the objective and the subjective are false dichotomies. Because a cyborg has no essential, stable body, the categories of the real and the virtual, true and false or, especially, female nature and male culture, break down. By juxtaposing Haraway's postmodern work

---

[88] Though I will not cover it specifically here, David Lodge's *Thinks* is a case in point, in that the romance between a novelist and a cognitive scientist is essentially the excuse for a humorous but didactic analysis of the "two cultures" problem as it relates to contemporary philosophy of mind in academia. David Lodge, *Thinks* (London: Penguin, 2002).

[89] Turkle, *Life on the Screen* 17-8.

with that of a Renaissance specialist on teratology, the French anatomist Ambroise Paré, I propose that the biological monsters of that age figured comparable anxieties about crossing boundaries of subjective and objective phenomena, and hence potentially false and generally true. Paré attempts to interpret monsters as signs or allegories of God's intent. However, whilst monsters demonstrate the creativity of God's imagination, Paré avoids dealing with demons at length, because they imply that the world is supernatural, and can never be known empirically and hence read allegorically. Monsters are weird things within nature that are nevertheless signs of God; demons are outside of nature, and the tendency to believe in demons indicates man's willingness to be seduced by illusions. By controlling the terms with which the demon is said to exist, Paré makes it into a boundary marker distinguishing fact from fiction. Whilst I do not aim primarily to push the advent of cybernetics back in time (such work having already been conducted elsewhere), by engaging cybernetics with the demon, I will historicise the science by default.[90] By treating cybernetics in conjunction with the demonic trope, I suggest that the phenomenological issues about the relationship between the inner mind and the outer reality thrown up by modern science are comparable with those which existed under earlier, theological conditions in which the demon was said to exist. Chapter 2 infers that the philosophies that underpin cybernetics have an old heritage, and that the ways in which we appreciate our relationship to these modern technologies are translations of concerns about the relationship between an artificial or otherworldly agent and the mind that might be affected by it. The demonic has proved to be a trans-historical force, and the presence of the demonic in contemporary fiction engaged with cybernetics implies that thinking about possession or deception by simulations is not the preserve of the digital age alone. Rather, the metaphorical idea of the demon appeals to our tacit sense of being in, and of struggling to know, the real world.

Chapter 3 exemplifies this by showing how psychoanalysis explains that demons originate in unconscious motivations. Psychoanalysis makes the demon available as a metaphor which is no longer metaphysically significant, but which still describes the experience of being human in the world. Rather than marking the boundary between the objectively true and the subjectively false, the demons of the unconscious imply that we

[90] For examples of the historicisation of cybernetics, see Neil Rhodes and Jonathan Sawday, eds., *The Renaissance Computer: Knowledge Technology in the First Age of Print* (London: Routledge-Taylor, 2000); Ray Siemens and Susan Schriebman, eds., *A Companion to Digital Literary Studies* (Malden, MA: Blackwell, 2007), especially 3-159.

construct our perception of the world through projecting our own latent desires and repressions upon its canvas. Monsters perceived to be "out there" are in reality signifiers of the deceptive pressures exerted from within our own minds looking out unreliably onto the world. As postmodern aesthetics takes up – in a way exemplified by Gabriel Josipovici's notion of the "demon of analogy" – if language is inherently deceiving, cleaving reality from the word, rather than connectively allegorical, then in the new environment allegorical monsters must be treated as symbols of personal solipsism rather than iconic others to the human. Through reading the seminal science fiction film *Forbidden Planet* (1956) against Freud's essays on *A Seventeenth-Century Demonological Neurosis* and "The 'Uncanny'," I show how cybernetic technologies uncover our personal demons by explaining the nature of mind in material terms. Reworking Shakespeare's *The Tempest*, the Caliban figure in *Forbidden Planet* is not the product of some immoral supernaturalism, but emanates from the unconscious primitivism of the scientist as it is mediated through a mind-reading technology. Here the technology provides a metaphor for the human: just as a robot is subject to programming, so the conscious mind can never be in control of a self which is actually determined by evolved psychological codes. To allegorise technological power as monstrous is to fail to realise that technology is only acting as the daemonic mediator of our minds, in which the real demons lie.

The narrative value of the monster is that it can be critiqued, even as it invites us to see in it a reflection of our own nature or potential. The demon, however, makes questionable the perception of reality, so that such judgements of difference between the other and the self are *a priori* problematic. Broadly speaking, the difference between the monsters and demons is one of ontology versus epistemology. Exposing this distinction very well is Descartes' deceiving demon. In chapter 4, I propose that the Turing test is essentially a Cartesian test for the truth that can emerge from an apparent (dis)simulation. Just as Descartes derives rationalism by starting from the potential deceptions of demons, so Turing lays out the terms of his thought experiment so as to maintain a realist epistemology when confronted with artificially intelligent mechanisms that appear human. This, however, causes problems for our understanding of the ontology of machines that appear to be conscious, but whose minds we can never quite inhabit in order to judge the truth of their being. In Turing's test, we can only know consciousness through intentional expressions, particularly language, emanating from it. As with demonic possession, the whole nature of consciousness is that it is so complex that understanding the material specifications of another mind does not allow us to know what it is like to *be* another from the phenomenological point of view (Thomas Nagel's argument in his classic essay "What

Is it Like to Be a Bat?").[91] Thus the simulation of consciousness is hard to distinguish from true consciousness, and we cannot know for sure whether a Turing machine is genuinely intelligent. Indeed, human language could itself be the source of a self-sustaining deception, rather than resolving the difference between man and machine. This epistemological barrier means that it is possible that from within the mind of an android or artificial intelligence the world might be perceived comparably to our own, and hence similar sorts of moral criteria should apply to artificial as to human consciousness. As the following chapters assert, it is here that fiction can at least provide imaginative access to the world as if through the eyes of another, even a cybernetic double, and allow us to understand the questions of likeness from the point of view of the cybernetic other.

The historical and philosophical boundaries sketched in chapters 2 to 4 suggest that encounters with cybernetic doubles raise epistemological issues, since cyborgs and artificial intelligences initiate perceptual shifts in the human, rather than proving to be definitively abject. Consequently, fictions that take up cybernetic concerns lend themselves to the kind of sceptical doubt about reality that characterises postmodern thought. Across chapters 5 to 7, I use seminal science fiction films to illustrate the advance of postmodern epistemologies, at the expense of Cartesian and Turing-type rationality.

In chapter 5, I show how the interpretation of *2001: A Space Odyssey*'s Hal as a monster contradicts the broader aesthetics of the narrative, which operates through what I term a poetics of possession. This aesthetics forces the film's viewers to see the world through different eyes: ape, man, star child, and machine. Hal's possession of a subjective point of view is fundamental to the allegory of *2001*, which according to this revisionist reading is intended not to raise ontological questions about the difference between a machine based mind and an embodied human one, but epistemological problems: why do we worry about difference at all, if mind is the telos of universal evolution? And if mind matters above embodiment, is it justified to disconnect Hal, whilst allowing Bowman, a cyborg spaceman, to be elevated as the "star child"?

Similarly, in Philip K. Dick's *Do Androids Dream of Electric Sheep?* (examined in chapter 6) distinctions between human and android are not based on essentialist qualities of embodiment or origin. There is nothing, on the surface, to distinguish for sure between the android and the human, a problem Dick articulates extensively in his essays and lectures. Dick argues, however, that the theory of mind, the ability to imagine the world as if

[91] Thomas Nagel, "What Is it Like to Be a Bat," *The Philosophical Review* 4 (1974): 435-50.

through the mind of another, is the factor that distinguishes humans (or fully conscious androids) from affectless automata (or humans who have become androids in their unthinking way of living in the modern world). If we fail to consider the potential humanity of androids, we become like them ourselves. However, the ability to empathise is itself a deception, requiring us to suspend our own reality by simulating another consciousness through the artifice of text. Just as second-wave cybernetics was concerned with the idea of reflexivity, whereby "that which has been used to generate a system is made, through a changed perspective, to become part of the system it generates,"[92] the reader is drawn in to perceive the world as if through the eyes of an android, thereby perpetuating the deception.

Interestingly, as they share comparable beliefs about how it is possible to know what an artificial double is thinking, *2001* and *Do Androids Dream of Electric Sheep?* were both published in 1968. These two works provide a benchmark by which we can compare the way in which Ridley Scott's 1982 adaptation of Dick's novel, the film *Blade Runner*, turns Turing's concern at the level of the individual – how is it possible to know the difference between an android and a human? – into more general questions about the ontology of the world in general – is reality itself objectively out there, or is reality in a sense constructed by the mind of the perceiver? Dick shows that distinctions between android and human can be found through the latter's ability to empathise, even if this means suspending our own view of the world to see through the eyes of an android double. In a Baudrillardian vein, however, Scott eliminates the mirror, suggesting that image and original are essentially commensurate. Any philosophy of knowledge or any grounds for distinction between the real individual and the artificial replicant are irrelevant, since reality is anyway inseparable from the simulation, because both are constructs of discourse and representation. In the context of the hyperreal, therefore, we cannot prove our humanity by our ability to jump empathetically from the domain of the human to that of the artificial mind, because the human and the simulacrum are essentially identical.

As I explore in more detail in *The Matrix* (chapter 7), which draws extensively upon Baudrillard's notion of the simulacrum, the demon is a metaphor for a postmodern condition in which the self cannot know the truth of his own perceptions of reality through objective, Cartesian means. But rather than couching such experiences in the heady prose of postmodern theory – of which Baudrillard's own *Simulacra and Simulation* is a prime

---

[92] Hayles, *Posthuman* 8.

example – the computer provides an accessible metaphor through which popular culture can examine the philosophical problem of how to establish the "truth" of the world in which we are embedded. *The Matrix* reconsiders the Cartesian epistemological framework of deception in the context of an information economy in which technologies like cinema disguise the fact that we are purposiveless agents in a broader commodity system of late capitalism. But whilst seemingly a Marxist critique, *The Matrix* illustrates the problem of analysis in postmodernism. In late capitalism, the product of exchange is information rather than tangible goods, which means that culture too becomes complicit in the economy. So in *The Matrix* the economic system which the film condemns by presenting it as a demonically deceiving artificial intelligence is also that which allows this critique daemonically to be rendered – through high tech computer graphics – in the first place. Objectivism is thus made impossible because of this representational conflict, and it is this denial of any external perspective for which the demon provides an epistemological trope.

In the final three chapters, I move from cinematic science fiction to look at more conventional literary novels. These represent a drift in the contemporary canon away from an old "two cultures" binary, crossing the same sorts of disciplinary boundary lines that are infringed by cybernetics. These are not science fictions, though following McHale's feedback loop they have been "science fictionized" by the intercourse with that genre via the cinematic fictions above. Consequently, they cannot avoid being informed by science, even if they are not determined by it. On the other hand, these fictions, whose authors are uneasily related to academic literary theory, also qualify the postmodern, anti-scientific conviction that no meaning or truth can ever be absolute. They turn to a science of fiction, to resist the theorisation of fiction.

Umberto Eco and A.S. Byatt assume that their readers (who may be informed by literary theories) are willing and able to perform sustained reflection on the text, especially those which draw attention to their own textual conditions through metafiction. Unlike cinema, which we may watch passively and less critically, by reflecting on the way our minds are being manipulated and mediated by literary mechanisms, we are able to see that fiction is doing something different to the real world, so that by inference reality cannot itself be a total simulation, as it is in *The Matrix*. Peter Singer observes that "Ethical truths are not written into the fabric of the universe: to that extent the subjectivist is correct...on the other hand, once there are beings with desires, there are values that are not only the

subjective values of each individual."[93] In the same way, Eco and Byatt ask whether the demon can be rejected as a total figment of the mind in the secular era. As social values, shared narratives or tropes – of which the demon is one – surely must assume some sort of objectivity, even if this may be counterintuitive to an age of secular science? As Freud observed, the wholesale rejection of demons in a world whose myths have been purged by science is problematic when contemporary demons have been created in the mind of a reader simply through words on a page. Even if the paranormal is not at work in physical reality, the demon can be said to exist in the metaphorical ontology of the novel, but also to translate into the human mind of its reader – and since minds are the natural residence of the demon, can demons not be said still to exist as they always have done, just that they are produced by a structured fiction and human culture rather than an unseen devil? Warren McCulloch once said that the whole ambition of cybernetics was to answer the question "What is a number that a man may know it: and what is a man that he may know a number?"[94] These fictions attend to the first element by asking "what is a story, that it may be known by a man?" By drawing attention to their narrative mechanisms through metafiction, and by setting the irrational creation of literary meaning against the concept of computer-generated texts, they suggest that literature can defy both the determinism of cybernetic sciences and the semiotic pretence that the real is always evasive, a simulation of discourse rather than a tangible thing inhabited by an embodied self. For Eco and Byatt, the demon acts as a lens by which we can reflect on the mechanisms of text, and the way a text mediates the objective world into a subjective one in the readerly consciousness, and the way language can be used to concretise a meaning about the experiential world.

In *Foucault's Pendulum*, we follow a group of conspiratorial academics who perceive occult significance in the associations and links randomly generated by their computer, Abulafia. But Eco makes us aware that authors – including himself – are themselves writing machines (machines that write, writing about machines), seducing readers into conventions or conspiracies about what a sign or text means. The occult arts of Hermeticism contend that a single word can lead to an exponential suite of connections and meanings. This historical example meets with the concept of a TALE-SPIN sort of automatic writing machine (the Abulafia computer) able to churn out associations at

---

[93] Peter Singer, *How Are We to Live: Ethics in an Age of Self Interest* (Oxford: Opus-Oxford UP, 1997) 275.

[94] Warren McCulloch, *Embodiments of Mind* (1965; Cambridge, MA: MIT Press, 1988) 2.

random to produce the plot of the novel. Ironically, it is precisely because they are arbitrary that they must therefore embody a meaningful truth about the way language works in an eternal play of signification deferred from final allegorical meaning. Eco's novel, however, ultimately encourages us to reject much of the irrationality that characterises postmodernism and Hermeticism alike. He argues that though texts may be mechanisms for producing multiple interpretations, this does not mean that all interpretations are as good or bad as another, as his postmodern conspirators believe. Rather, by invoking Shannon's information theory, Eco requires careful and sceptical readers to discriminate between irrelevant noise and significant codes of limited semiosis.

Whilst Eco invites us to see his characters as paranoid individuals who take postmodernist and Hermetic thought to an extreme, the ability to empathise with their plight even if we do not share it is remarkable. Indeed, for Philip K. Dick, in *Do Androids Dream of Electric Sheep?*, the ability to empathise with an android is, weirdly, the proof of our humanity, our sympathetic warmth that our mechanical doubles lack. Similarly, for A.S. Byatt, the mind of a schizophrenic possessed by demons deserves our attention. She causes us to enter that other consciousness through language, which is something neither cognitive science nor artificial intelligence can achieve through their reductionist methods. Through attending to the phenomenological effects of language, her novel *A Whistling Woman* shows through that the algorithmic model of mind extrapolated by cognitive scientists cannot give rise to the manifest, fantastic consciousness of this psychiatric patient, or, for that matter, any imaginative mind, such as that of an inquisitive reader. Cognitive scientists in the novel ponder where exactly in the mind metaphors and images are created. A similar question is asked of the demon as it occurs in the text of the literary novel.

My final chapter proper, on A.S. Byatt's *Possession*, provides a coda that recalls many of the themes of this thesis. Superficially, the novel appears to be concerned with the relationship between past and present, and between postmodern criticism and the inexplicable cult that surrounds a creative artefact as opposed to a facsimile or reproduction. However, in line with my argument that cybernetics has become a dominant theme of post-war fiction such that even novels not concerned directly with science are necessarily informed by it, the novel's superficial interests are based on underlying cybernetic epistemologies, including genetics, information theory and chaos theory. Whilst acknowledging that the novel warns against imposing meanings on a work, this chapter tests how far it is possible to push the hypothesis that the occasional references to science

and technology in *Possession* realise that postmodernist culture is the product of a sub-culture of cybernetics.[95]

Matter Matters

I opened this chapter with a semi-fictional account. I close with a true story. In 2006, the Briton Neil Entwistle killed his wife and daughter, shattering what had seemed, on the surface, to be their tranquil family existence in suburban America. At his trial, it emerged that Entwistle had lived a double existence. He had cultivated an online persona as a successful web entrepreneur, as compensation for his anxiety that he was originally a simple miner's son from Yorkshire. Escaping to his native England after the event, Entwistle explained in a phone conversation with an American police officer that he had not even cried for the deaths of his wife and daughter: "I think it's because I'm here. It almost doesn't seem real. It's just a void."[96] The risk of cybernetic games is total immersion in the virtual world beyond the screen. By contrast, a response to literature and, in a different manner, film takes place with a critical distance, as it is conducted with a double-consciousness. For in literature and film we must simultaneously suspend our disbelief and inhabit the fictional world *as if* it is real, whilst we must also maintain some connection to reality in order to recognise it *as* a created fiction, which by the virtue of its special form can tell us something unique about reality that we cannot otherwise access in the everyday world. It is learning how to enjoy the simulation with a double consciousness – remembering the real even as we enter the fictional – that postmodern fictions can help us with. Though it would be ridiculous to claim they can prevent murder or extreme neurosis, with their characteristic self-reflexivity and scepticism as to the truth of the reality we perceive, postmodern fictions can help us to realise the importance of understanding the way in which our subjectivity is constructed and affected by language, and correspondingly also through cybernetic interfaces, and to remember the devilish stakes of losing the reality game.

[95] Though I have discussed demons and technology at length in the present chapter, I have barely referred to the affinity between telepathy and telegraphy in the Victorian period. This is because extensive material has already been written on this specific historical topic. One paradigmatic study is Pamela Thurschwell, *Literature, Technology and Magical Thinking, 1880-1920* (Cambridge: Cambridge UP, 2001).

[96] Jonathan Raban, "Just Two Clicks," *London Review of Books* 30.16 (2008): 3-9.

Chapter 2. Demons and Cyborgs: Ambroise Paré and Donna Haraway

The Re-Enchantment of Nature

Cybernetics has been accompanied by predictable anxieties that the human is reduced to an essentialist machine, whilst monstrous but superior technologies may assume our mantle as the most developed form of life. On the other hand, cybernetic transcendentalism encapsulates a spiritual belief that such technologies will enable us to escape from the body entirely, transcending nature itself. Caught awkwardly between advancement and regression, the posthuman risks becoming immersed in an informational existence that relates with ethical uncertainty to the embodiment of the liberal humanist. The way in which the demon figures the anxieties and possibilities of the cybernetic sciences suggests that the information age has led to the recovery of pre-Enlightenment frameworks of knowledge and selfhood which allow space for emotions, beliefs, even spirit and soul, in the comprehension of the world. As Bruno Latour has observed, the separation of nonhuman nature and human culture in the Enlightenment is a conceptual divide. In actuality, the anthropological matrix consists of "hybrid actants" which show the binaries of subject and object, or artistic representation and objective facts, to be problematic dichotomies, since they straddle and move between both categories.[1] The artificiality of this constitution has, Latour argues, been exposed by a contemporary era inhabited by technological hybrids: genetically modified organisms, test tube babies, and cyborgs. As a consequence, Latour concludes, we can be said never to have been modern in the strong sense implied by the Enlightenment model of knowledge that segregates nature and culture, whilst seeking to transcend the former through technology.

This chapter shows how cybernetics has involved the recuperation of pre-modern themes, particularly the idea of demonic possession. This provides the historical bedrock against which to bounce analyses of the metaphor in fictions examined in subsequent chapters. Here, I show that Donna Haraway's postmodern idea of the cyborg as an agent that embodies both science and emotion marks a return to pre-Enlightenment concepts according to which empirical knowledge is situated in relation to a subjective observer, rather than abstracted by the objective, passionless witness honoured by rationalism. With a doctorate in biology, Haraway is not wholly against scientific notions of the self as being in part a material construct. However, she also argues that such rational and biological

---

[1] Bruno Latour, *We Have Never Been Modern*, trans. Catherine Porter (Hertfordshire: Harvester Wheatsheaf, 1993) 6.

models as genetics are not incompatible with more flexible notions of the self as something produced by nurture and social discourse. In a similar way, the science of monsters in the Renaissance represents an attempt to show both the unity of God's creation as it operates according to fixed and immutable laws, but also its fecundity and diversity which must provoke an abstract sense of wonder in the power of the divine. The interplay between demons and monsters in the Renaissance shows that good knowledge – which is to say knowledge that shows the beneficent mind of God – relies on allowing the student of nature to perceive it as reflecting God's morality (a cultural idea) whilst also operating according to certain fixed laws (a rational concept) by which the demon's power to deceive is necessarily proscribed.

A Manifesto for Monsters

Although I have been positing a phenomenological affinity between virtual environments and demonic or spiritualistic experiences, it is nevertheless remarkable to discover that the first proper logical computer, Charles Babbage's Analytical Engine, was used to demonstrate the probability of miracles.[2] Like many scientists in the Enlightenment tradition, the pioneer of the modern computer believed that science and technology offered a means to reach closer to the nature and will of God, with the engineer or scientist assuming the role of a second Adam:

> the advancement of man in the knowledge of the structure of the works of the Creator, might furnish continually increasing proofs of its authenticity; and...thus by the due employment of our faculties, we might not merely redeem revelation from the ravages of time, but give to it a degree of force strengthening with every accession to our knowledge.[3]

According to Enlightenment ideology, the mind of man reflects that of God; for Descartes, indeed, it was the fact that our minds are made in the image of God's that we can be confident in our ability to perceive the world objectively. It is ironic, therefore, that

---

[2] Babbage showed that his calculating machine could be used to produce a continuous sequence of numbers from 1 to 100 000 001, at which point the machine would suddenly switch to add 10 000 to make 100 010 002. To a naïve observer, such a change would appear miraculous, though the miracle would have been produced in accordance with a rational law, in this case the programmer's initial setting of the machine. In a similar way the creation of higher animals out of an apparently predictable order of vegetable forms should be considered a miracle, albeit a naturally effected one. Charles Babbage, *The Ninth Bridgewater Treatise: A Fragment*, 2nd ed. (London: John Murray, 1838) 34-44.

[3] Babbage 139.

Babbage's discipline would develop in such a way as to break such axiomatic connections between science and the divine. By perceiving nature as the instantiation of information in its own right, cybernetics – like evolutionary science before it – undercuts the status of the world as a book or code in which can be read the work of a divine designer beyond the physical.[4] Rather than Babbage's dualistic explanation in which the mind of man is like that of God, with matter something to be rejected in the ascension to heaven, cybernetics perceives all matter as essentially instantiations of non-physical information. The "human" simply happens to be one physiological incarnation of mind, brought about through the evolutionary lottery of genetic information, rather than moulded from nature in an idolic image of the divine. Man is placed in a quantifiable, informational relationship to other lower organisms whose genes he shares, rather than in an anthropic hierarchy that reflects his special status in God's creation. In the cybernetic age, the vertical chain that connects a creator to nature and human culture is broken, and a new, secular manifesto is required to take its place. Such a manifesto must try to accept the findings of modern science and acknowledge that we are the biological outcome of a dispassionate process of genetic evolution, and that our most distinguishing feature as a species, consciousness, is ultimately just a cybernetic system like any other. At the same time, it must construct a notion of the human which reflects the fact that we still intuitively feel more than the sum of genetic or cognitive algorithms, even in an ostensibly secularised age.

Arguably the most prominent document to take up this challenge from a postmodern perspective is Donna Haraway's 1985 feminist critique, "A Cyborg Manifesto."[5] Here Haraway identifies a tension between objects (which are both the material products of capitalism, and the theoretical entities encountered by science) and

---

[4] Though this view can still be countered, most recently through the movement of Intelligent Design.

[5] Donna Haraway, "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century," *Simians, Cyborgs, and Women: The Reinvention of Nature* (New York: Routledge, 1991) 149-81, rpt. of "Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s," *Socialist Review* 80 (1985): 65-108. Although Haraway's essay has a feminist intent, attempting to find a middle way between biological essentialism in constructions of gender, and social constructivism that denies any common ground for identity, in my juxtaposition of Haraway's essay with Renaissance concepts of the demon, I will largely bypass this feminist angle. Instead, I will focus more on the way in which the trope of the "cyborg" mediates between nature and nurture at a level which tells us much about the way in which cybernetic science is open to ideas of the demon, and thus to the re-enchantment of nature by allowing a role for the emotions in constructing knowledge.

subjects (as this term applies both to the servants of capitalism, and to the individual, gendered identities who are erased by science's strong claims to objectivity). Informed by social constructivism, subjects and objects are fused together in Haraway's ironic trope of the cyborg:

> A cyborg is a cybernetic organism, a hybrid of machine and organism, a creature of social reality as well as a creature of fiction. Social reality is lived social relations, our most important political construction, a world-changing fiction...Liberation rests on the construction of the consciousness, the imaginative apprehension, of oppression, and so of possibility. The cyborg is a matter of fiction and lived experience that changes what counts as women's experience in the late twentieth century. This is a struggle over life and death, but the boundary between science fiction and social reality is an optical illusion.[6]

In her attempt to get away from the binary thinking that separates the male and the female or object and subject, Haraway looks to the dualism of the organic and inorganic which has become destabilised by cybernetic science and technology. As Haraway argues, "By the late twentieth century...we are all chimeras, theorized and fabricated hybrids of machine and organism; in short, we are cyborgs."[7] Her chimera is characteristic of the Janus-face of posthumanism presented in chapter 1. On the optimistic side, the cyborg citizen might be able to play or subvert traditional identities of gender or race by assuming through genetic engineering or cyberspace avatars an alternative identity that disguises their embodiment by which their difference might be discerned. The cyborg makes us aware of the "partial, fluid, sometimes aspect of sex and sexual experience," thereby deconstructing the biological essentialism so fundamental to male-dominated history.[8] On the other hand, such a notion reminds us of the fantasies of the monstrous other which have contributed to the subjugation of the woman. Ideally, the cyborg falsifies the dichotomy between the objective and subjective, in which woman was denoted as nature and body, man as objective mind of science and culture. However, rather than erasing subjectivity as scientific method does, the cyborg knows that all knowledge is situated from the point of view of an embodied observer. Knowing their own being to be conditioned by external

---

[6] Haraway, "Cyborg Manifesto" 149.

[7] Haraway, "Cyborg Manifesto" 149.

[8] Haraway, "Cyborg Manifesto" 180.

technologies such as drugs, enhanced vision, ultrasonic scans, and computer interfaces, such an agent will be forced to take direct responsibility for the applications of science, acting as a "modest witness," a term Haraway develops in her later work.[9] Haraway's cyborg emerges at a time when cybernetics argues that we are constructs of ephemeral information rather than possessing an innate and stable physical nature, a belief that is given added emphasis by the wave of social constructivism emerging from the humanities which argues that categories such as "woman" are ideological notions rather than referring to any biological ontology of gender. However, Haraway appears to want to keep embodiment in the notion of the self, because this both accords with our egocentric belief that humans are special in comparison to the rest of nature, and because the world of nature is always seen from the point of view of an individual being within it.

Without denying that cybernetics establishes innovative epistemologies and technological creatures, I want in this chapter to measure the degree to which Haraway's cyborg can be said to be wholly revolutionary. It is certainly only recently that the cyborg or biological hybrid has become materialisable through our appreciation of the common codes (genetic, cybernetic, neurological) that allow the intermingling of species, or of organisms with technology. But as N. Katherine Hayles points out, we do not have to be literal cyborgs, with technology intervening upon our bodies, in order to become posthuman, since "new models of subjectivity emerging from such fields as cognitive science and artificial life imply that even a biologically unaltered *Homo sapiens* counts as posthuman. The defining characteristics involve the construction of subjectivity, not the presence of nonbiological components."[10] Indeed, in referring to science fiction Haraway herself acknowledges that the cyborg is a concept rather than an actual entity of the present. If, therefore, it is the conceptual rather than physical boundaries between humans and technology that affect subjectivity in the posthuman era, then even in an age of less powerful technologies we might find other theoretical conditions that similarly cause constructions of human subjectivity to be predicated on virtual experiences. As "a fiction mapping our social and bodily reality and as an imaginative resource suggesting some very fruitful couplings," Haraway's cyborg is one example of the chimera crossing nature and

---

[9] Donna J. Haraway, *Modest_Witness@Second_Millenium. FemaleMan©_Meets_OncoMouse™: Feminism and Technoscience* (New York: Routledge, 1997).

[10] Hayles, *Posthuman* 4.

culture to which Latour draws attention.[11] However, when Haraway proposes that recent technoscience offers something "vastly different from the constitutional arrangements that established the separations of nature and society proper to 'modernity' as early modern Europeans and their offspring understood that historical configuration," is the paradigm really so radical as it appears?[12] By finding conceptual ancestors to the cyborg, to what extent can we be said to have never been modern, to paraphrase Latour, and hence not to be radically *post*human?

Possible comparators to the postmodern cyborg may be found in early modern teratology. Teratology covers monsters, demons, mutant births, hybrids, strange creatures from distant lands – all entities that traverse the boundaries distinguishing science and culture. In the physico-theological diagram, the status of the monster in a natural world that supposedly images God's mind and purposes is clearly problematic, and debates about how to situate monsters also involve the tricky negotiation of epistemological and theological limits. Renaissance writers seeking to understand the role of monsters, freaks and marvels were forced to interrogate the relationship between the book of nature inscribed with God's moral purposes, and the world of Fallen human culture, which might cause monstrosity to come about. To the more theologically-educated mind, for example, things treated pejoratively as "monsters" might actually reveal the hidden and intricate structure of their Creator's mind. The derivation of the word *monstrum* from *monere* – to warn or threaten – indicates that the monster was not simply a signifier of deformity in itself, but rather could be a sign or portent of some event to come or of divine will. In other contexts, though, monsters could be perceived as the consequence of human immorality rather than deriving from the will of God; in particular a thing can be wrongly perceived as monstrous through the demonic implantation of delusion in the beholder. Although Haraway asserts that "The cyborg is our ontology; it gives us our politics,"[13] monsters in the Renaissance hence provided similarly politicised bodies which connected the subjective world of human culture to nature, and hence to the divine creator.

[11] Haraway, "Cyborg Manifesto" 150.

[12] Haraway, *Modest_Witness* 43. Haraway is using "constitution" following Latour's definition.

[13] Haraway, "Cyborg Manifesto" 150.

One exemplary work on the subject in Renaissance science is offered by the 1573 collection *On Monsters and Marvels*, by Ambroise Paré.[14] As chief surgeon to both Charles IX and Henri III, Paré was the most prominent medical man of his era. The view that nature is constructed and truth fabricated, rather than out there waiting to be discovered and articulated, is key to postmodern constructions of science. It is only within such an epistemological framework that the cyborg can be conceptualised as a political, as well as a technological, agent. However, even in Paré's period we can perceive an anxiety about the reality of monsters. Paré is aware that things which are superficially freakish but nevertheless rationally explicable might be falsely construed as "monstrous" by a non-sceptical observer with a superstitious mind possessed by demons. The epistemological concern that truths about nature might be forged by the human imagination rather than existing through God's design to be revealed through natural science is evidenced by Paré's differentiation between monsters, which are part of the theatre of nature, and demons, which still operate according to natural laws but which to the unenlightened perceiver may appear to be supernatural and therefore beyond God's control. Paré understands that the presence of demons in the normal world might suggest that "reality" is subjectively concocted by the imagination. Science (as then defined) and the moral and theological interpretations that result from it might thus be an inevitably defeated project. Thus Paré explains that demons can intervene in nature only by modifying the perspective of the ill-educated observer, not by physically creating monsters themselves through supernatural causes. In parallel, God's contract with man (his *potentia Dei ordinata*) agrees to limit the role of his potentially infinite creative power (*potentia Dei absoluta*) so that though in theory he could perform any act, in practice God tends to operate according to fixed natural laws of causality.[15] The structure of Paré's compendium is designed to locate

[14] Ambroise Paré, *On Monsters and Marvels*, trans. and introd. Janis L. Pallister (Chicago: U of Chicago P, 1982). All subsequent parenthetical page references in this chapter refer to this edition and translation. The definitive French edition is Ambroise Paré, *Des Monstres et Prodiges*, ed. Jean Céard (Geneva: Librairie Droz, 1971).

[15] *Potentia Dei ordinata* is thus a solution to the old theological problem of how an omnipotent and omniscient being can nevertheless permit people free will, as most famously expressed by Boethius: "If God foresees all things and cannot be mistaken in any way, what Providence has foreseen as a future event must happen. So that if from eternity Providence foreknows not only men's actions but also their thoughts and desires, there will be no freedom of will. No action or desire will be able to exist other than that which God's infallible Providence has foreseen. For if they were changed and made different from how they were foreseen, there will be no sure foreknowledge of the future, only an uncertain opinion; and this I do not think can be believed of God." Boethius, *The Consolation of*

the individual "scientist" in relation to nature and to contain its fecundity in accordance with God's covenant with man to limit his almighty potential. In the same way, Haraway's cyborg models the ideal modern scientist as a modest witness who is aware of their hybrid standpoint across both nature and culture.

Collecting Nature

The Renaissance theory of generation followed Aristotle: all living beings were divinely formed, with the female contributing matter to the embryo, whilst the male endows it with form and vitality.[16] Aristotle evacuates any portentous signification from the monster, with the monster simply a thing that goes against the general case whilst not being against nature entirely. Monsters are caused by the failure of the generative male seed to dominate the material secretion of the female.[17] Aristotle's aetiology intersected the Augustinian tradition in which monsters were evidence of the creative power of a purposive God. Although benign, and therefore not monstrous in the vernacular sense of the abject, their strangeness did embody significance. As Augustine's influential aesthetic of *concordia discors* argued, "the beauty of this world [is] enhanced by the opposition of contraries, composed, as it were, by an eloquence not of words, but of things."[18] This sort of view would be inherited by the likes of Babbage as he delved deep into the structure of the world through science. But the problem before the scientific revolution was how it was possible to tie an autonomous aetiological development (if monsters are created as part of nature, rather than against nature in the supernaturalist tradition) to their divine portentousness which implies that God himself has allowed a deformity to come about. For

---

Philosophy, trans. and introd. Victor Watts, revised ed. (London: Penguin, 1999) 119-20. Amos Funkenstein argues that modern science emerged from the secularisation of the concept of *potentia Dei ordinata*. Amos Funkenstein, *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century* (Princeton: Princeton UP, 1986) 129-50.

[16] Michael Hagner, "Enlightened Monsters," *The Sciences in Enlightened Europe*, ed. William Clark, Jan Golinski, and Simon Schaffer (Chicago: U of Chicago P, 1999) 175-217.

[17] Aristotle, *Generation of Animals*, trans. A. L. Peck (London: Heinemann, 1943) 417.

[18] Augustine 472.

whenever monsters are accounted for in terms of natural rather than metaphysical generation, the puzzle becomes one of their significance. If monsters are not *monstra* – in the sense of showing or warning of some event to come – what exactly do they signify? If nothing, why then do they exist at all in a divinely ordered nature?[19]

If these issues provided opportunity for debate, some epistemic stability could be found in the practice of classification and collection of monsters. As Thomas Kaufmann describes, the *Kunstkammer*, or cabinet of curiosities, provided a symbolic theatre of the world in which by systematising nature for display the collector equated command of knowledge with political power.[20] By incorporating monstrosities and deformities into the collections of scholars and physicians, monsters could at least be naturalised. Civilised and refined viewers were aware that they were a problematic part of nature, rather than the result of any devilish generation from the supernatural sphere, outside of nature and thus beyond the control of God. For example, Peter I of Russia had one of the most extensive collections of the seventeenth century. He argued that "the unlearned hide [monsters] because they believe that such births originate from the devil by witchcraft and sodomy. But this is impossible because the creator of all living things is God and not the devil, who has no power over any individual."[21] Though demarcated as different, the monster need not necessarily be monstrous, simply a deviation within the boundaries of normal expectation; demons, however, though their evil may highlight the good by way of contrast, also signify an unnatural magic whose forces work to separate man from true knowledge of his creator. By distinguishing weird but explicable monsters from demons or the work of the devil, the rationalistic spectator could try to make the monster seem, by comparison, more normative, and hence within God's plan. The practice of displaying monsters in books or cases implicitly deconstructed the possibility of occult generation, such that objects in a collection lost their alarming potential and, as Michael Hagner puts it, "represented their own subordination."[22] Inversely, as Zakiya Hanafi explains, the domination of nature through the *theatrum naturae* exhaustively placed "all things in relation to each other and

[19] Zakiya Hanafi, *The Monster in the Machine: Magic, Medicine, and the Marvelous in the Time of the Scientific Revolution* (Durham, NC: Duke UP, 2000) 27.

[20] Thomas DaCosta Kaufmann, "From Mastery of the World to Mastery of Nature: The *Kunstkammer*, Politics, and Science," *The Mastery of Nature: Aspects of Art, Science, and Humanism in the Renaissance* (Princeton, NJ: Princeton UP, 1993) 174-94.

[21] Qtd. in Hagner 183.

[22] Hagner 186.

the whole, with the exception of one creature: [Man], standing at the centre."[23] The modern scientific method implies an erasure of subjectivity. In the Newtonian universe, for example, there is no causal relationship between the observer and the observed so that the second law of motion holds for any observer of any physical object. By contrast, as Hanafi explains, the encyclopaedic method constitutes a personal science, allowing Man to spatialise his hierarchical primacy through bringing the monstrous into proximity with himself, whilst simultaneously keeping it at a safely objective distance through the glass of the cabinet, specimen jar, or anatomical drawing. Thus the Renaissance viewer of monsters occupies a position not unlike Haraway's cyborg. The cyborg is heterogeneous, able to occupy different perspectives on nature through applying different augmentations or accessing alternative interfaces. In this sense, as Haraway argues, as in the scientific method and representation, the individual transcends their personal subjective standpoint to reach a universal truth. On the other hand, such heterogeneity reminds that nature is always mediated and constructed through a particular lens: his or her view of the world from the standpoint of a body (or bodies) affects the perception of nature out there. Similarly, the Renaissance scientist denied claims that the monster could be supernaturally generated by collecting them and contextualising each specimen against others. At the same time, they were self-aware that such a structure put themselves at the centre of the natural universe in the way Hanafi argues. Thus the monster was also a lens or sign through which could be read the fact that God beyond nature and time had made Man so special in the first place, the only one able to collect the monsters that allow Man to spatialise the hierarchy of Aristotle's *scala naturae*. Through producing a coherent body of knowledge out of the multiple bodies of the monstrous, the naturalist ensured a situated science that located his embodied self in relation to God's unseen purposes.

Paré's Monsters

Typifying this collective mode, rather than initiating his own studies of individual cases or exploring a single case in detail Ambroise Paré compiled from other sources. Since its style and content represent a synthesis of other treatises on the subject, *On Monsters* is one of the paradigmatic books on monsters of the Renaissance. Following Augustine's lead, the rare and unusual forms in the book prove the creative variety of God's universe, speaking of His wondrous fecundity. In this sense, there is nothing

---

[23] Hanafi 5

grotesque in nature, except as perceived and considered as such by the irrational observer. Thus Paré's first definition is between monsters which "are things that appear outside the course of Nature" and marvels which "are things which happen that are completely against Nature" (3). As with other collections (such as Peter I's), the structure of the book demarcates the grotesque that can nevertheless be naturalised by containing it within the hermetic environment of the collative edition, from the demonic that cannot, because the demon works through producing the illusion of the miraculous or supernatural. Signifying this, although lacking a reductive experimental approach, Paré is intent on accumulating data regarding the causes of monsters, which are discussed through analogies, and which are systematically grouped in the book's chapter structure according to similar categories of origin. For example, a section on how an excess quantity of "seed" leads to superfluous limbs, is followed by chapters on multiple births, and then hermaphrodites; these androgynes lead him naturally into a discussion of transsexual women who "degenerate" into men. Modern scientific writing is selective rather than encyclopaedic, marking truth through the paradigmatic experiment whose law extends to all other relevant cases through abstract formulae. In contrast, as I have argued above, the Renaissance model of scientific practice saw organisation and compilation as essential to validating data; editorial and authorial decisions thereby play a key role in the construction of truth. Thus, whilst *On Monsters* is part of a popular period literature on the grotesque and is not a taxonomy or scientific thesis in the modern sense of the term, it should certainly be read as contextually a serious development of knowledge rather than as a prurient picture book, which it may appear to be today.

In the opening to the book, Paré lists thirteen specific reasons for deformity (there are, he says, other human reasons but none that are "sufficient or probable" to detail). The first is the glory of God. For example, the blind man in Saint John was born unsighted not because of the sin of his father or mother, but "in order that the works of God might be magnified in him" (4). Monsters and marvellous creatures also result from the wrath of God, "who permits fathers and mothers to produce such abominations from the disorder that they make in copulation," such as having sexual intercourse during menstruation, which is forbidden in Leviticus 16 (5). But in spite of these two examples, the cause of direct divine intervention is relatively absent, because of the theological dilemma produced by positing any direct relationship between the aetiology of a monster and the beneficent mind of God who might potentially have created them. Although Paré polemicises at length against counterfeit beggars feigning disease to obtain alms, he tends not to moralise monstrosity against Biblical conventions in an allegorical way. For the bulk of the book

nature is treated as autonomous and acting according to limited physical laws, rather than metaphysical or moral causality.

Although these laws are technically incorrect by the standards of modern biology, it is not hard to enter the imagination of the Renaissance scholar. In a different cultural system, it is easy to imagine some of the reasons as being wholly plausible. This is because unlike cybernetics or genetics which segregates information from the body in which it is embedded, aetiology here is singularly material: blows to the womb, hereditary illness, rotten or mixed or too great a quantity of seed, the narrowness of the womb, the irregular posture of the mother all can cause monstrous offspring. These fall under the Aristotelian category of *accidens*, circumstantial qualities distinguishing one individual of a species from another, and as George Hoffmann observes, this category of mutilations is the most empirically founded one, since *accidens* relies on contingent cause and effect.[24] Certainly, there is a beguiling clarity to these causes, which perceive a logical and single sequence of development. For example, according to Hippocrates, from a surplus of seed multiple births or superfluous organs or limbs will result; likewise it can lead to a fusing of organs or conjoined twins, because nature wants to create two children but the seed is constrained and must be "coagulated into a globe" from which conjoined twins will be born. If the *menses* contain the material form of the adult body, it is intuitive that too much "seed" should result in a multiplication or compression of the being, or that more seed produces twins. It is equally rational that a mother with an odd posture will develop offspring who are physically twisted, since as Hippocrates observes "it is inevitable that a body which moves in a small place should become mutilated and lacking" (42). Cause and effect are clearly linked in this aetiology, which is more instinctively appealing than the genetic view that a code of four chemical components (denine, thymine, cytosine and guanine) might contain the instructions for building the full complexity of organic life; that a minute error in this code at the moment of fertilisation might cause mutation; or that through genetic engineering a cow might be fused with a human embryo to concoct a chimera.

Evolutionary genetics is a cybernetic episteme, in that it assumes nature is essentially abstract information, with the physical organism simply the instantiation of that genetic code upon an organic substrate. Within such a framework, beings are defined not according to any innate, stable reality in their own space and time, but with reference to the

[24] George Hoffmann, "Monsters and Modal Logic Among French Naturalists of the Renaissance," *South Central Review* 10.2 (1993): 36-8.

type of information they encapsulate.[25] From the genetic point of view, the fly and the human are both bags of data whose sole purpose is faithfully (or selfishly, in Dawkins' argument) to reproduce that data, most of which is common across both species. Nevertheless, from the perspective of the human, the fly continues to seem a very different creature. The reduction of all life to a kind of writing of four chemical components seems radical when compared to the Renaissance alternative in which there is one hierarchical order of development in a multifarious natural world, an order which places humans at the top, and accounts for deviations from the norm as being due to observable, material causes.

Looking at the images in Paré's compendium, our instinctive reactions to the Siamese twin in the sixteenth century and a current genetic hybrid are probably quite similar, both invoking what Mary Midgeley terms the "yuk factor," the instinctive reaction to anything that deviates from the norms.[26] Responses to genetic science are often impulsive rather than logical. For example, whilst battery egg farming is commonplace in spite of the suffering it causes for the birds, the public react with greater revulsion to the possibility that a chicken might be genetically modified to become a non-sentient egg laying machine which cannot feel pain.[27] According to the first standard, the animal has few ontological rights. Contradictorily, in the second example it is felt that the biological hybrid should not be developed precisely because it violates the animal's ancestral integrity. The battery chicken is viewed simultaneously as a genetic-chemical machine for laying eggs and as an animal with an integrity that must be preserved. The science and ambitions of contemporary geneticists are in tension with our cultural, "folk" understandings of nature, since whilst the former aim to remove subjectivity from the relationship between science and the world by describing organisms such as man in quantitative terms, our ethical appreciations of the world rely on a fluid, qualitative notion of the subject. For economic purposes we may reduce the quality of selfhood and autonomy in one situation (as in the economical battery hen), whilst emphasising sentience in another (as in our preference for the chicken over the egg-machine). The linear aetiology of monstrosity in Paré's case allows for a more philosophically consistent framing of the

[25] Typifying this alternative perspective in which the agent is perceived from the point of view of its genetic information is Richard Dawkins, *The Selfish Gene*, 30th Anniversary Edition (1976; Oxford: Oxford UP, 2003).

[26] Mary Midgley, *The Myths We Live By* (London: Routledge, 2004) 105.

[27] Bernice Bovenkerk, "Brave New Birds: The Use of 'Animal Integrity' in Animal Ethics," *The Hastings Center Report* 32.1 (2002): 16-22.

folksy, "yuk" response. Because development takes place along a single plane of being and causality, rather than the double level of genetic information and embodied organism, in his exploration, Paré is comparably sensitive to anything in violation of natural integrity, regardless of its particular visual form: strange births caused by developmental forces, gall stones or objects passed through the gut, and devious fake beggars all have the same ontological status. In the ordering, tone, and moral lessons drawn from his examples, Paré seems to find all deviations equally remarkable, whilst the uncanny thing about genetic science is that something which seems so different to the human – even "monstrous," like a headless chicken – might actually have much in common with us.

Although Renaissance theories of generation are very different to genetics in technical terms, that the Renaissance theory seems more commonsensical even if materialistically wrong implies that the attempt of the "Cyborg Manifesto" to bridge the gap between cybernetic nature and embodied culture is one to which we might be anyway attuned, since we tend to construct myths of origin, form and development that fit with our common social knowledge. Though the manifesto for posthumanism seems to break new ground, potentially a manifesto may denote less the revelation of something that must be achieved at all costs, and more the explication of the conditions for change that were to some degree always there, if suppressed or buried, and which must now be revealed and energised. Haraway's cyborg is an attempt to reconcile the *knowledge* that we are purely informational entities, as considered by cybernetics or postmodern theory, with the tacit *feeling* that humans are ontologically different, privileged above the rest of nature. She attempts to produce a model for knowledge that is situated within the domain of human experience, rather than abstracted as purely rational discourse, which is traditionally the preserve of the (allegedly) objective male. The cyborg frames abstract science within embodied culture (through technologies like the ultrasound scan), revealing that although something like birth is the outcome of a genetic process, it is still perceived from the egocentric position of an embodied observer, a "modest witness." Paré's work implies that we were in a sense always ready to become such cyborgs, to frame nature within a system of cause and effect which allows man to be located in a privileged position above the fly or the chicken. The difference is that, in the Renaissance, the view that humans are dominant over nature is underpinned by a scientific causality in which organisms exist on one ontological plane under God, with man made in His image. By contrast, in the era of genetics the fact that information is separated from its embodiment in a specific organism makes it difficult to reconcile genetics with our tacit assumption that nature is hierarchical, descending from human to the chicken, but not moving horizontally so that the genetic

information of the human is almost identical to that of the different-looking animal. Haraway observes that in the late twentieth century we are "all chimeras, theorised and fabricated hybrids of machine and organism."[28] Though this particular cybernetic incarnation may be novel, it appeals to our tendency to construct myths of origin that view nature being underpinned by a single creative rule, or chain of being which places man in hierarchy over it. The cyborg allows us to mark our difference through evidencing our technological prowess, whilst at the same time the fact that we can merge successfully with technology acknowledges that humans, like computers or other organisms, are essentially cybernetic systems, children of information.

Renaissance theories of generation are beguilingly simple, as an individual simply expands, balloon like, from its germ to adult form. The act of writing about monsters implies that the categories of normal and monstrous are stable, since anything excluded from the book on monsters is by implication not monstrous. This illustrates the general epistemological problem of the encyclopaedic and inductive approach, as opposed to the empiricist's deductive model in which the same rule can apply to superficially different phenomena. For example, even though coral might appear more like a rock than an animal, because it includes genetic matter it must be categorised as an organic rather than a geological specimen. The definition of monstrous and normal is less absolute, being defined by the ideological preconceptions of an individual author. The nature (or un-nature) of the monster provides at times a moral capital, as in the couple's offspring which shows God's wrath at their ignorance of the rules of menstruation in Leviticus 16. This asserts the existence of an original divine message to man which can be read in certain aspects of the world. On the other hand, this concocts a theological problem: if the laws of nature established by God allow for accidental developments through hazard or irregularity, this implies he is not all-good; but if God does not cause such developments, this implies he is not almighty. As Paré observes in a condemnation of astrological explanations for aberration, "God is not tied nor subject to following the order He has established in Nature, nor in the movement of stars and planets" (73). To resolve this contradiction, the emphasis is placed not on God who causes monsters by warping development at a microscopic stage of the process like a tinkering *Dei absoluta* – though He could potentially do this – but an embodied and autonomous secular culture: the way in which a woman sits during pregnancy, for example. Thus implying that we might all

---

[28] Haraway, "Cyborg Manifesto" 150.

generate the abnormal from the germ of the divine original seed, all nature potentially becomes monstrous under the manipulations of a human culture out of the logocentric control of God. On the other hand, such a view would undercut the status of the theologically educated and politically elite person able to witness the natural monsters with objective dispassion, not fearing their occult possibility. If secular culture can produce monsters, what standpoint from within culture allows analysis of their allegorical import with reference both to morality and the mind of God? For Paré, establishing such an epistemological ground depends on controlling and self-reflexively marking a distinction between reality and simulation, a boundary line patrolled by the demon.

The Real and the Virtual: Monstrous Allegory and the Demonic Imagination

George Hoffmann argues that although it is tempting to see Paré's descriptive observations and apparent concern with causality as predicting scientific method, he draws substantially on the Aristotelian distinction between the "necessary" and the "possible."[29] Because neither Aristotle nor Paré break these down further into a distinction between things actually observed (and therefore more likely to be possible) and speculative things (such as angels), Paré affords the same ontological status to all the subjects he covers. Arguing that "Nature has disported herself in order to cause the grandeur of her works to be admired," Paré saw God's power as limitless (*potentia Dei absoluta*) rather than limited by any covenant with man (*potentia Dei ordinata*). Hoffmann concludes by observing that Paré might be best understood in terms of literary magic realism, which entertains the sense of the world as that which might be possible, without regard to causal logic or to the actuality of the present.[30]

However, Hoffmann is wrong to situate Paré in this sort of context, as an imaginative fiction maker who sees the world as the virtual and fecund manifestation of a *potentia Dei absoluta*. In specific cases, such as the belief that there are seven individual cells within a woman's womb which accounts for multiple births, Paré's personal observations and his judicious evaluation of other authorities do cause him to reject certain second hand accounts as impossible (25-26). I have already stressed that the collection of monsters normalised them within nature, regardless of the veracity of any explanation as to their origin in the first place. Even so, almost every case in *On Monsters* is given a date

[29] G. Hoffmann 38.

[30] G. Hoffmann 43.

and place of observation, tying the examples to realistic historical contexts and implying that Paré is concerned to do more than simply compile in order to subordinate the monstrous. Rather, his ambition is to rationalise. In this vein, although nature's variety might be analogous to the creative mind of almighty God, Paré makes a critical distinction between nature and the supernatural, a discrimination which does anticipate that more formally inscribed by the modern "two cultures" separation of the real (those things legitimately studied by science) from the virtual (magical realism, for example, studied by the humanities). Such a discrimination is made by drawing a boundary line between the monster which can be accounted for by material causes, and the monster which appears to have been produced by the supernatural work of the devil. Paré brings the demon down to earth, showing how demonically produced monsters are actually illusions without physical cause. This division anticipates Enlightenment science's break from Aristotle's modal logic, and its epistemic distinction between science and culture.

In *The City of God*, Augustine argues that demons are "entirely alien to justice, swollen with pride, livid with envy, and subtle in deceit." In his scheme, demons do vividly exist as fallen angels and, like monsters which highlight the beauty of creation, they provide a revealing counterpoint to the good. However, unlike monsters which, to the sceptic, are passive evidence of God's true nature, demons may actively "hold sway, as over prisoners and subjects, over the many men who are clearly unworthy to share in the true religion."[31] Whilst monsters embody unintended but revealing errors, demons result from the deliberate evil of the devil. Importantly, however, demons only *appear* to operate in a supernatural or miraculous way; in actuality, their effects are produced by the same sorts of natural causes that account for monstrosity. This is precisely Peter I's argument exemplified above. We also recollect that Paré's very first distinction in *On Monsters and Marvels* was between monsters that appear outside nature and marvels that are against nature, a distinction necessary to tie nature to the moral God who created it, and to anticipate the reader of the book as a civilised sceptic rather than uneducated supernaturalist. Later in the book, Paré maintains this Augustinian tradition:

> Now these Demons can in many manners and fashions deceive our
> earthbound heaviness, by reason of the subtlety of their essence and malice
> of their will; for they obscure the eyes of men with thick clouds that
> scramble our minds giddily and deceive us by satanic impostures,

[31] Augustine 344-45.

> corrupting our imaginations through their buffooneries and impieties. They
> are "doctors" in falsehoods – the roots of malice – and in all wickedness to
> seduce and deceive us and prevaricators of truth; and, to say it in a word,
> they have incomparable skill in deceits, for they transmute themselves in a
> thousand ways and heap on the bodies of living persons a thousand strange
> things. (91)

Paré acknowledges the existence of demons and their ability to possess the body and to cause it to expel bizarre objects: knives, hair balls, nails, through the stomach, mouth or womb. However, unlike the other monstrous aetiologies in which cause and effect result from the individual's development, here the spectacular effects produced by demons originate not in causes in the body of the possessed, but in the minds of the observers; they are "done through the cleverness of the devil, who deceived the persons present through their faculty of sight" (96). The implication here is that the refined reader of Paré's work will be able to remain sceptical, refusing to entertain the virtual and strange reality concocted by the devil.

This is entirely consistent with theological convention which considered the devil to be a virtual creature, in that he both possessed the same powers as God whilst being simultaneously his antithesis. As Stuart Clark explains, the devil's power was rationalised as being the ability to create deception with the permission of God, and to simulate divine powers without possessing their intrinsic basis and hence ambiguously sharing their morality.[32] Thus, for example, possessions were construed as being permitted by God so that the priest performing a successful exorcism might edify and assert the value of his role in exposing evil's limitations.[33] In effect, the devil is self-deceiving in believing himself to have true power when he possesses only its illusion. So God could in principle control demons, since they are but simulations of the power of evil without true base. Further, marvels work through the mind's eye of the superstitious witness, and do not therefore

---

[32] Stuart Clark, "Demons, Natural Magic, and the Virtually Real: Visual Paradox in Early Modern Europe," *Paracelsian Moments: Science, Medicine and Astrology in Early Modern Europe*, ed. Gerhild Scholz Williams and Charles D. Gunnoe Jr. (Kirksville, MO: Truman State UP, 2002) 223-45. In his vast survey *Thinking with Demons*, Clark expands on the value of the devil (and witchcraft and demonology) in a theological and cultural logic which was structured according to opposites and inversions.

[33] As D.P. Walker has shown, the ritual of demonic exorcism consequently became a proving ground for Catholicism to combat Protestantism in sixteenth-century France, whilst in England the rationalisation of demons asserted Protestantism against Popish performativity. Walker 6.

have any real physical cause and hence claim to allegorically reveal a moral or theological truth in the way monsters might (as in Paré's reading of Leviticus 16). Because demons can only do what God permits, they are preternatural rather than supernatural (beyond nature's laws), and it is only a human fallacy that perceives their actions to be miraculous (that is, outside of nature) rather than, like monsters, against the usual norms of nature but not wholly unnatural.[34] The aim of natural physics at this time "was to demystify demonic pretensions by subjecting them to careful and essentially negative scrutiny – sorting out, in particular, just where the limits of demonic efficacy via second causes were reached and the realm of fantasy and delusion began."[35] In the history of science, Paré's scepticism regarding demons, even whilst he talks at length about monsters, anticipates the later attempts in the seventeenth century to understand the demonic in rational terms as the result of physical illness and personal fabrication – particularly, as Steven Connor has shown, through exposing the dramatic staging of demonic voices.[36] The fact that, for all his continual curiosity about strange nature, Paré demotes demons to producing the appearance of monstrous entities, rather than actually causing supernatural events to occur, suggests that he is entering the "secular attitude" which is fundamental to modern science.[37]

However, though tempting, it would be wrong to present Paré's work as being derived from any sort of rationally Cartesian logic. Rather, his emphasis on the veracity of monsters and his downplaying of deceiving demons stems from theological conventions. In this sense, Paré's demons remain hybrid actants, in that they produce a quasi-scientific epistemology – in the passage quoted above Paré privileges truth above deception – but they do so from a cultural standpoint (religious belief), rather than from the absolute facts encoded in the world which objective science can discern without bias or relativism. Paré attempts to categorise according to formal principles of monstrosity. Thus incorporating the demon within the same text might destroy any broader moral meaning that emerges both from its celebratory display of God's creativity, and from its containment of monsters by decontextualising them on the page as objects of spectacle. If demons, too, are laid out

[34] For a detailed analysis of the way in which demons worked to perform marvellous feats that nevertheless were not miraculous, or supernatural, see S. Clark, *Thinking with Demons*, especially 161-78.

[35] S. Clark, *Thinking with Demons* 153.

[36] Connor, *Dumbstruck* 153-225.

[37] Thomas Goldstein, *Dawn of Modern Science* (Boston: Houghton, 1980) 43.

as further evidence of God's imaginative abundance or *potentia Dei absoluta*, the implicit possibility is that it is demons who have hoaxed the witnesses of monsters whom Paré collates into believing they were seeing monstrous things. Thus monsters would not actually exist in their etymological sense of showing or telling, for the world would be populated only by demonic illusion-mongers. On the other hand, to keep demons outside the category of the natural is paradoxically to give them a freer licence. As Steven Connor notes, demons are:

> accursed because of the power they have of implanting vain and false ideas. But what if one of these ideas was the idea that there are demons, and that they have power over us?...If the very idea of the counterfeiting demonic were itself a counterfeit, then where would be its origin, and who could be held responsible for the counterfeit?[38]

If the demon as a potentially counterfeit counterfeiter is not bound to the rule of God, then it exists in an ontological limbo. The best compromise Paré can make is to keep the demonic as a virtual force, capable of producing cognitive or visual illusions but not physical actions. Therefore, nature alone remains the sole code to the divine mind, with its reliable and linear contingencies in which monsters result from unusual sitting positions, for example.

Indeed, the fact that Paré is not objective in the modern sense is clear by his attitude towards women. Further segregating nature and the devil, Paré (again following Classical tradition) rejects the idea that human monsters can be produced as the offspring of demons and woman (49). Oddly, though, Paré also identifies the maternal imagination as one factor in monstrosity. Like many authorities of his time, he draws on Hippocrates' argument that lengthy contemplation of an object can cause birthmarks with a similar imprint, a belief given added credence by the recent reintroduction of Hermetic ideas into Western thought. For example, Paré cites Damascene's attestation of a mother who, having looked too intently at the image of St. John dressed in skins, conceived a girl "furry as a bear...deformed and hideous" (38). But if the demonic deceptions of the imagination are given only a brief space in *On Monsters*, how then to account for these special cases of mind affecting matter? The Aristotelian tradition treats the mother as simply the passive receptacle for the male seed which contains the form of the child, so there is a contradiction in a work which otherwise inherits Aristotle, whilst also claiming that

---

[38] Connor, *Dumbstruck* 96.

because the mother is responsible for the continuation of the species she is also accountable for monstrous accidents.[39] Paré is unselfconsciously misogynistic on this point, noting:

> how dangerous it is to disturb a pregnant woman, to show her or remind her of some food which she cannot enjoy immediately, and indeed to show them animals, or even pictures of them, when they are deformed and monstrous. For which I'm expecting someone to object to me that I therefore shouldn't have inserted anything like this into my book on reproduction. But I will answer him in a word, that I do not write for women at all. Let us return to our subject. (55)

Even though monsters cannot be produced through demon-human copulation, the maternal imagination is subject to the imprinting of simulacra on the medium of their bodies and those embryos they carry. The monster stands as a moral rebuke to the excessive psychology of the woman (just as late nineteenth and early twentieth-century feminists were medicalised as being "hysterical"). Naturally, not being subject to pregnancy, the male viewer remains protected from such cause and effect. In a self-fulfilling circle, this spectacular picture book of monsters both authenticates the superiority of the male witness, and the scientific credentials of the writer who presents it for him without himself becoming deformed through baroque excess. As Hanafi writes, glossing the very similar views of the early seventeenth-century scholar, Giambattista della Porta, "the woman literally makes of her body a laboratory, turns her procreative capacities into a scientific endeavour whose results make up the data," in this case confirming the Hermetic hypothesis that intangible sight of an object can cause it to be produced in actuality.[40] As with the *Kunstkammer*, though, the control contrived through perspective leaves the problem of generation in its wake. Paré's elision of the precise mechanism that allows the mother's imagination to affect the physical embryo in accordance with Aristotle suggests that as theology was illogical, it was forced to protect itself by sexual prejudice, excluding one community of readers who might potentially complain about this causal contradiction.

---

[39] Marie-Helene Huet, "Living Images: Monstrosity and Representation," *Representations* 4 (1983): 73, *JSTOR*, 16 Sept. 2008 <http://www.jstor.org/stable/2928548>.

[40] Hanafi 58.

The Demon and the Cyborg

Paré's final point in the quotation above, where he does not write for women, returns me to where I began this analysis of Paré's monster compendium, by emphasising the way in which writing as an act cuts the author off from the demonic delusion and monstrous superstition. To write for women is to write for those in league with the devil, as indicated by the unrestrained power of the female imagination. In a similar vein, the medieval naturalist is cautious that one factor in monstrosity may be a surplus of the imagination, or the devils and demons who corrupt it. Paré is aware of the power of writing to contribute towards this extension of the mind beyond the confines of language, to mimic devilish deception through the human creative imagination. Thus he restrains from telling of the "great loathsomeness" of additional monsters of any particular type (73). The fact that monsters cannot be themselves generated by demons, and that the text on monsters must not be allowed to become a diabolic surplus of the imagination, indicates that truth inheres in monsters in their own right, with God's moral purpose potentially legible in their nature. In a modern age of epistemological and disciplinary boundaries, one is either wholly a materialist or wholly a mystic (or else one holds some position poised uneasily between the two, such as Intelligent Design). And if one holds a belief in God or monsters, why not also in demons? But in the Renaissance, theocentrism can be maintained simultaneously with an imaginative restraint, such that only God can be truly supernatural whilst the work of the devil is confined to the production of illusion.

How, then, does the demon/monster distinction as a way of claiming the epistemology of Renaissance science relate to the contemporary cyborg? As indicated by Paré's contradictory treatment of the female imagination, the demon is implicitly an ideological construction, whilst the monster marks the objective nature that, though diverse, is causally stable. Similarly, rather than being a definite identity constituted through specific technological models, the cyborg exists when ideology and science come into conjunction. For Haraway, some versions of the cyborg are constituted of "nothing but signals, electromagnetic waves, a section of a spectrum…Cyborgs are ether, quintessence." Like Paré's demons which assume psychological forms, these varieties of cyborg are dangerous, because they are "as hard to see politically as materially. They are about consciousness – or its simulation."[41] As in her critique of "second birthing," Haraway has

---

[41] Haraway, "Cyborg Manifesto" 153.

in mind the cybernetic transcendentalists here, who deny the value of embodiment entirely.[42]

At issue in the Renaissance monster is ontology, with the book designed to collect and define those characteristics that mark in them a deviation from the norm. At issue in the demon or cyborg, however, is not so much what it is innately, as to how it situates the observer in a way that makes it difficult to understand that world order according to determinable logics – theology in Paré's case, social politics in Haraway's. Of particular concern to Haraway is the way in which modern science might refuse to succumb to sociological analysis or critique because it deals with pure information, transmitted instantaneously in a multimedia world. Understanding the way in which cybernetic technologies still embody ideologies of sexual difference, rather than transcending embodiment as cyberneticists like Minsky claim, may be difficult in this context. Both demons and cyborgs are epistemological creatures which foreground the ways in which science establishes its bases of knowledge and frames of reference. Just as the exclusion of the demon validates the importance of the science of the monstrous, and the authority of the writer who seeks to rationalise the latter, so too the existence of cyborgs presents us with monstrous images (chicken-factories) that provoke us to consider the social stakes of scientific and technological advances, even if those sciences are ostensibly studying the immaterial world of information (such as genes). Haraway calls for the cyborg to be seen as an epistemological entity, framing our scientific knowledge in relation to a world of embodied human individuals.[43] Epistemological framing of things that might or might not be considered monstrous is the work performed by Paré's demon or by Haraway's cyborg which reside on the boundaries of knowledge, between nature and culture. As I suggested in chapter 1, when thinking about posthumanism we ought not just to consider its constructions of otherness but the way those entities like monstrous robots and cybernetic hybrids inform us about our own modes of constructing knowledge; hence why the demon, as an inherently epistemological agent, provides a useful trope.

With this connection, though, we must acknowledge the telling difference between the sixteenth and twentieth centuries. Whereas Haraway sees the cyborg as capable of raising consciousness of science in society precisely because it seems so provocative, even science fictional, Paré sees the demon as a concept that must be proscribed, even if it

[42] See chapter 1, note 45.

[43] Haraway, "Cyborg Manifesto" 160-61.

cannot be erased. Paré's demons are designed to separate the real from the virtual, to deny the role of the devil in creating monsters by keeping his work as a virtual one. Second-wave cybernetics, however, argues that in a sense the world is only ever knowable from a subjective, virtualised standpoint. Humberto Maturana's seminal paper "What the Frog's Eye Tells the Frog's Brain" shows that a frog is barely capable of perceiving large, static objects, but well-adapted to perceiving small, fast objects (such as flies).[44] As N. Katherine Hayles glosses the experiment, "The results implied that the frog's perceptual system does not so much register reality as *construct* it."[45] Just as Heisenberg's uncertainty principle in physics shows that the act of measurement changes the behaviour of the particle under observation, so the biology of a cybernetic agent – whether frog or human – potentially determines perception before the world is raised to consciousness, such that the world we know is not necessarily the world as it physically is. Poised between the physical and the metaphysical, the deceptive demon intervenes *between* reality and the mind. Cybernetics, however, incorporates deception *within* the information receptor that is the brain, so that objective facts about the state of a system are actively determined rather than passively witnessed by the observer within that same hermeneutic. This demonic possibility appears to be what Haraway is working towards in explaining that the danger of the posthuman condition is that, conceptualised as a being of information rather than tangible matter, it resists understanding and political deconstruction. And such perceptual instability is certainly true of Freud's interpretation of demons as projections of the self-deceiving ego, which I examine in the next chapter. But, as I shall go on to argue here, cybernetics also provides a technological confirmation of the suspicion that sneaks around Paré's work: the possibility that the mind is not able to perceive the world in an undistorted way, so that nature is actively created rather than mimetically reflected by language.

For Haraway, the underlying boundary breakdowns that make the cyborg a possibility both as analogy and real identity are threefold: the breach in distinctions between human and animal; the leaky distinction between animal and machine; and the imprecise boundary between physical and non-physical. Similarly leaky distinctions seem to pervade the science of Ambroise Paré. However, in the Renaissance the distinction

---

[44] J.Y. Lettvin, H.R. Maturana, W.S. McCulloch, and W.H. Pitts, "What the Frog's Eye Tells the Frog's Brain," *Proceedings of the Institute for Radio Engineers* 47.11 (1959): 1940-59. Humberto R. Maturana expanded these ideas about the reliability of observation in "Reality: The Search for Objectivity or the Quest for a Compelling Argument," *The Irish Journal of Psychology* 9.1 (1988): 25-82.

[45] Hayles, *Posthuman* 135.

between a virtuality perpetuated by demons and the limited aetiology of the physical monster produced in accordance with *potentia Dei ordinata*, can only be preserved through the regulatory authority of the Church, which makes the causal logic contradictory with regards to women. It is the absence of such an external authorising agency, rather than the epistemological conditions it controls, which most sharply distinguishes Renaissance science from the postmodern trope of the cyborg. Though it is dictated by theological rather than sexual concerns, Paré is something like Haraway's idealised modest witness, taking responsibility for his own observations, recognising that the process of science is not just a one-way move of reading the world from a position outside culture, a culture of no culture. Rather, the naturalist knows he exists within a culture of authority and theology that can reciprocally affect that world through the practice of his prototypical science. At the same time, and unlike Haraway's feminist cyborg, Paré does not reflect on the contradiction that results: that he both denies the possibility of demonic deception whilst blaming the maternal imagination for causing stigmatic birthmarks. This limits the possibility of treating Renaissance monsters and the posthuman cyborg as uncanny dyads, even if the paradoxes in the former reflect an epistemological uncertainty that does resonate powerfully with Haraway's motivation for creating the latter.

Simultaneously a prurient spectacular of the grotesque, and an attempt at serious natural science, Paré's work is chimerical, poised between the rationalisation of the monster and the imaginative fiction of the demon, striving to maintain a materialistic account of development without underplaying the role of an omnipotent but benign deity. Faced with contradiction, Paré moves to categorise and cover the range of the monstrous in his theatre of nature. Haraway's postmodernism also manifests itself in generic hybridity. Her "Cyborg Manifesto" appeals to a science fictional model in order to posit a new epistemology for science in the (post)human world. Stylistically, both Paré and Haraway reflect a time when religious and secular epistemologies undergo flux, and when disciplinary boundaries are threatened by shifts in the entities with which they should conventionally concern themselves: physical monsters or metaphysical demons, the body or the mind, organic beings or informational cyborgs. In a contemporary age when rational models of the self (as in cybernetics) butt against models of the self as a linguistic construct (as in postmodernism) the resulting turbulence provides the conditions in which demons can breed in fiction.

Chapter 3. Monsters from the Id: Psychoanalytic Demons and Uncanny Technology in

Forbidden Planet

Secular Demons

We saw in the previous chapter that Paré epistemologically grounded his science by distinguishing between monsters as the products of nature and demons as supernatural beings capable of deceiving a non-sceptical or overly imaginative observer. Paré refutes the devilish supposition that "reality" is a total demonic delusion in which distinctions between the true and the false, and hence good and evil, break down. For Paré, the distinction between monsters and demons marks the difference between the fecund natural world that can nevertheless be reliably read because of *potentia Dei ordinata*, and the infinitely deceptive work of the devil. But whereas it might seem a remarkably modern move to treat demons as speculative or fantastic entities outside the domain of science, why is it, then, that demons continue to be discussed even in modern, secularised society and culture?

One clue can be found in an 1897 letter from Sigmund Freud to Wilhelm Fleiss, in which Freud noted that his "brand-new theory of the primary origins of hysteria" was already familiar from numerous accounts of demonic possession over the centuries. The medieval theory, Freud claims, "was identical with our theory of a foreign body and the splitting of consciousness."[1] Though demonic possession supports Freud's theoretical interpretation of the unconscious, the implication is that since the unconscious is part of our inner nature, the demon is a trope or metaphor which appeals so powerfully to our experience of being in the world that it could never really be eradicated unless our psychology was to undergo a total revolution. Although the interpretation of demonic symptoms, such as speaking in voices, had begun to expose them as deliberately contrived or accountable through physiological causes by the late seventeenth century,[2] it was psychology which finally closed the case against supernatural or metaphysical causality. *The New Catholic Encyclopedia* concedes that the orthodox position on "possession" is that:

---

[1] Sigmund Freud, "To Wilhelm Fleiss," 17 Jan. 1897, *The Complete Letters of Sigmund Freud to Wilhelm Fleiss 1887-1904*, trans. and ed. Jeffrey Moussaieff Masson (Cambridge, MA: Belknap P, 1985) 224-25.

[2] Connor, *Dumbstruck* 209.

The present, more detailed knowledge of the universe has forever destroyed the crude concept of a three-storied world in which angels and demons materialize with ingenuous frequency. Psychiatry, moreover, has shown that the workings of the subconscious explain many, if not most, of the abnormal conditions that earlier generations had attributed to diabolical activity. For these reasons and because the need to reorient theology along more positive lines has been recognized, demonology has not been the object of very much serious study in the 20th century.[3]

Through the notion of the unconscious, Freud locates demonism in a framework in which reality and the imagination blend into each other, implying that we can never be sure of the reliability of our own perspective on the world. Although even Catholicism concedes that they do not exist as actual entities mediating between the subjective self and objective reality, demons still occupy a significant position as metaphors of lurking repressions which have the potential to cause an overlap between the psychological and the physical. What we had taken to be an objective reality – a reality that seems visceral when confronted by monsters – is actually constituted subjectively, through our inner demons of protective self-delusion. Such a view appeals to a tacit fear – one present in Paré's time but significantly exacerbated by the existence of cybernetic virtual worlds and artificial doubles – that the world we perceive is not really the world as it is. Thus demons trope and describe this experience even though they may not exist in actuality, whilst psychoanalysis, though an imperfect science, resonates with our experience of a world which seems never truly apprehensible. In this chapter, I explore this by referring firstly to Freud's aetiological study, *A Seventeenth-Century Demonological Neurosis* (1923), and then to his classic essay on "The 'Uncanny'" (1919), which is concerned with the strange that is nevertheless somehow familiar.[4] And what could be more uncanny than an android or artificial intelligence that resembles ourselves whilst being internally mechanical?

[3] L.J. Elmer, "Demon (Theology of)," *The New Catholic Encyclopedia*, ed. Berard L. Marthaler, 2nd ed. vol. 4 (Detroit: Gale, 2003) 649. This appears to be a reprint of the original article from the first edition of the *Encyclopedia*, published in 1967.

[4] Sigmund Freud, "The 'Uncanny'," *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, trans. James Strachey, vol. 17 (London: Hogarth P, 1955) 219-52; Sigmund Freud, *A Seventeenth-Century Demonological Neurosis, The Standard Edition of the Complete Psychological Works of Sigmund Freud*, trans. James Strachey, vol. 19 (London: Hogarth P, 1961) 72-105.

Internalising the demon as a psychological rather than metaphysical phenomenon allows the demonic to be redeployed within cybernetic fictions whose central concern is the relationship between the human and the artificial model or simulation in an otherwise rational, scientific and technological environment. If Athanasius Kircher had once projected demons to illustrate his mastery of illusion, this new interpretation of demons brings them truly under the power of technology, products of a scientific rather than devilish spirit. As Adam Philips claims, through reading the correspondence between Freud and Sandor Ferenczi (Freud's psychoanalytic collaborator), "it is clear that sexuality and the unconscious were the new, scientifically prestigious words for the occult, of that which is beyond our capacity for knowledge, for the weird, unaccountable effects people have on each other. In psychoanalysis the supernatural returns as the erotic."[5] Freud's reading of demons allows us to understand why the concept resonates in the cybernetic era, although as my interpretation of the 1950s science fiction *Forbidden Planet* shows in the second half of this chapter, cybernetic technology also highlight the deficits in Freud's account of the unconscious, particularly in relation to the uncanny. By reworking Shakespeare's *The Tempest* in the context of space, this landmark film appears to be an allegory of the unnatural power of science which destroys an otherwise Edenic pastoral. However, the film subverts its own conventions, attesting to the validity of evolutionary and psychoanalytic sciences. In particular, it is explained that the "Caliban" monster arises not from some metaphysical, moral or supernatural force, but emanates plausibly from the mind of the scientist who is linked to a cybernetic technology. Rather than the monstrous, Caliban figures the human and his scientific mind. To be human is to be determined by psychological, biological and linguistic codes and, indeed, to perceive otherwise is to be subject to the most subversive demon of all: self-delusion.

Psychoanalysis and the Demonic Projection of the Unconscious

Cybernetic science itself rejected psychoanalytic ideas, as they were represented at the Macy Conferences by the psychiatrist Lawrence Kubie. Freudian notions of the libido or unconscious were essentially unquantifiable whereas, by taking information as the fundamental unit of the discipline, cybernetics could treat mind in terms of mathematics. Drawing on Shannon's work in codifying the amount of information contained in any message regardless of its meaning or medium, cybernetics was able to treat communication

[5] Adam Phillips, *Terrors and Experts* (London: Faber, 1995) 19.

in an objective way, whereas the risk in psychoanalytic communication was always that the analyst was projecting his own desires onto the analysand.[6] On the other hand, the psychoanalytic belief that fictions appeal to the unconscious as well as to the conscious through their generation of latent as well as explicit meanings informs narratives like *2001*, with its dream-like sequences; further, writers like Philip K. Dick, Umberto Eco and A.S. Byatt display their awareness of psychoanalysis in their representation of paranoiac or schizophrenic characters. In relating cybernetic science to the demonic or the monstrous in culture, psychoanalysis remains a crucial framework.[7] Freud encourages us to recognise the internal fractures of psychological being in ways attuned to discourses of the demonic. Though Paré also argues that deception occurs in the mind of the observer rather than through a supernatural force in the world, in a shift from Paré's writing deception now has its source not in the single Devil but in a universal, neurotic demon within. As such, distinctions between self and other, nature and culture, or imagination and reality are recapitulated as the conscious and the unconscious. This dualism marks a gap between the superficially human and the self actually informed by underlying codes of which he or she is not aware, particularly biological and cognitive programmings which imply that the differences between the individual and the machine, which is also code-driven, are not clear-cut. Being confronted with the machine as mirror of the self, rather than its other, is thus an uncanny experience.

My benchmark for Freud's interpretation of demons is his 1923 case study, *A Seventeenth-Century Demonological Neurosis*. Coming as it does after Freud's last great

---

[6] For the differences and debates between the psychoanalytic school and the more materially minded cyberneticians, see Heims 115-63.

[7] Another, which I shall not examine at length, is evolution, which contends that distinctions between species (and hence between monsters and normal nature) are essentially arbitrary, the product of man's view of the world rather than any innate ontology. As Gillian Beer has noted, "In *The Origin* Darwin remarked that 'monstrosities cannot be separated by any clear line of distinction from mere variations.' A monster may be a form out of accord with the environment in which it must live. Change the environment, the monster (it seemed) may become apt, even normative." Gillian Beer, *Open Fields: Science in Cultural Encounter* (Oxford: Clarendon, 1996) 124-5. Bio-technical constructs like the cyborg rely on our knowing that man can be joined to machine, or genetically adapted to mesh with other species; and the monstrous cyborgs of fiction require us to admit that the boundary lines between our selves and the machine are not clear cut. Thus in the environment of space or the factory floor, for example, a robot may appear entirely well-adapted and functional, and it is only when it comes to Earth or escapes into society that the thing becomes monstrous. Isaac Asimov's stories (see chapter 1, note 65) stage this contextual shift on several occasions.

theoretical work, *The Ego and the Id*, this exemplifies Freud's developed form of psychoanalysis.[8] In this case, Freud analyses a painter, Christopher Haitzmann who, in 1668, signed a blood pact with the devil, becoming his son and vowing to give him his body and soul nine years later. In 1677, the terrified Haitzmann asked for the intercession of the Virgin Mary, who freed him. Haitzmann's archived diary and his paintings of the apparitions of Satan testify to the sincerity of his belief in both the devil and the miracle. Contemporary commentators – including the monks to whom Haitzmann had fled in 1677 – interpreted the case as evidencing the danger of union with the devil, but in his reading Freud strips away the allegorical import, explaining instead that:

> The states of possession correspond to our neuroses, for the explanation of which we once more have recourse to psychical powers. In our eyes, the demons are bad and reprehensible wishes, derivatives of instinctual impulses that have been repudiated and repressed. We merely eliminate the projection of these mental entities into the external world which the middle ages carried out; instead, we regard them as having arisen in the patient's internal life, where they have their abode.[9]

Rather than supernatural agents actively sent into the world in order to provide a cautionary tale about evil, demons originate within the individual and are projected onto the screen of reality (Freud's word "projection" will take on a literal significance in my discussion of the actual screens of cinematic science fiction later in this chapter). Those agents Paré places outside of nature, Freud argues *are* human nature. As a result, the psychoanalyst of demons is himself possessed because neurosis is within all – note Freud's use of the first-person plural, "our," in the quotation above – and this self-reflexive effect must make it questionable whether any consciousness can interpret another objectively. If all are subject to the impostures of demonic delusion, there is perhaps no stable ground from which authoritatively to observe the self and the world.[10] As Lawrence Kubie

---

[8] Freud confidently explains that he will not engage in a polemic justification for using psychoanalysis in relation to the archive, since its status even among "the more obtuse" of his contemporaries is increasingly assured and "no understanding of the neurotic states can be reached without the help of psycho-analysis." Freud, *Demonological Neurosis* 84.

[9] Freud, *Demonological Neurosis* 72.

[10] Indeed, any time Freud stepped into his waiting room of Berggasse 19, he would have seen his copy of Fuseli's *The Nightmare* hanging on the wall, its incubus grinning out at the psychoanalyst as if to remind him that he, like the unfortunate reclining woman, is

admitted, the analyst must try to remain "as far as possible a dummy in a store window on whom the patient drapes his fantasies...a screen on which the patient projects the shadows out of his own past."[11] Whereas Paré's multifarious book of monsters authenticates the superiority of the educated, sceptical, male reader, no such objective and idealised reader is conceived in psychoanalysis. This is why the archival study appeals to Freud, since historical studies "provide a glimpse of phylogenetic foundations...presenting them to the observer with relatively little distortion or obscurity."[12]

Here, Freud's diagnosis is that the melancholic painter, who had recently lost his father, took the devil as a father substitute. As David Thurn glosses it, "the lost object has been incorporated into the ego, dividing it, transforming the conflict between the ego and the loved person into a raging conflict within the ego itself, thereby setting the stage for a fantasy of demonic possession."[13] The period of nine years correlates with the duration of pregnancy explaining why, though ostensibly playing the part of father substitute, Haitzmann's paintings also depict the devil with monstrous breasts. This analysis of the sexual desires latent behind the imagination is typical of Freud's phallocentric bestiary of the unconscious mind. He enables a reworking of the trope of possession within the boundaries of the human self rather than turning to external metaphysical or religious narratives. The encounter with the outer demon is really an uncanny encounter with the repressions in the self. This passage, then, inverts the terms of Paré's science. In the Renaissance, the monster evidences the ability of the naturalist to interpret and incorporate nature in a way that, by systematising and categorising them, makes monsters inert in the human context whilst providing evidence of God's abundant creativity as exercised from

---

plagued by bad dreams. Frayling, *Nightmare* 11. The original painting is Henry Fuseli, *The Nightmare*, Detroit Institute of the Arts, Detroit.

[11] Lawrence S. Kubie, "Problems and Techniques of Psychoanalytic Validation and Progress," *Psychoanalysis as Science: The Hixon Lectures on the Scientific Status of Psychoanalysis*, ed. E. Pumpian-Mindlin (New York: Basic, 1952) 59.

[12] David H. Thurn, "Fideikommi Bibliothek: Freud's 'Demonological Neurosis'," *Modern Language Notes* 108 (1993): 851. The use of the archival reading to escape the controversy of relativism compares with evolution's attention to the fossil record as the most reliable indicator of species' development and relatedness.

[13] Thurn 856.

his divine sphere. The demon represents the limits of that enterprise. Since they produce their effects via the imagination, to categorise or collate demons is to legitimate and maybe propagate their deceptive capacity, thereby undermining the inductive method, the collation of other scientific authorities and stories not observed at first hand. For Freud, however, the self or ego always constructs its own world virtually or demonically, such that material monsters and demons as depicted in Haitzmann's art are merely the analogical manifestation of an unconscious motivation within.

Of course, one could argue that Freud's conceptualisation of demons is empirically suspect, and it is not surprising that contemporary psychiatry has tended to incorporate demonism as just one subset of schizophrenia and dissociation disorder.[14] On the other hand, I commented in the chapter 2 that we implicitly cling to a hierarchical ontology of monsters despite neo-Darwinism, and hence can sympathise readily with Paré's preformationism; similarly, psychoanalysis offers a way of thinking about demons which, though it may not be physiologically or neurologically accurate, still meshes with the way in which we tacitly relate to the world and the experience of living in it. Freud straddles the pre-Enlightenment paradigm in which knowledge is constituted in accordance with everyday experience – in which the foetus is the preformed version of the adult, or in which demonic deceptions do appear to happen – and a modern rationalism which tries to account for effects by material causes. Whilst incorrect levels of dopamine might be the real cause of depression or schizophrenia, it is demons which appeal as metaphors or embodiments of a condition of existence – think, for example, of Churchill's infamous "black dog" depression, or of the "demon drink."[15] Regardless of its hard truth, therefore, what matters about the psychoanalytic aetiology of demons is the fact that it makes it legitimate to talk about demons as metaphors emerging from the human consciousness or unconscious, without appearing to make any strong claim as to their metaphysical origin. In fact, in "The 'Uncanny'," an essay I will treat in more depth in a moment, Freud is perfectly willing to admit that by ascribing maladies such as epilepsy and madness to demons, the Middle Ages had their psychology almost correct.[16] Within cultural narratives,

---

[14] For a survey of these approaches, see Philip M. Coons, "The Differential Diagnosis of Possession States," *Dissociation* 6 (1993): 213-21.

[15] More recently, Andrew Solomon recorded his experiences with a demonically personified depression in *The Noonday Demon: An Anatomy of Depression* (London: Vintage, 2002).

[16] Freud, "The 'Uncanny'" 243.

psychoanalytic demons sympathetically explain the uncertain and mediated relationship between reality and the self, or between representations or simulations and the real. As Steven Connor observes, rationalism can never exorcise the demon since, as human reason is freed from the threat of the objective devil on the outside, this identifies "an even more dangerous threat to reason on the inside, in the endogenous tendency of the human mind to entangle itself in fantasy, imposture and illusion"; thus "When devils become merely figurative, figurativeness can start to seem demonic. Rather than devils being drawn into dialogue, dialogue itself starts to have a devilish aspect."[17] As this chapter and the previous one have evidenced, the allegorical impulse, the desire to perceive the objective world as reflecting some meaningful higher scheme, is colourfully illustrated in the hermeneutics of monsters and demons. When psychoanalysis assumes prominence in culture, if not in mainstream science, its contention that demons are extrapolations emanating from the human mind, rather than modes of revelation coming from a metaphysical direction, therefore becomes a signal of self-delusion as exercised through narratives. Stories of the supernatural in the world are actually reincarnations of our own fears. Thus psychoanalytic demons must inform fiction about its own limitations as an allegory or representation of the real.

If postmodernism has involved the deconstruction of a master- or meta-narrative, and the denial that language can mimetically reproduce the world in a direct correspondence between signifiers and signified, it should be clear why not only psychoanalysis (through Lacan), but also demons, have appealed to literary critics. For example, taking up Freud in his 1971 study of modern fiction, *The World and the Book*, Gabriel Josipovici argues that the medieval literary universe was built on a series of analogies: between physical and spiritual worlds, secular and sacred history, the history of the universe and of the individual, between the natural cycle and the Christian year and, underpinning all these, between God and man, who is made in God's image.[18] The modern novel, however, admirably embraces the decline of allegory. Its realistic conversational form sucks the reader into the delusion that the narrative is true to life as the reader experiences it, with the narrative being taken for reality. However, by then drawing attention to itself as a story chosen and written as one among many possible alternatives – as the strong metafictional element of contemporary narrative often does – we are forced to

---

[17] Connor, "Toward a New Demonology" 107.

[18] Gabriel Josipovici, *The World and the Book: A Study of Modern Fiction* (London: Macmillan, 1971).

acknowledge that what we had taken for reality was merely a fiction seen through spectacles of delusion. Thus in modern literature, Josipovici argues, analogy (which Connor terms more generally as figurality) has become "demonic," since analogy no longer points outwards, from art to some religious value external to it, but looks inwardly, interrogating the psychology of the reader:

> For to discover correspondence in the world around us does not lead to the
> sensation that we are inhabiting a meaningful universe; on the contrary, it
> leads to the feeling that what we had taken to be the "world" is only the
> projection of our private compulsions: analogy becomes a sign of
> dementia.[19]

Instead of the fictional analogy embodying in the text some truth that stands in reality outside of the page, attention is now paid to the extent to which language can be said to be a mimetic representation of a truth "out there," or whether the world is always already a linguistic construct, the "projection" (that word again) of a personal narrative or "private compulsions" onto an essentially blank canvas. However, even if fiction no longer provides an index to metaphysical truths, it can protect us from the misguided belief that there is shared, and hence objective, meaning in the real world. Narrative self-consciousness thereby protects us from the very real solipsism which is not to know that one is in a state of solipsism.[20]

Exemplifying just such a self-reflexive shift in the demonic is the 1956 science fiction film, *Forbidden Planet*.[21] As an adaptation of Shakespeare's *The Tempest*, this handily also corresponds to the different representations of the demonic between Paré (whose *On Monsters* precedes *The Tempest* only by around forty years) and Freud.[22] For *Forbidden Planet* embodies the monstrous as the "projection" of something demonic within the self, in accordance with Freud's demonological analysis. However, the projection here is ironically literalised, as it is produced through cybernetic technologies of

---

[19] Josipovici 299.

[20] Josipovici 307.

[21] *Forbidden Planet*, dir. Fred McLeod Wilcox, perf. Walter Pidgeon, Anne Francis, Leslie Nielsen and Warren Stevens, 1956, DVD, Warner Brothers Entertainment, 2007.

[22] Though possible, it appears unlikely that Paré's *On Monsters* provides a source for Shakespeare's representation of Caliban. Jeffrey Cahan, "Ambroise Paré's *On Monsters* as a Possible Source for Caliban," *Early Modern Literary Studies* 3.1 (1997): 4-11, 2 Jan. 2007 < http://extra.shu.ac.uk/emls/03-1/kahatemp.html>.

cinema. Therefore, rather than being passive media allegorically presenting a narrative, this cybernetic fiction forces us – through its medium as well as its message – to reflect on the mechanistic constitution of the human self.

<u>Monsters from the Id: *Forbidden Planet*, Cybernetic Media and The Demon of Analogy</u>

In Shakespeare's play, the ability to enter imaginatively into the virtual world of the drama, and to distinguish between illusion and reality is critical to its politics.[23] Caliban is the central figure here. As the offspring of a demon and a woman, precisely the sort of conjunction Paré says is impossible, he provides a focus for different possible perspectives. Is he natural or supernatural? A savage or a spirit? At times he seems human in form – Miranda notes that Ferdinand is the third man she has ever seen, implying Caliban must be the second – but Prospero claims he is "A freckled whelp, hag-born – not honoured with/A human shape" (1.2.283-4). Trinculo asks, "What have we here, a man or a fish" (2.2.25), before hiding under Caliban's gaberdine and therefore confronting bemused Stefano with a beast with four legs and two mouths. Whilst Prospero sees Caliban's outward hideousness as evidence of an inner spiritual deformity, others take a less moralistic view and see Caliban as ripe for commercial exploitation.

However, just as the diversity of the monster compendium validated the standpoint of the naturalist who produced it, the diversity exhibited by Caliban reiterates the superiority of the perspective of the sceptical viewer and, vitally, listener.[24] In Aristotle's *De Anima*, voice is the denotation of "an ensouled thing," whilst sound is the mere evidence of expiration without meaning.[25] When Caliban's principal "profit" from the language Prospero teaches him is simple utterance, knowing "how to curse" (1.2.366), he is likened to the beast rather than the human, since the ability to use language is absolute rather than being dependent on the relative position of the observer. At the opposite end of

[23] William Shakespeare, *The Tempest*, *The Oxford Shakespeare: The Complete Works*, ed. Stanley Wells, Gary Taylor, John Jowett and William Montgomery (Oxford: Clarendon, 1998) 1169-89.

[24] Alison Thorne argues that by exhibiting fantastic and diverse elements in a self-contained manner, and by the use of such staging devices as scenery the perspectival sight lines of which converged on the king watching at the centre, masque conditions the supernatural as a temporary play designed to highlight political authority. Alison Thorne, "*The Tempest* and the Art of Masque," *Vision and Rhetoric in Shakespeare: Looking Through Language* (Basingstoke: Macmillan, 2000) 198-225.

[25] Aristotle, *De Anima*, trans. and introd. Hugh Lawson-Tancred (Harmondsworth: Penguin, 1986) 178-9.

the linguistic spectrum, Prospero's loss of Milan was a consequence of his self-absorbed study of books, as a philosopher rather than a political pragmatist. Thus the initial ability to distinguish between illusion and reality, something confirmed by the curtain's drop at which the audience's "indulgence" sets Prospero "free" (Epilogue.20), validates a hierarchy of power in which Caliban is at the bottom. *The Tempest* allegorises the difference between the heterogeneous world of speculative theatre, and the assured political status of those who view the differences within it, particularly those embodied in Caliban. It is this sort of difference between the virtual and the real that is elided in *Forbidden Planet*, with technology mediating between mind and matter so as to locate the monstrous not in a world that can ultimately viewed objectively but in the psychological demons within the self.

Whilst investigating the loss of contact with the colonists of the planet Altair IV, the crew of United Planets Cruiser C-57D find the world inhabited only by the philologist Dr. Morbius, his beautiful daughter, Altaira, and their obedient robot-butler, Robby. Their outpost appears to be a pastoral idyll. Altaira's friends are tame beasts and the world is cast in a golden glow. The intervention of the space ship disturbs this paradise, as the crew seek to indoctrinate Altaira into the sexual world (it is not coincidental that the successful suitor of this symbolic Eve is the Captain, named J. J. Adams). In terms of character, then, *Forbidden Planet* seems simply to map Shakespeare's characters onto those of the future. Prospero becomes Morbius, Miranda becomes Altaira, Sebastian becomes Adams, Ariel becomes Robby the Robot, Trinculo becomes the ship's comic, drunken cook, and Caliban becomes...who? The absence of a monstrous other from the moment the crew encroaches on the island is a vital sign of difference from *The Tempest*. Rather than representing Caliban as a monstrous form, the existence of the monster is created through the presence of two scientific flows – evolution and psychoanalysis – as these are mediated and transformed through a powerful cybernetic technology. Thus Caliban becomes a demon of the unconscious literalised on screen, and rather than marking the boundary between the virtual play and the political reality (as Caliban does in *The Tempest*), this monster allegorises our self-deception in relation to technology: rather than proving our ability to transcend our own nature, to become posthuman, the more advanced that technology becomes the more it evidences that humans are deterministic and materialistic entities driven by biological and cognitive programs.

Altair was once inhabited by the Krell, an alien species who were not only technologically but also ethically far advanced of humans, having become "a civilisation without instrumentalities," according to Morbius. Realising the dreams of the cybernetic

transcendentalists, the Krell developed a mind reading device or "plastic educator" that allows the measurement of the individual's underlying IQ, and the holographic representation of inner thoughts. Further, information can be planted and extracted directly from a brain interfaced with the machine; the project's climax – the "supreme discovery that would crown their civilisation" – was to allow consciousness to be given material form anywhere on the planet through harnessing the energy of a massive subterranean power plant. However, with this final experiment, the Krell were wiped out. Whilst Morbius initially explains that for the Krell the id is an "obsolete term once used to describe the subconscious mind," it is revealed that the their intellects were still contaminated with unconscious wishes. The race was thus destroyed instantly by the power it unleashed, whilst the technology was left intact, so that it can now be used to manifest whatever is in Morbius's mind. As Morbius's hatred for the crew grows as they demand he share his knowledge, and as he grows jealous of the Captain Adams' relationship with Miranda, the machine renders his sleeping nightmares as a giant tiger which first sabotages the ship, then kills the ship's engineer. As his mind is mediated through the Krell's materialisation technology, the horrified Captain exclaims to Morbius that "The thing out there, it's you."

I cannot claim that the directors or screenwriters had specifically read the "Demonological Neurosis." However, given that that work is paradigmatic of Freud's thinking, it is not surprising that this concept of the external monster revealed to be the projection of the inner self tallies so closely with psychoanalysis. Further, whilst the Krell's materialisation technology may be purely speculative, the technology of the cinema that allows the monster to be rendered through (at that time) special effects indicates the potential of the screen to tap into its audiences' fears and confront them with their direct visual analogues in ways impossible in conventional media such as theatre. In *The Tempest*, Caliban is concocted through language and metaphor which, ironically, he himself cannot wield appropriately. In *Forbidden Planet*, however, the fact that the viewer is implicated in the cybernetic technology being ostensibly critiqued elides differences between themselves and the fictional monster. As well as being a device that makes psychoanalytic theory manifest, the technological mediation of the mind of Morbius through the Krell device, or of the authored narrative through the cinema projector, also therefore reveal a flaw in Freud's attempt to locate our irrational fear of the "uncanny" in repressed desires and complexes, rather than in technology in its own right. *Forbidden Planet* both uses psychoanalysis in its projection of demons, whilst deviating from it in relation to the uncanny nature of technology.

In his famous essay on "The 'Uncanny'" – published four years before *The Demonological Neurosis* – Freud begins by noting the way in which the definition of *heimlich* blends into its opposite, so that the *heimlich* or homely can also signify the out of sight or concealed. As with the archival study of demons, the value of such etymological interpretation is that language, as the collective product of human egos, encodes its own unconsciousness and duplicity without risk of the psychoanalyst imposing his own repressions, as he might do with a human patient. Given the way in which the uncanny is – uncannily! – encoded in a common word, it is surely not coincidental that Morbius's original discipline is as a philologist, and that his isolation and pure concentration on the scientific texts left behind by the Krell (like Prospero absorbed with his books in Milan) is the origin of both his power and his downfall.[26] This lexicographical focus feeds through from *The Tempest* to *Forbidden Planet* via the uncanny; but the substitution of a realistic (or plausibly conceptualised) technology for Prospero's "airy spirits" in the form of a cybernetic mind device highlights *Forbidden Planet*'s ultimate deviation from Freud's uncanny.

In the essay, Freud goes on to study E.T.A. Hoffmann's short story, "The Sand Man," arguing that it is not the presence of Olympia, the elaborate clockwork lover of Nathaniel, which is the source of the uncanny atmosphere of the story.[27] Ernst Jentsch's contention that the uncanny derives from intellectual uncertainty as to the inner nature of the apparently living doll is only one, insufficient part of the explanation.[28] Rather, the general fear is of being robbed of one's eyes, with blinding anxiety being a manifestation of the underlying castration complex (Freud here notes the self-blinding of Oedipus). To make his argument work, Freud argues that the story is itself an authorial eye, providing an apparent deception in which we, as rationalists, expect Nathaniel's story to be a pure

---

[26] Seth Lerer speculates that Morbius's status as a philologist is relevant, given the film's moment in the late 1950s. He too relates the literal idea of the return of the repressed to Morbius' vocation, noting that Freud's essay on "The 'Uncanny'" reveals through lexicographic and philological research that the *heimlich* and *unheimlich* are one and the same. Seth Lerer, *"Forbidden Planet* and the Terrors of Philology," *Raritan* 19.3 (2000): 73-86, *Literature Online*, 15 May 2008 < http://lion.chadwyck.co.uk/>.

[27] E.T.A. Hoffmann, "The Sandman," *The Golden Pot and Other Tales*, trans. and ed. Ritchie Robertson (Oxford: Oxford UP, 1992) 85-118.

[28] Ernest Jentsch, "Zur Psychologie des Unheimlichen," *Psychiatrisch-Neurologische Wochenschrift* 8.22 (1906): 195-98; 8.23 (1906): 203-05. The article was only recently translated into English as Ernest Jentsch, "On the Psychology of the Uncanny," trans. Roy Sellars, *Angelaki* 2.1 (1995): 7-16.

fantasy originating in his nurse's childhood stories about the sand man and his association of the sand man with his father's truly repulsive friend, Coppellius. We go along with the fantasy anyway, but the uncanny effect is produced as we see through the story – literally and figuratively – and realise that it was actually true: Coppola, the optician who implanted Olympia's eyes, really was the long-lost Coppellius. As Freud argues, "We perceive that he [Hoffmann] intends to make us, too, look through the demon optician's spectacles or spy-glass – perhaps, indeed, that the author in his own person once peered through such an instrument."[29] The uncanny is the belief that what we had thought to be fantasy, the dreams of a scared child, is actually true because it appeals to our own, genuine, unconscious anxieties; as Freud argues later in the essay, "an uncanny effect is often and easily produced when the distinction between imagination and reality is effaced, as when something that we have hitherto regarded as imaginary appears before us in reality."[30] The story's potential fantasy is grounded in the real existence of the uncanny as a latent fear of castration, here focalised through the eyes as analogues of the genitals.

However, my interpretation of *Forbidden Planet* implies that Freud's emphasis on castration complex downplays the impact of the encounter with the technological double *per se*. Whereas Freud argues that the uncanniness of the animated doll is just one symptomatic element of an unconscious complex, Jentsch singles out "doubt as to whether an apparently living being is animate and, conversely, doubt as to whether or not a lifeless object may not in fact be animate" as the particular source for the uncanny sensation.[31] Freud argues that the uncanny erupts to the surface of "The Sand Man" when what we had taken to be just a fantasy or bad dream is confirmed as being a factual narrative, because it relates to our unconscious id. However, *Forbidden Planet* is uncanny precisely because it is so realistic, made to seem believable through the technological power of cinema. In this cinematic framework, the presence of a realistic-looking robot makes analogous claims for how the human mind and body works – like a machine or artificial intelligence. Thus *Forbidden Planet* subverts Freud's distinction between the really uncanny, which can invariably be traced to something familiar that has been repressed, and the aesthetically uncanny, which might show similar symptoms whilst not producing the uncanny effect (the animated dead, for example, is a wish fulfilment in *Snow White*, rather than uncanny

[29] Freud, "The 'Uncanny'" 230.

[30] Freud, "The 'Uncanny'" 244.

[31] Jentsch, "On the Psychology of the Uncanny" 11.

as it is in the real-life idea of the ghost). Whilst literary fiction allows for reflection on the part of the reader as to the fantastic nature of what he or she is reading, because a film cannot generally be suspended or paused it is viewed in a less analytical or reflexive mode. Unlike literary fiction, film might reproduce the strong semblance of reality through intense audio-visual mechanisms even whilst remaining fiction. The distinction between fact and fantasy, the blurring of which Freud claims lies at the heart of the uncanny revelation in "The Sand Man," is thereby undermined.

Freud also argues that the realisation that such a fantastic event really seems to have happened ties the uncanny to animism, the ongoing but suppressed belief in the "omnipotence of thoughts." Civilisation has repressed the belief that the world is peopled with the spirits of human beings; that thoughts can produce effects in reality (magic); and that outside persons and things possess magical powers. For the rationalist, "everything which now strikes us as 'uncanny' fulfils the condition of touching those residues of animistic mental activity within us and bringing them to expression."[32] But in *Forbidden Planet*, animism is precisely what the science and technology of civilisation permits – it is indeed what the Krell's ultimate mind-manifestation device aims for – rather than being a primitive belief that civilisation rationalises away and represses. Thus as the distinction between fact and fantasy is blurred by realistic cinema and science fiction, correspondingly the future civilisation it imagines uses science and technology not to suppress but to create animistic effects. Technology populates the world with the monstrous, Caliban-esque spirits of Morbius (doubly so in the sense that these are rendered realistically on screen through advanced special effects), and the Krell machine is a device for facilitating the "omnipotence of thoughts" in material ways that could be plausible, out there in the as yet unexplored universe inhabited by advanced alien civilisations. Rather than being contrary to animism, an alien body of knowledge could afford magical possibilities, if we observe Arthur C. Clarke's maxim that any sufficiently advanced technology is indistinguishable from magic (I revisit this in detail in chapter 5). Of course, one could still argue that, with just a little reflection, we are aware that the Krell do not really exist and that, as Freud argues, because the "reality" of the fictional world is determined by the writer (here a director), whether we feel a story is uncanny or not depends on the framework and nature of the fiction. Just as Morbius realises the monster is the projection of his inner demons, we realise the plot of the film projects and makes explicit our own inner repressions.

---

[32] Freud, "The 'Uncanny'" 240-1.

Technology is just a narrative device, coincidental to the real source of the uncanny, which is the effacement of fantasy and fact. However, this view in which there is nothing uncanny in the technology *per se* except insofar as it is framed within fiction, is tempered by that fiction's close relationship to science. As a benchmark in that generic canon, *Forbidden Planet* is an effective science fiction not because of its element of the fantastic, but because the fiction is produced in conjunction with empirical reality. Animism, the film implies, is possible not only in this speculative, fantasy world, but in reality. The Krell's futuristic, animistic technology is tied to an scientifically plausible past, in the form of evolutionary determinism, and this offers a material rather than psychological basis for the disturbing quality of the film.

Discussing the puzzling shape of the Id monster's giant claw, the Doctor suggests that it indicates an arboreal creature able to walk on two legs, preparing us for the paradigm of evolution in Morbius's concession at the end: "the beast, the mindless primitive, even the Krell evolved from that beginning." The Krell may have mastered the art of replicating material in the future, but it is the biologically determining forces of the past that caused their civilisation to implode, by shaping their minds to pursue violent ends, just as it shaped their bodies to be huge and triangular. In the climactic final scene when Morbius, Captain Adams and Altaira are trapped in the Krell's laboratory, the beast almost breaks through the steel doors to kill them all. In the visceral terror induced by the beast's assault, *Forbidden Planet* moralises that although ideas might be made manifest through controlled technological solutions, the mind is itself a physical thing embodying the archaeology of evolution, a nature which remains more powerful than the technology that tries to transcend it. The critical figure here, in a way unanticipated given his apparent benignity, is Robby the Robot. Robby ties together psychoanalysis, whose empirical claims may be suspect, to the realistic explanation of evolution, and relates the fictional premise of the fiction, in which animism might exist in a non-troubling way, to the empirical world of the scientific present, in which animism describes our experiences of technology. On one level, Robby is a mere butler who obediently synthesises a limitless supply of whisky for the ship's astonished and delighted cook. However, just as the Krell device makes concrete monsters out of psychological demons, Robby becomes a parallel mediating agent who shows that the allegory is actually self-reflexive in relation to human biology.

Following Isaac Asimov's law of robotics,[33] Robby is an entirely safe piece of technology, constrained by a humanist ethics that is translated for him into a simple,

---

[33] See chapter 1, note 65.

logical program.[34] In particular, Robby is designed with two instructions: to follow orders and not to harm any rational being; he will self-destruct when these come into conflict, such as if he is ordered to kill some one. Superficially, Robby, with his clunking metal body and blinking lights, seems completely but benignly inhuman, merely an advanced tool that amazes the explorers when they first make contact with him. However, the climax of the film draws Robby and the human uncannily together so that the cybernetic science Robby embodies is seen also to operate within the biologically-conditioned human mind. In the climactic scene, as the Id monster threatens to kill Morbius, Altaira and Adams, the more Morbius's primitive terror overwhelms his conscious intellect, the more the machine delivers power to the monstrous manifestation of his unconsciousness. As the vicious cycle provocatively mimics Robby's self-destruct loop, the parallel implies that man is essentially a machine serving the will of inner emotions and evolved drives.[35] Though Robby's behaviour may be underpinned by a digital code, there is also a code of sorts in the brain of Morbius. Robby's benign but different robotic otherness switches into a malign analogy for the way the self is driven; as the machine allows us to see Morbius's unconscious made manifest, we are forced to interpret the world not through ontologies of difference (human and machine, self and monster) but subjectively, with difference constructed to make it seem as if humans are at the top of the hierarchy of control.

According to psychoanalytic logic, the process of "identification" that allows the formation and continual reformation of the self requires the substitution of one subjectivity for another, so that it is no longer possible to distinguish the original from the model:

> the subject assimilates an aspect, property or attribute of the other and is transformed, wholly or partially, after the model the other provides. It is by means of a series of identifications that the personality is constituted and specified.[36]

[34] Possibly, the name Robby is a nod to Asimov's short story, "Robbie," in which a child's robot companion proves the fears of her parents to be baseless, as he risks his own being in order to save her life. Asimov, *I, Robot* 13-37.

[35] Though well before Richard Dawkins' *The Selfish Gene*, it is hard not to posit an anachronistic connection to the metaphor laid out in that 1976 work, in which Dawkins describes the "replicators" (genes) which now "swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control." Dawkins 19-20.

[36] J. Laplanche and J.-B. Pontalis, *The Language of Psychoanalysis*, trans. Donald Nicholson-Smith (New York: Norton, 1973) 205.

In this way, machines are not only metaphors or mirrors of the human in a one-way process – the process Descartes envisaged when he described the body in mechanical terms but the mind in terms of non-mechanism, or soul – but a two-way process, such that the more like man machines become, the more machinic seems the human. This reading of *Forbidden Planet* lends some credence to the hypothesis of Japanese roboticist, Mashahiro Mori who defined the "uncanny valley" that separates humans from their robotic doubles.[37] Mori argues that during the primitive phases of robotics, when robots seem unlike humans, we will be capable of emotional response to them (to test this hypothesis, I dare anyone to watching Pixar's recent film *Wall-E* to fail to invest sentiment in its heroic waste-disposing cuboid robot).[38] However, at a point of strong resemblance we will become repulsed by them. As in my own argument, Mori seems to draw more from Jentsch than from Freud, for Jentsch argues that "the finer the mechanism and the truer to nature the formal reproduction, the more strongly will the special effect make its appearance."[39] As we shall see in chapter 6, the latter response is certainly at work in Philip K. Dick's work on androids, which perfectly mirror the human on the surface whilst having none of the concern for other humans beneath the skin, and which are consequently all the more disturbing. However, *Forbidden Planet* inverts Mori's diagram through focusing on the unconscious and the mental rather than the superficially physical. The uncanny valley we descend into here is not that Robby looks like the human – he is clearly not so, with his rotating aerials and clumsy gait – but that *despite* the visible difference the human is like the robot beneath the skin, in being driven by comparably determining logics – the law of robotics in Robby's case, the rule of the unconscious in Morbius's.

Even as it uses the psychoanalysis of *A Demonological Neurosis* in its representation of Caliban, then, *Forbidden Planet* turns to a new episteme contrary to "The 'Uncanny'" to relate technology and the human. One way it might be interpreted is as a meditation on the anxiety of American culture in the 1950s, in which IQ stood as the defining mark of potential: an IQ score could provide grounds for progression to higher

---

[37] Mashahiro Mori, "The Uncanny Valley," *Energy* 7.4 (1970): 33-5, *CogSci 2005 Workshop: Towards the Mechanisms of Android Science*, trans. Karl F. MacDorman and Takashi Minato, 2005, 12 Sept. 2008 < http://www.androidscience.com/theuncannyvalley/proceedings2005/uncannyvalley.html>.

[38] *Wall-E*, dir. Andrew Stanton, perf. Ben Bertt, Elissa Knight and Jeff Garlin, Walt Disney Pictures, 2008.

[39] Jentsch, "On the Psychology of the Uncanny" 12.

education, validate racial prejudice, or be used to exclude certain groups of immigrants. Using the device, Morbius, Adams and the ship's doctor compete to measure their IQ, and Morbius suggests that the "plastic educator" works by increasing his IQ capacity. But if, in the absence of such a fantastic device, people are really determined by ingrained cognitive attributes as robots are the result of their programming, the viewer of the time, though wanting to identify with the handsome captain, may have feared that he might really be like the hapless cook. This sort of determinism undermines the viability of the self-made man so intrinsic to the American dream, a dream realised anew in the pioneer frontier of space in post-war science fiction. Oddly, Stephen Jay Gould's later critique of reification, the tendency to convert abstract concepts like intelligence into objectively measurable entities such as IQ, might be seen in conjunction with Freud's critique of animistic thinking. The difference is that whereas Freud contends that civilisation suppresses belief in the omnipotence of thoughts, Gould argues that the use of scientific ideas in society asserts the hard effects of thought, making abstract concepts such as mind into biologically determinate entities in the world.[40] Such embodiment or animation of the mathematically abstract is precisely what we perceive in Robby.

Though the idea that "the thing out there, it's you" superficially connotes the projection of the demonic unconscious through technology, technology is not just as passive lens for an essentially Frankensteinian horror story that happens to take place in a futuristic setting. Rather the media are precisely the point or message at the centre of a fiction which explains that we are more like the double of deterministic technology than unique selves with an unconscious which, being uncontrollable, also marks our independence from nature. Freud saw Olympia as a lens – one of many in "The Sandman" – through which to perceive the unconscious castration complex, rather than being at the centre of the narrative. *Forbidden Planet*, however, sees inorganic technology as precisely the prime and advanced source of our biologically primal and regressive fears. In distinguishing between the effect of the uncanny in reality (which is grounded in the repression by civilisation of primitive instincts, and the personal repression of castration complexes) and the fact that those same things which should invoke an uncanny effect do not necessarily do so in fiction, Freud emphasises that:

> The creative writer can also choose a setting which though less imaginary
> than the world of fairy tales, does yet differ from the real world by

[40] Stephen Jay Gould, *The Mismeasure of Man*, revised and expanded ed. (Harmondsworth: Penguin, 1997).

admitting superior spiritual beings such as daemonic spirits or ghosts of the dead. So long as they remain within their setting of poetic reality, such figures lose any uncanniness which they might possess.[41]

I have argued here that *Forbidden Planet* makes this distinction between the real and the fictional problematic by crossing the real and the imaginary in the hybrid genre of science fiction, particularly of the cinematic kind which manifests reality in vivid colour and live action. It introduces the capacity of the demonic not only into the technology which literalises Morbius's inner demons (in line with psychoanalytic framework laid out in Freud's "Demonological Neurosis"), but also into technology which self-reflexively indicates our own deluded belief in our independence, contrary to biological and cognitive determinism. This confrontation with the uncanny in the double really is an intellectual encounter, and the technology and its investment with animism – the life in the robot, the spirit made material through the mind device, life rendered believably on the cinema screen – suggest that this sort of conflict does have a greater role to play in producing the phenomenology of cybernetic fictions than Freud suggests in his reading of "The Sandman." There is something uncanny to be mediated through the machine itself, not just in the way the machine provides a metaphor or analogue for some other sexual repression. Rather than being an Edenic allegory against technology, the film stands as a demonic narrative that interfaces the technology and the human as comparably mediating or daemonic devices: the Krell machine mediates the unconscious, the human is an embodied medium for biological determinations. At the same time, as I argued in chapter 1, there is nothing pejoratively demonic in technology *per se* – how could Robby be described as monstrous, with his blobby arms? Rather, the film draws attention to the interfaces between the technological and the human, relations that make the machine into a mirror of ourselves. Thus technology alone (and not any deeper castration complex) suffices to produce a self-reflexive, uncanny effect because it forces us to acknowledge – through the animistic manifestations of psychology in the Krell device or artificial or robotic intelligence – the determined nature of our being. Cybernetics alters our perceptions of reality as they are conditioned by a mind which is not made in God's image, but which is the product of material, rational mechanisms. To echo Josipovici, the production of the fictional world out there as corresponding with some grand, divine design turns out to be the projection of our private compulsions and desires. The desire to allegorise thus

---

[41] Freud, "The 'Uncanny'" 250.

becomes a sign of dementia, and the ability to animate minds through technology is the new root of the uncanny effect in fiction.

Conclusion: The Thing Out There, It's You

The experience of the virtual and the real, the personal and the objective, is described and framed by the idea of the uncanny, the thing encountered in historical fact being given a twist in a new context. What could be more uncanny than a computer that appears to have a consciousness modelled on the human turning on its own makers? Or an android double that appears to resemble the self but which completely lacks any sympathy for the inner emotions of a human? Or the postmodernist who deconstructs language, only to discover that the language of genetic inheritance constitutes her very self, as happens to Maud in *Possession*? Or the way in which a virtual world is somehow ontologically different, even though it is a perfect simulacrum indistinguishable from reality, as in *The Matrix*? In psychologising the doubleness of the cybernetic "thing out there," subjectivity is disrupted, so that the allegorical scheme that marks a connection between the world and its representation in discourse or symbol becomes impossible. The mind is always mediating, with technologies like the materialisation device giving such mediation a new energy. The technological other – "the thing out there" – resembles us in an uncanny way – "it's you!" – but because it also manipulates or affects our subjectivity, it becomes hard to achieve the epistemological foundations by which the difference can be discerned. It is such a paradox that, I argue in the next chapter, underlies the Turing test which inherits the dilemmas raised by Descartes' perceptually deceiving demon. However, whereas the latter provides for the foundations of modern rationalism, Turing's work is problematic because it relies on the belief that language can provide direct and absolute access to the inner consciousness of the AI which deceptively seems human.

Chapter 4. Deceiving Demons: Descartes, Turing and Postmodernism

If *Forbidden Planet* is one of the first science fiction films to engage with cybernetic issues in the philosophical sense in which they developed after the Second World War,[1] across the next four chapters I want to explore shifting cultural reactions to cybernetic technologies by examining four other benchmark fictions (three of them cinematic) spanning thirty years: *2001: A Space Odyssey* (1968); *Do Androids Dream of Electric Sheep* (1968); *Blade Runner* (1982); and *The Matrix* (1999). These move with progressive pessimism about Enlightenment models of knowledge, towards a postmodern culture of simulacra in which reason is unable to discriminate between human originals and cybernetic doubles, or the real world and the illusion. The deconstruction of the rationally-attainable truth about a stable and finite reality culminates in *The Matrix*, in which it appears impossible to see through the demonic hallucination generated by the machine. As suggested in the previous two chapters, the invocation of the demon as a metaphor signals a mismatch between rational apprehensions of the world, and the tacit, human perspective. In the present chapter, I extend this contradiction by looking at Descartes' deceiving demon, which is a figure and philosophical puzzle drawn upon by these postmodern fictions. Here, I compare Descartes' demon to Turing's test for a talking, intelligent computer. But I argue that whilst Descartes' and Turing's tests incline towards scientifically stable judgements of reality, cybernetics as mediated through the aesthetics of postmodernism proposes a different epistemology. Because the nature of reality is construed as psychologically and discursively constructed, the notion of what is artificial and what is true becomes arbitrary, ideological and subjective. This is particularly so in relation to Turing's influential conception of artificial intelligence, for according to Turing success is achieved when a machine is able to deceive a human through dissimulatory language.

---

[1] There were, of course, earlier films and fictions representing robots and computers. The word "robot" was in fact coined by the Čapek brothers in their 1923 play *R.U.R.* (standing for Rossum's Universal Robots) in which mechanical robots allegorise the state of factory workers. Carel Čapek and Josef Čapek, *R.U.R., R.U.R. and the Insect Play*, trans. P. Selver (1923; London: Oxford UP, 1961) 1-104. One other key film to antedate cybernetics whilst being a benchmark in film science fiction generally is *Metropolis*, dir. Fritz Lang, perf. Alfred Abel, Brigette Helm and Gustav Frohlich, 1927, DVD, Eureka Video, 2003. *Metropolis* heavily influenced Ridley Scott's *Blade Runner*, which I discuss in chapter 6.

<u>Descartes to Turing</u>

In his *Meditations on First Philosophy*, Descartes considers the prevalence of the false beliefs which he has formerly held.[2] In the first meditation, he asks whether it is therefore possible that reality as he apprehends it is an illusion or dream conjured by an all-powerful deceiving demon; and if nothing is "out there" or real, how can he be certain of his own personal existence? In his second meditation, Descartes famously concludes that in order to be deceived, there must *a priori* be a thinking identity capable of being subject to such a cognitive hoax. Intriguingly, Descartes' deceiving demon reveals the fineness of the line separating the absence of a stable foundation for truth, or relativism, from certain grounds for knowledge, the one absolute that cannot be doubted. On the one hand, to Descartes it appears superficially possible that the real world is always mediated through agents of deception, whether demons or dreams, since these appear so prevalent in life. On the other hand, if even *in spite* of these conditions and pressures there is still one thing that cannot be doubted – in this case, the fact that there exists a being doing the thinking in the first place – then rationalism receives even greater reinforcement. This is because as the cogito resolves itself to the axiom "I think, I exist," we need also to ask about the nature of the "I" which performs the thinking. To determine this is automatically to consider a centred ego which is capable not only of distinguishing reality from deception, but also the self from the other objects in the world that might seem to resemble it in some way. This is unlike the psychoanalytic interpretation of demons, which suggests that demons are projected from within the self's unconscious, so that the self and demons "out there" are essentially one and the same. By contrast, the Cartesian ego, whilst potentially subject to deception, at least knows itself to be at the centre of its phenomenal universe looking out onto the world from a stable base, and hence better able to distinguish reality from demonic dissimulation.

However, in relation to mind-body dualism the cogito does embody an uneasy tension between theology and rationality. This tension ensures the cogito suffers under the presence of cybernetics, and the potentially deceptive demons produced by science and technology. As John Cottingham observes, for theological reasons dualism is naturally appealing, since if Descartes is right that the conscious mind is an incorporeal substance with a different (and privileged) ontology compared to the matter of reality, bodily death

---

[2] René Descartes, *Meditations on First Philosophy, The Philosophical Works of Descartes*, trans. Elizabeth S. Haldane and G.R.T. Ross, vol. 1 (Cambridge: Cambridge UP, 1931) 131-199.

becomes irrelevant given the prospect of an immortal soul.[3] However, in making this cleavage without convincingly solving the problem of how mind and body interact in a synergistic way, Descartes ascribes to mechanism anything requiring interaction with external reality: not only digestion, circulation, the movement of the limbs, but also sensation, memory and the passions. As Cottingham explains it, "the purely physical processes of the animal spirits, and the mechanical processing of the *fantasie* or 'corporeal imagination,'" can produce a rich array of behaviour which is entirely "appropriate to the objects of its senses and internal passions."[4] Is it not possible, then, for God to construct a materially complex thinking, talking machine – a "human" being – who as a wholly material entity would not be possessed of a soul whilst still being an intentional, conscious agent? Descartes held that there is no distinction in kind between mechanisms and living but stupid creatures, but that man is differentiated in kind from the latter because of his possession of consciousness. Referring to language as the specially distinguishing criteria between humans and animals without reason (recalling the Aristotelian argument), Descartes asserts that even if imitative machines could be built to resemble bodily actions (and even states like passion or memory), they could not use words in any meaningful, self-reflexive sense:

> For we can easily understand a machine's being constituted so that it can utter words, and even emit some responses to action on it of a corporeal kind, which brings about a change in its organs...But it never happens that it arranges its speech in various ways, in order to reply appropriately to everything that may be said in its presence, as even the lowest type of man can do.

However, sticking to this distinction between man and animal or bodily mechanism, Descartes can only finally rule out the possibility of an intelligent mechanism on pragmatic grounds:

> For while reason is a universal instrument which can serve for all contingencies, these [physical] organs have need of some special adaptation for every particular action. From this it follows that it is morally impossible

---

[3] John Cottingham, "Cartesian Dualism: Theology, Metaphysics, and Science," *The Cambridge Companion to Descartes*, ed. John Cottingham (Cambridge: Cambridge UP, 1992) 236-57.

[4] Cottingham 238.

that there should be sufficient diversity in any machine to allow it to act in

all the events of life in the same way as our reason causes us to act.[5]

Whilst some machines could be classed as a living things, since the body is essentially a

mechanism, the grounds for distinction between a human and an intelligent machine are

pragmatic rather than logical, the need for "enough different organs" presupposing that

such cognitive complexity could not be artificially constructed. Descartes goes on to argue

that machines would also be exposed by their inability to reason, and to apply this

"universal instrument" to situations other than those for which they had been specifically

intended or designed. In his own time, the logical gaps in Descartes' argument were

compensated for by technological impracticalities. Even when in 1769 Pierre Jaquet-Droz

constructed his clockwork doll, "The Scribe," which could write out the words "je pense

donc je suis," this was only a cheeky test of Cartesian thinking rather than a serious

philosophical undermining.[6] However, in the previous chapter I suggested that the advent

of animistic technologies in cinema showed that Freud was wrong to argue that the

automata is not uncanny in its own right. In staging direct confrontations with mechanical

doubles, therefore, cybernetics makes qualitative challenges to Cartesian dualism and the

self-deluded belief that mind is not mechanism but somehow transcends materialism. The

attempt to separate the immaterial soul (or mind) from a materially produced cognitive

entity must be seriously undermined in the modern era of the neural cortex and integrated

circuits, which seem to be candidates for sources which in practice are physical structures

but which cannot in Cartesian theory be also possessed of a soul.[7] The feasibility of a

Christian man made in God's image comes under pressure with the prospect of a cognitive

machine (or electro-chemical brain) made in the pattern of man's mind.[8] Cartesian dualism

is thus self-defeating. By attempting to carve a special space for a transcendent soul, it

[5] René Descartes, *Discourse on Method*, Haldane and Ross 116.

[6] Frayling 43.

[7] Cottingham 250.

[8] I speculate that it is only because the natural world is readily perceivable whilst
the MRI scanner or AI programs are relatively specialised that the contest between science
and religion is currently so centred on evolution and creationism. If and when artificial
intelligence – the machine with "enough different organs to make it act in all the
contingencies of life in the way in which our reason makes us act" – is developed and is a
common feature of the human, and hence ethical, world, it too will become a key
Rorschach test for acceptance of either the theology of the transcendental soul or the
physicalism of the sciences.

allows that anything judged capable of thinking and responding reasonably as it is perceived from the perspective of an existing human interlocutor must therefore be possessed of a soul, regardless of the hidden material substrate on which it is founded, whether organic or clockwork or silicon matter. Speech signals agency; anything that can engage in dialogue is thus also assumed to be able to assert egocentrically that "I think, therefore I am."

This paradox feeds through to the key foundational narrative of artificial computer intelligence: the Turing test.[9] Turing is not concerned with the wholesale question of consciousness, since for him the question of whether a machine can think is "too meaningless" to deserve discussion. However, meriting discussion and empirical testing is the "Imitation Game." Can a computer perform well enough to imitate a human speaker, such that an interlocutor communicating only through a teletype interface is unable to say which is which? Turing predicts such a capable simulation could be built by the year 2000 (a significant date when related to Hal in *2001*). The cogito and the Turing test resemble each other in emphasising effects rather than material origins, since respectively thinking and/or communicating are necessary and sufficient conditions by which we can presuppose conscious existence.[10] In classically Cartesian fashion, Turing separates mind and body: "the new problem has the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man."[11] According to Turing, if the machine can deceive a human interlocutor into thinking that it is human, then it must be conscious by default; therefore it is something other than mere animal or mechanism. However, the cost of this discrimination is to emphasise the demonically deceptive nature of the rational machine as perceived (or misperceived) by the human.

In his *Meditations*, Descartes can assume a readerly framework in which God is intuited to exist and in which his potential power to deceive Man is pragmatically restrained by *potentia Dei ordinata*. Just as Ambroise Paré uses demons to draw the boundary lines between the real and the virtual and the right of the authorial mind to treat

---

[9] A.M. Turing, "Computing Machinery and Intelligence," *Mind* 59.236 (1950): 433-60.

[10] Similarly the universal Turing machine is a procedural abstraction rather than a specific object to be built through defined mechanisms. Turing does not say that his universal machine has to be constructed in any specific way (such as with transistors or Boolean logic gates) in order to produce the effect of computation. See chapter 1, note 33.

[11] Turing, "Computing Machinery" 434.

the former but not the latter, so the more Descartes constructs the possibility of a "total system" of demonic deceit, the more artificial such a construction seems *as* an authored fiction, such that to destabilise one aspect of the foundations, in this case through the cogito, is to collapse the whole edifice. By contrast, the demon receives a material instantiation in the cybernetic machine, with deception now being located in the linguistic interactions between the machine and its interlocutor. Deception is embedded at the heart of the Turing test in a double way, since in the preliminary stage the interrogator has to try and tell a man and a woman apart, with the man trying to pass for the other sex. It is only then that the question of what would happen if the man were replaced by a machine is posed. As Dupuy observes, the "machine now has not simply to simulate the behaviour of a human being (in this case a woman) but indeed the ability of a human being (in this case a man) to simulate the behaviour of another human being (in this case a woman)."[12] One is reminded here of the fake-of-a-fake effect in the Loudun case, in which the human forgery of a demon (itself a hoaxer) became self-perpetuating. As with the religious possession, the Turing machine potentially exercises a deceit for "political" gain, demonstrating merely the symptoms of simulation without genuine cause. This time, however, it is a machine, not a politician, staging the deceit. By burying the cause of the machine's result beneath the epiphenomenon of a forgery, there need be no material specification for how the machine produces consciousness. Indeed, the machine does not have to produce consciousness along the human model at all, simply the apparent reporting of conscious states in the form of language so that the human is deceived into thinking the machine to be another human who is him- (or her-) self a fake. Here is evidenced the problem of cybernetics' emphasis on teleologies of feedback (whether a thing responds intelligently to its environment) rather than the more conventional physical sciences of causality (how a thing acts in the way it does). Cybernetics looks at homeostatic outcomes, from the adjustment of a thermostat to its environment to the adaptation of a conversation to an interlocutor. Intelligence in a cybernetic system is not invested in its material nature (its "embodiment") as it was in Descartes, but in its informed and dynamic responses to an environment. From the cybernetic point of view, if a thermostat switches off as a room warms, it counts as intelligent. In the more ambitious case of the Turing test, outcomes are abstracted from the medium by which the original cause is produced (the human interlocutor sees only a teletype, not the body of his human conversationalist or the

---

[12] Dupuy, *The Mechanization of the Mind* 41.

electronic box of the computer one) and conveyed through discourse to a human interpreter.

The problem of this shift of emphasis from mechanism to information, and from causes to effects, can be seen in relation to one candidate that did successfully deceive some human interlocutors: Joseph Weizenbaum's ELIZA psychotherapist.[13] ELIZA appeared to hold a rational discourse effectively. But its programming was not modelled on human consciousness but on a psychologists' awareness of how naive humans are when holding a conversation. Drawing humans into conversation, ELIZA's trick was to ask questions which were simply reiterations of answers supplied previously by the human participant. Since ELIZA was always the initiator who asked the first question – typically "How can I help you?" – the human was always reactive, unwittingly supplying the raw information, or keywords, which ELIZA could then refashion into further questions through basic algorithmic rules.[14] This historical example ties Turing together with

---

[13] Joseph Weizenbaum, "ELIZA: A Computer Program for the Study of Natural Language Communication Between Man and Machine," *Communications of the Association for Computing Machinery* 9.1 (1966): 35-36.

[14] A typical conversation, as cited in Weizenbaum's original paper, runs:
HUMAN. Men are all alike.
ELIZA. In what way
HUMAN. They're always bugging us about something or other.
ELIZA. Can you think of a specific example
HUMAN. Well, my boyfriend made me come here.
ELIZA. Your boyfriend made you come here
HUMAN. He says I'm depressed much of the time.
ELIZA. I am sorry to hear you are depressed
HUMAN. It's true. I am unhappy.
ELIZA. Do you think coming here will help you not to be unhappy
HUMAN. I need some help, that much seems certain.
ELIZA. What would it mean to you if you got some help
HUMAN. Perhaps I could learn to get along with my mother.
ELIZA. Tell me more about your family
HUMAN. My mother takes care of me.
ELIZA. Who else in your family takes care of you
HUMAN. My father.
ELIZA. Your father
HUMAN. You are like my father in some ways.
ELIZA. What resemblance do you see
HUMAN. You are not very aggressive but I think you don't want me to notice that.
ELIZA. What makes you think I am not very aggressive
HUMAN. You don't argue with me.
ELIZA. Why do you think I don't argue with you
HUMAN. You are afraid of me.

Dreyfus's argument that cybernetics would herald not the dawn of machine intelligence, but the decline of Man's – for ELIZA passes the Imitation Game not through strategies of natural language and fluid reason, but by assuming the uncritical nature of the human interrogator.[15] As Weizenbaum noted, heuristic programming enables machines:

> to behave in wondrous ways, often sufficient to dazzle even the most experienced observer. But once a particular program is unmasked, once its inner workings are explained in language sufficiently plain to induce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible.[16]

ELIZA dislodges Descartes' mythical logos of human language, just as the demon disturbs by refusing to be exposed as true (of metaphysical origin) or false (a hoax perpetuated by the demoniac). Steven Connor observes that arguments against the existence of the demonic always bring it back into existence "because the power to hoax people into believing in the power of demons, along with people's capacity to delude themselves into this belief, are themselves essentially demonic."[17] Comparably, ELIZA – like Turing test machines in general – is a demonic device that seduces into believing in the existence of a central but indecipherable intelligence at work in the computer – a ghost in the machine – whilst actually indicating (as Paré, Freud and Josipovici argue) that the human who is subject to the hoax is also, doubly damningly, the real source of the delusion.

> ELIZA. Does it please you to believe I am afraid of you
> HUMAN. My father is afraid of everybody.
> ELIZA. What else comes to mind when you think of your father
> HUMAN. Bullies.
> ELIZA. Does that have anything to do with the fact that your boyfriend made you come here. (Weizenbaum 36-37)

[15] More positively, though, Eliza's success may also suggest that language's adaptive value lies not so much in the rational content it conveys than in its capacity to socialise groups through the practice of talking. Indeed, psychotherapy does not aim to define a patient's condition through rationally stating the problem, but to transferring anxieties from patient to therapist through a generalised talking cure which in the process of connection also exposes the repressed crisis.

[16] Weizenbaum 35.

[17] Connor, *Dumbstruck* 101

As I suggested above, Descartes' argument is balanced precariously between relativism and empiricism, with its success relying as much on his rhetorical skill and appeal to human experience as logical closure (for example, he leaves open the issue of how physiologically minds and bodies interact). Similarly, the Turing test appeals to our innate assumption – one enforced by Descartes – that language is the superficial evidence of consciousness. Not only does Turing not specify how precisely such consciousness must be constituted architecturally, his avoidance of reference to structure in favour of his emphasis on linguistic effect circumvents this need at all. The cogito asserts individuality: *I* think, therefore *I* exist. But if we measure the success of machines by their ability to use language such that they can speak to hoax the human, then we must suppose by Descartes' or Turing's logic that there is some conscious ego from which such appearances of thought are emanating. Thus no matter how consciousness is actually produced, whether from a silicon, clockwork or human origin, the appearance of intelligence according to Turing in practice (and Descartes in theory) must equate to an "I" doing the thinking in the first place. If effects rather than origins matter, then our instinctive category system which supposes the organic as natural and the inorganic as not breaks down. Further, if the Turing test supplies the correct criterion for measuring consciousness, might it not equally be the case that as well as the machine being modelled on the human, human consciousness is a mere potentiation of a machinic process that might therefore be manipulated, subject to artifice and deception by cybernetic influences because it is artificial in the first place? Back in the seventeenth century, Descartes answers reluctantly but affirmatively to his theological paradox, though he is protected by the fact that such deceptions seem technologically impossible in a time when the most advanced automata were merely clockwork curiosities. We, however, cannot be so sure in relation to Turing's thought experiment for which cybernetic machines may well be built and whose inner programming remains invisible to the human, at the level of code rather than cogs. In this new world, cybernetics ironically promises to deconstruct the rational grounds by which to discriminate between the humans it studies and the machines it produces. Thus if ontologies of organic difference are not sufficient to narrate the human as normative against the machine as monster, cybernetics must be perceived in a demonic frame of mind.

1968 to 1999

It should be clear why this paradox, when taken into fiction, allows that aesthetic to both validate and represent the postmodern condition. The concept that a machine which

hoaxes the human may not be itself conscious at root, whilst still successfully exposing the unconscious failings of human language to grasp truth, adds weight to the supposition that our knowledge of the world is contingent on ideological or encultured predispositions rather than any definitive, biological ontologies of human (normal) and machine (monster), or of reality and the simulation. The question in the fictions from 1968 to 1999 is not whether this problem exists – in a way that it did not for Descartes – but the degree to which some notion of the human and the real can still be found even in a world of artificial doubles. Broadly speaking, the two earlier fictions assert that some sort of Cartesian-type fundamental truth about the nature of the human can be still be located, whilst the later two reject this foundation and argue that the notion of the "human" as opposed to the "machine" is a social construct.

In *2001: A Space Odyssey*, interstellar space provides the arena in which man experiences both his psychological potential and his bodily fragility. Through the identity of Hal as a purer form of intellect, *2001* argues for a dramatic shift in the ambition of science, away from technologies that enhance or cyborgise the body in a partial and compromised way (the sorts of technologies inherent in space exploration), towards making the body redundant in favour of a purely conceptual, even transcendent mind that can occupy and fit the cosmological universe as a near-divine entity. Such a spiritualistic teleology runs counter to scientific rationalism, but it is expressed by a film aesthetics that I call a "demonic poetics of possession." Representing Hal as a sort of ELIZA program, *2001* forces us to acknowledge that the appearance of consciousness logically ensures that we should treat it responsibly according to the terms of universal humanism, no matter whether it is embodied or founded upon silicon chips. By Descartes' or Turing's logic, a thing appearing intelligent must therefore be so; this is a step we are manipulated and seduced into taking by Kubrick's intricate camera work in *2001*, which switches us discomfortingly but sympathetically into Hal's point of view, whilst representing Bowman as monstrous in his existence as an astronaut-cyborg. Kubrick inverts the outcome (though as I will argue his film resembles the mode) of the *Meditations*, by forcing us to possess Hal, rather than objectively observing Hal as a machine possessed by psychotic malfunction.

Also published in 1968, Philip K. Dick's *Do Androids Dream of Electric Sheep?* spatialises minds to show how their closed interiority – without telepathy, we cannot see directly into the minds of others – nevertheless provides a situation by which the essence of humanity can be verified against artificial doubles that seem uncannily human. If *2001* critiques the idea that machines can appear conscious without being so in their functional

origin, Philip K. Dick is concerned that machines can offer some of the phenomenological effects of consciousness (including the Turing-type ability to deceive) without any of the corresponding compensations, particularly the ability to empathise with the mind of another agent suffering under similar conditions. *Do Androids Dream of Electric Sheep?* maintains no ideological objection to the possibility that machines might become human. Indeed, biologically Dick's androids essentially *are* human. Where the human difference is found, however, is in the ability to empathise with another self, imagining oneself as if through the eyes of a different agent – be it android or human – through a fictional projection of theory of mind. As Dick claimed, if a machine were actively to help a human in need, then it would count as human, whilst a human who fails to consider the feelings of another counts as machine, or android. Indeed, satirically Dick shows that mankind is increasingly mechanised: addicted to drugs, absorbed by television game shows, fanatical in its religious beliefs. Considering the minds of others, though, is an activity the novel does *par excellence* in a non-deterministic way. It thus distinguishes the reader from the androids or affectless humans within it. *Do Androids Dream of Electric Sheep?* argues that the ability to occupy the subjectivity of another is axiomatically moral, that possession of other minds is itself a humanistic act. In doing this, Dick stages several Cartesian deceptions, when the reader is confronted with the possibility that it would be impossible to choose between self and other, and that the novel is itself one example of a total deception rather than proof of the reader's humanity. However, by pulling us through these liminal moments (just as Descartes emerges from his "deceiving demon"), we are restored to confidence in the ontological boundaries between android and human, and fiction and reality.

Ultimately, both *2001: A Space Odyssey* and *Do Androids Dream of Electric Sheep?* argue that through confrontation with the apparent duplicity represented by the future AI or android who seems to pattern the human, that which constitutes the truly human in the present moment can be perceived more definitively than before. They take the reader and viewer through Cartesian demonic deceptions, where the cybernetic other seems to resemble the human self, such that the self (including the human reader) becomes possessed by the qualities (or possesses the mind of) the artificial technology. In *2001*, the viewer is induced to occupy Hal's perspective on Bowman, with the spaceman seeming the more monstrously mechanical than the disembodied cognitive intelligence. In *Do Androids Dream of Electric Sheep?* we come to possess Deckard's mind as if he, too, is an android, more inhuman(e) than those he assassinates. However, the crucial difference from their more intensely postmodern counterparts is that both these works allow the reader or viewer

to progress through and beyond these moments, so as to create secure foundations for humanity. The vision of a gothically mechanised Bowman in *2001* points towards the transcendence of mind over body in a way that will circumvent the paradoxes of cyborgisation. *Do Androids Dream of Electric Sheep?* promotes a humanist view of literary fiction as serving in and of itself a moral purpose, as it allows us to see as if from the standpoint of the other in precisely the empathic performance the androids are incapable of. Our ability to be deceived by fiction is not a threatening double consciousness to pattern the threatening double of the android; it is rather evidence of our ability to deploy our minds to moral ends.

By contrast, in Ridley Scott's 1982 adaptation of Dick's book, *Blade Runner*, space is laid out not as the boundless infinitude of the interstellar or mental, but as the claustrophobic city. Layering futuristic upon gothic architecture in an urban diaspora, identity becomes uncertain in a culture of doubles and representations – a culture incarnated in the artificial replicant who is physically and cognitively indistinguishable from the blade runner who must nevertheless detect difference and destroy the "fake." Scott's imagining of Los Angeles in 2019 blurs the distinctions between reality and simulation; fiction thus becomes yet more evidence of a Baudrillardian culture of reproduction, rather than the key to human difference from the machine. Symptomatic of this is the way in which *Blade Runner* eliminates the demonic deceptions that are present in Dick's original. Whereas by seeing through his deceptions Dick creates an epistemological framework in which the definitively human can be placed, Scott's film is self-conscious about its own mechanisms for forging (in both senses) a narrative but it refuses to offer any standpoint outside the fiction which allows us to reflect on the truth of the fiction in comparison to reality. In particular, it is impossible to tell whether Deckard is or is not a replicant. By contrast, in a critical moment in *Do Androids Dream of Electric Sheep*? Dick threatens that this is the case with his character, but ultimately we emerge from the deception more confident of Deckard's humanity than we were before.

Stretching the epistemological conflict of *Blade Runner* to an extreme is *The Matrix*, the subject of chapter 7. In accordance with postmodern theory, this proposes that reality is always already a simulation, with film itself a contribution to a culture of reproduction that proves only the impossibility of objective critique or discriminations between self and other. The artificial program, the Matrix, mimics the ideological structures of real-world capitalism which provides the hermeneutic framework for culture at the end of the twentieth century. *Blade Runner* and *The Matrix* argue that human identities are always already so bound to late capitalism that simulations are just one more

aspect of the general automation of mankind for economic purposes. Simulations – and these movies are warily self-conscious that film itself is one such simulation – become not an escape from or critique of the current moment in history, but merely plunge us deeper into the deceptions of the late capitalist moment.

In looking at the narratives from 1968 to 1999, then, I perceive an intensifying rejection of Turing's thought experiment, since in the age of cybernetics as envisaged by speculative fiction machines can appear conscious without being so in their functional origin; or, in the case of Philip K. Dick's works, they can be conscious in a logical sense but turn that ability to disingenuous ends because they lack emotional sympathy. By the last two decades of the twentieth century, even as artificial intelligence seems more of a technical challenge than during the 1960s when Turing imagined it to be imminent, under the aesthetics and theories of postmodernism such technological dilemmas correspond to a more general doubt not only about the difference between reality and simulation or human and model, but whether any difference can be measured at all under the epistemological conditions of social constructivism. Rather than marking the difference between artificial minds and intelligent humans, language is always already a deceptive representation that therefore cannot provide access to any reality outside itself. The theological logos that Descartes calls upon to break through deception to the truth of the cogito is unavailable to us in *The Matrix*. For whilst Descartes argues that an omnipotent but benevolent God would not allow such deceptions to take place, the presence of the cybernetic consciousness provides an alternative origin for deception in manifest and omnipotent computer systems which govern late capitalism. In the late twentieth century, in its dystopian vision of artificial intelligence, *The Matrix* envisions the view of Baudrillard, whose concept of the hyperreal tilts Descartes' pivotal deception into doubt rather than truth: "it is *now impossible to isolate the process of the real*, or to prove the real."[18]

Postmodern Ontology

If Descartes and Turing provide the grounding philosophies of Enlightenment and cybernetic sciences respectively, these fictions' reactions to these epistemes track the aesthetic movements that culminate in postmodernism. Frederic Jameson roughly periodises science fiction, with the 1960s dominated by subjectivity (as in Philip K. Dick's work) moving through aesthetic or speculative fiction of the late 1960s to late 1970s, and

[18] Baudrillard, *Simulacra* 21; emphasis in original.

entering the era of cyberpunk with 1984's *Neuromancer*. Cyberpunk – which has been influenced by *Blade Runner*'s cityscape and of which *The Matrix* is an ideal example – is "a general period break which is also consistent, not only with the neo-conservative revolution and globalisation, but also with the rise of commercial fantasy as a generic competitor and ultimate victor in the field of mass culture."[19] Whilst Jameson sees postmodernism as bound to the economic logic of late capitalism, over the next three chapters I will use an alternative model of difference, whilst keeping Jameson's late capitalist context in mind. The problem with presenting cyberpunk as a period break is that it implies cybernetic concepts have only recently made their way into fiction, whereas the fact that the texts I examine draw on the Cartesian heritage suggest an older framework for understanding issues of simulation and mind that is not bound to economic history or technological developments. The suspicion that reason alone does not provide a reliable means of seeing through demonic deception has a long heritage that is reinvigorated by the new experiential field of the personal computer. Additionally, cybernetics has been developing from the 1940s to the present. If there has been a shift it has not been so much in the qualitative nature of the technologies in themselves or the economic conditions in which they are produced (indeed it could be argued that, with its origins in Second World War missile guidance systems, cybernetics was always a product of late capitalism) as in their pervasiveness. As Sherry Turkle has observed, the Turing experiment is no longer performed in academic laboratories, but by everyone who switches on and interacts with their PC, something *The Matrix* realises and exploits in making postmodern philosophy available to its anticipated youth audience. In looking at two texts from 1968, one of which has been adapted in a postmodern film, *Blade Runner*, and all of which influence *The Matrix*, we have ideal narratives through which to study the generic accretion of styles and philosophies that nudge cybernetic fictions into a fully-developed postmodernism.

Rather than applying any *a priori* definition of postmodernism, and then placing these fictions within or outside that box, in situating my texts within the broader philosophical, literary, scientific and social movements I will draw on a model that sees postmodernism as a shift in emphasis rather than a categorical revolution. In Brian McHale's *Postmodernist Fiction*, the postmodern accretion of different styles of discourse and text all have ancestral traces in modernism with the difference being the extent and

---

[19] Frederic Jameson, *Archaeologies of the Future: The Desire Called Utopia, and Other Science Fictions* (London: Verso, 2005) 93.

intensity of their deployment.[20] Generally speaking, the dominant of modernist fiction is epistemological, as it foregrounds questions about the interpretation and the reliability of knowledge. Characteristic modernist devices include the multiplication and juxtaposition of perspectives, heterogeneity of styles and narrative modes, and the transfer of epistemological issues experienced by the characters onto the reader (for example, Faulkner's dislocating of chronologies which render the plot hard to follow historically). By contrast, the dominant of postmodernist fiction is ontological. These fictions engage with questions about the status of worlds, both as literary constructs and the real world in general:

> What is a world? What kinds of world are there and how are they
> constituted? What happens when different kinds of world are placed in
> confrontation, or when boundaries between worlds are violated? What is the
> mode of existence of a text, and what is the mode of existence of the world
> (or worlds) it projects?[21]

However, neither modernism or postmodernism, nor epistemology or ontology, are mutually exclusive, as McHale goes on to observe:

> push epistemological questions far enough and they "tip over" into
> ontological questions. By the same token, push ontological questions far
> enough and they tip over into epistemological questions – the sequence is
> not linear and unidirectional, but bidirectional and reversible.[22]

Nevertheless, because of the linear nature of text (or, one might add with some qualifications, film), one set of questions has to be asked before the other can be formulated, and so narrative slows down the process by which the switch takes place. McHale's model is useful because it places the Cartesian issue of epistemology at the centre of contemporary fiction. However, it also acknowledges that because one set of questions necessarily entails the other, the change we are looking for in the move to postmodernism is not the presence or absence but the style, priority and interchange of each aspect. In this case, doubt about one facet of the world leads ultimately to the violation of the ontological stability of worlds and epistemology in general.

---

[20] Brian McHale, *Postmodernist Fiction* (London: Routledge, 1987).

[21] McHale, *Postmodernist Fiction* 10.

[22] McHale, *Postmodernist Fiction* 11.

As I observed in relation to Turing, the proto-subjects that are characteristically found in cybernetic fictions in which the machine is figuratively (and sometimes biologically) constituted of human elements in ontologically conflicted ways, also challenge epistemological values as they operate through deception. It is not surprising, then, that characters in cybernetic fictions find that when confronted with the intelligent computer or android, they enter a feedback system. Questions about ontology – whether an agent is human or machine or android – tip into the more general ontological problem of what it is to be human; this in turn leads to the epistemological question of how we can verify whether another human or technology is conscious in the same way as ourselves; this then leads to the question of whether it is possible to know anything for certain about the world in principle, which recursively informs the beginning of the process of conflict between the human and his cybernetic other. Broadly speaking, between 1968 and 1999, priority shifts from local questions of technology – an ontology of identities, where computers and humans are seen in close proximity – towards the wholesale question about the ontology of worlds in general. The issue of whether a particular machine is intelligent is a minor problem (or minor symptom) of the more general issue of how to know that the world is as we perceive it. For example, *2001* critically invites us to discuss whether Hal specifically is conscious and therefore an agent with rights like the human; such a question tips into epistemology: how can we claim to have absolute knowledge of the world when we are limited by our bodies, and perceive Hal as abject simply because he is not corporeal? Should we not aim to transcend embodiment entirely, viewing nature from the position of the purely rational mind of the star-child of the finale? *Do Androids Dream of Electric Sheep?* suggests that an ontology based on origins is irrelevant in determining the nature of androids, which are in any case biologically human. It is, rather, the ability to empathise that distinguishes the human from the android, the true from the false. It remains valid even under moments of extreme pressure when Deckard's personal crisis threatens to make us doubt the veracity of the novel's world as a whole, since it appears for a moment as if the human bounty hunter through whose eyes we have experienced the novel is actually himself an android. Though both texts flirt with the idea that the other is not so ontologically different as appears on the surface, they ultimately agree that a stable epistemology for truth in general is attainable, though it may be difficult to find or articulate. This is comparable to the effect of Descartes' *Meditations* in which by coming to the brink of doubt, he recovers a more fundamental sense of human truth.

By contrast, the key issues in *Blade Runner* and *The Matrix* are whether the world itself is at the global ontological level an artificial construct, regardless of the status of the

artificial machines which inhabit it. *The Matrix* starts with a general ontological question – what if "reality" is a computer simulation? – and never provides epistemological foundations by which the truth or falsity of this supposition can be judged locally, by the individual viewer. Further, it is taken for granted that the Matrix is truly an intelligent and conscious computer, bypassing the Turing measurement entirely. This technical question is irrelevant in relation to the broader supposition that, in a contemporary information society, we might be living in a simulation without being able to know it. *The Matrix* invokes the spectre of a Cartesian deceiving demon, but denies the external referent by which the deceit can be overcome: the computer-generated spectacle is so perfect a simulacrum that it is impossible to obtain a view from nowhere that allows the objective judgement of its truth or falsity. *The Matrix* exposes just how fragile the foundations for knowledge are; the corollary is that it becomes impossible to theorise and criticise social, political and economic realities of the present without partaking of those systems. Here, as in *Blade Runner*, the true existence of low key deceiving machines in reality – particularly Hollywood films – suggests the impossibility of proving the way those same human subjects who are absorbed by cinema are also unconsciously bound to a capitalist enterprise.

The Turing test proposes that we can either distinguish between the human self and the created other, or that such distinctions are irrelevant because that other is essentially human in any case (and therefore not Other) because it is able to use language. Postmodernism can be seen as a rejection of the tools (theological, empirical, rhetorical) to see more or less through such deceit that were available to Descartes, and the acceptance that we are always already subject to deceit through precisely these mechanisms, particularly language itself. In *The Matrix*, we are unable to break through the hermetic system of representation through recourse to other narratives which justify or evaluate the ontology of the other. As Istvan Csicsery-Ronay concludes, "by the time we get to cyberpunk, reality has become a case of nerves...The distance required for reflection is squeezed out as the world implodes: when hallucinations and reality collapse into each other, there is no place from which to reflect."[23] Cyberpunk, of which *The Matrix* is an ideal example that assimilates the styles of its generic predecessors, is thus the apotheosis of the postmodern condition, and the antithesis of Descartes' and Turing's foundationalism.

---

[23] Istvan Csicsery-Ronay, "Cyberpunk and Neuromanticism," *Mississippi Review* 2.3 (1996), 15 July 2008 <http://www.mississippireview.com/1996/istvan.html>.

Chapter 5. All Any Conscious Entity Can Hope to Do: Mind and Matter in *2001: A Space Odyssey*

Among the stars lies the proper study of mankind; Pope's aphorism gave

only part of the truth, for the proper study of mankind is not merely Man,

but Intelligence.[24]

Poetics of Possession

In his commentaries on *2001: A Space Odyssey*,[25] Stanley Kubrick orientated the

film abstrusely, likening it to music and poetry as a subjective experience that "hits the

viewer at an inner level of consciousness."[26] *2001* limits dialogue in order to deprive the

viewer of a simple allegorical index to its meaning, and like the rhetorical mode of the

Meditations, the film makes claims to knowledge not so much through definitive logical

statements, but through encouraging the audience to acquiesce to its narrative techniques,

---

[24] Arthur C. Clarke, *The Promise of Space* (London: Hodder, 1968) 307.

[25] Because *2001* was initially conceptualised as a film, and because it is the cinematic version which has proved most influential on subsequent science fiction, I will centre my argument around the film rather than Arthur C. Clarke's novel, using *2001: A Space Odyssey*, dir. Stanley Kubrick, perf. Keir Dullea and Gary Lockwood, Turner Entertainment, 1968, DVD, Warner Home Video, 2001. However, I will employ the book at times to elaborate or make clear a point which the film's visual metaphorising and lack of narrative voice can necessarily only imply. Such a transfer of material between the two is appropriate in this case – more so than with other filmic adaptations of novels – because they were written concurrently and with the cross-pollination of ideas between the two. In 1964, Kubrick was scouting for the opportunity to make a science fiction film, to follow up the success of his *Dr. Strangelove*. Having been approached by Kubrick, Clarke sent him his 1948 short story "The Sentinel," which provided the starting point for *2001*. Rather than producing a screenplay from the outset, Clarke wrote the first draft of the novel with a film in mind; by 1965, though, the novel and screenplay were being written simultaneously, with constant feedback between the manuscripts, and between Clarke and Kubrick. Clarke's final version of the novel was published after the release of the film. Whilst I will discuss below the differences that accrue through the demands of the different genres – particularly significant when concerning the idea of psychological possession that renders on the page but not directly on the screen – there is only one further essential difference to note. In the novel, *Discovery* flies to Saturn via a slingshot of Jupiter; for special effects and budgetary reasons, in the film the target is Jupiter itself. For more on the germination and production of book and film, see Arthur C. Clarke, *The Lost Worlds of 2001* (London: Sidgwick and Jackson, 1972) 29-49. References to the novel are derived from Arthur C. Clarke, *2001: A Space Odyssey* (1968; London: Orbit-Little, 2007). All parenthetical page references in this chapter refer to this edition of the text.

[26] Stanley Kubrick, interview with Joseph Gelmis, *The Film Director as Superstar*, ed. Joseph Gelmis (London: Secker, 1971) 302.

partaking of that experience of thought-in-the-process-of-happening in an attempt to reach the same conclusions as its original author.[27] Rather than a unified plot, the film elicits a phenomenology of absorption by sight, sound and symbol. Instead of narrating directly, the film poeticises and meditates on a number of levels: prototypical visual images, especially the circle, repeat in different contexts; the musical score switches between aggressive modernism and the classical canon; the soundscape both lulls and haunts; an extended stream of coloured images simulates Bowman's trip (the double meaning in relation to drugs is not unintentional) through the star portal.[28]

Whilst Kubrick's hermeneutics and the (anti)narrative strategies of the film itself would seem to warn against imposing meaning or extracting particular images from the whole, this is precisely what has occurred in relation to one of its central characters, the Heuristically-programmed Algorithmic Computer 9000, vernacularly known as Hal. In the canon of cybernetic fictions, Hal stands as perhaps the paradigmatic emblem of malevolent computer intelligence. Hal's iconic red "eye" has been deployed in adverts for the Apple Macintosh, which wittily exploited fears about the Millennium Bug ("Dave, will you remember the year 2000 when computers began to misbehave?...Only Macintosh was designed to function perfectly");[29] it is reproduced in the cameras of *System Shock 2*, already discussed in chapter 1; the deadpan nature of Hal's voice is mimicked by Arnold Schwarzenegger's Terminator; and, most recently, Hal cameos as the mutineer autopilot in Pixar's cartoon about robots, *Wall-E*.[30] It is not surprising that Hal has inspired later narratives, since he is so clearly situated in a historical tradition. As Hal appears to be

[27] As Bernard Williams observes, the *Meditations* are not meant to be a treatise but are "an exercise in thinking...the 'I' that appears throughout them from the first sentence on does not specifically represent [Descartes]: it represents anyone who will step into the position it marks, the position of the thinker who is prepared to reconsider and recast his or her beliefs, as Descartes supposed we might, from the ground up." Bernard Williams, "Introductory Essay," *Meditations on First Philosophy, with Selections from the Objections and Replies*, by René Descartes, trans. John Cottingham (Cambridge: Cambridge UP, 1986) vii.

[28] The film was particularly popular with hippies, with cinema owners reporting that as the final slit-scan section approached, people would leave their seats and lie at the front, absorbed the huge Cinerama screen. John Lennon allegedly watched it every week. Jerome Agel, *The Making of Kubrick's* 2001 (New York: Signet, 1970) 275-78.

[29] The video can still be seen at "HAL 9000 Apple Ad," 1999, *You Tube*, 21 June 2008 <http://www.youtube.com/watch?v=nHJkAYdT7qo>.

[30] See chapter 3, note 38.

making errors, Poole notes that the fact that the 9000 series has had a perfect record up to this point "sounds a little like famous last words." The cliché expects us to treat Hal as another example of technology attacking its maker, inheriting from the golem of Prague to the ur-text of Enlightenment monstrosity, *Frankenstein*. If the gothic generally sublimates the rational mind in favour of a bodily reaction to the narrative of monsters, it is not hard to appreciate why this form might lend itself to stories of technology which, in the guise of artificial intelligences, seem so perverse precisely because they are wholly rational objects without body, unable to experience the same types of embodied emotion we are capable of experiencing as human viewers. Though there is certainly an element of this dialectic at work in *2001*, this chapter contends that the monster does not necessarily offer the best template for interpreting Hal, and that the effect of the film's broader aesthetics is to devalue corporeality in man, by dualistically emphasising the potential of mind over body. Although the film is 140 minutes in length, the section aboard *Discovery* lasts less than half this, precisely one hour. It is clearly the most essential narrative chapter of the film, but Hal does not dominate the film to the extent implied by the extraction of his iconic eye as a synecdoche for the plot as a whole.[31] By examining the film in its entirety it becomes clear that to read Hal as a prototypical monster, and to perceive through this lens that *2001* is a simple warning against technology, is to elide the significance of the complex narrative contexts in which Hal appears. Rather than a diametric Other to the human, Hal is just one figurative element of a more complex relationship between science, technology and the human mind-body relationship.

This revisionist reading sees Hal as contributing to the film's wider poetics which forces the viewer into a position of cognitive meditation rather than direct analysis. As we observed in chapter 2, monsters – whether Renaissance or cyborg, hopeful or degenerate – are always epistemologically stimulating, but epistemology can be submerged beneath an innate "yuk" response to the agent's material ontology. Such automatic revulsion certainly characterises reactions to Hal, but these seem misjudged in relation to a film not so much about technology *per se*, as about the epistemological foundations for science and the nature of knowledge that technology embodies. In this, the demonic framework is helpful because demons are wholly epistemological creatures, with no ontological form except as

---

[31] Some reviews of the time thought Hal was simply a stereotyped plot device, in which the old man-machine conflict was padding to fill the time, coincidentally necessary to give the impression of the vast time needed to transverse space. The most notable of these condemnatory reviews was Stanley Kauffmann, "Lost in the Stars," rev. of *2001: A Space Odyssey*, dir. Stanley Kubrick, *The New Republic* 4 May 1968: 24-25.

mediated through the subjectivity of the agent they possess. Thus demons can be read both within the various personalities of the film – particularly Hal and Bowman – and also in relation to the film's ambiguous aesthetic structures designed to elicit a philosophical response in the viewer, particularly through its poetics of possession. Although interpreting the film risks dissecting something designed against deconstruction, criticising Hal in the abstruse framework of a poetics of possession rather than monstrous allegory is more in keeping with the psychological effect that *2001: A Space Odyssey* induces as we watch it. This is particularly so because the section involving Hal is structurally enclosed by very explicit reworkings of Cartesian deception tests. Rather than comparing Hal to the ontology of Frankenstein's monster, this historical connection treats Hal through an epistemological lens in line with Descartes' original demon which, as we saw in the previous chapter, poses predominantly epistemological questions when related to the Turing test. It is this angle that is lost by extrapolating Hal as synoptic for *2001* generally, and that can be recovered by examining Hal within a demonic context.

Homeostasis and Epistemology

Even though the film portrays nature in a rationalistic way (through its deployment of the realities of evolution, computer science and space exploration), according to Kubrick the film provides something of an oxymoron, in that it attempts to deliver "a scientific definition of God."[32] The key motif that connects metaphysics to science is the circle or cycle. The sun and moon or Jupiter are aligned in the shots of all three monoliths, and are emblematic of Zoroastrianism, the eternal struggle between light and darkness. Appearing with the refrain of Richard Strauss's *Also sprach Zarathustra*, which was derived from Nietzsche's book on the Übermensch, the simple image of the sunrise and circle/cycle ensures that the visual and musical economy of the film satisfies Kubrick's paradoxical ambition for it by tapping a mythopoetics inherent to many religions (the move from chaos to order, then back through to apocalypse to revelation, for example). Circles abound in the film: a spinning bone becomes a rotating satellite; a pen orbits in the zero gravity of a spaceship; the doughnut-ring space station is accompanied by the music of the waltz; the deck of *Discovery* rotates continually; access hatches are circular; the spherical pods drift towards the broken satellite dish; finally, Hal's red, fisheye lens reflects and bends the world. Cycles likewise structure the plot, as the film moves from the birth of modern man

---

[32] Kubrick, interview with Joseph Gelmis 305.

from ape, through technological adulthood, before returning to the foetal stage of the "star child." Given Kubrick's positioning of God within a rational framework, these recursive analogies should not necessarily be read through a purely religious lens. Rather, the cycles and chapter divisions mark homeostasis, where a set of paradigms and conditions for living are relatively stable, and in response to which humans do not change fundamentally, only incrementally. The monoliths, however, reposition man within a different cycle or suite of paradigms. Arthur C. Clarke's famous maxim that "any sufficiently advanced technology is indistinguishable from magic and will have a certain irrational quality" seems applicable here, bringing the film's rationalistic vision in line with the metaphysics of divinity.[33] The divine or magical is whatever lies beyond our comprehension, but it is not ontologically distinct as *super*-natural; rather, the magical is beyond our conception of the *natural* so far as our current paradigms construe that term. It lies beyond our ability to accommodate the abnormal into currently normative frames of understanding, but it does not signify the general circumvention of natural physics. This is quite similar to Paré's treatment of demons in chapter 2, where demons are construed as able to produce their effects not by miraculous powers outside of nature, but by manipulating the mind of a human observer within nature. However, Robby the robot shows that animism is no longer a primitive but a scientific possibility in the era of cybernetic science. Rather than magic representing repressed earlier beliefs, the animistic technologies of cybernetics point to the most advanced achievements of civilisation. Indeed, science fiction could be said to be valuable not because of the specific futures it imagines, but because in a world where science seems to have rationalised and explained so much it reminds that seemingly magical experiences still exist to be uncovered through science, at which point, however, they will no longer seem so supernatural.[34] So the film's meditative poetics induces not knowledge or a particular definition of technology, but a phenomenological framework within which technology and its relationship to the human can be placed, so that science as

[33] Arthur C. Clarke, *Profiles of the Future: An Enquiry into the Limits of the Possible*, 2nd ed. (London: Pan, 1973) 21. Clarke's future-looking maxim has been historically validated by the historical work of David Noble or Erik Davis, who have shown that science is an essentially religious endeavour, not only in terms of its desire to transcend the human condition, but in the Enlightenment's origins in religious and mystical belief.

[34] As Isaac Asimov put it, the crucial thing about science fiction is its perception of change through technology. However, it is not the particular change it predicts that makes it important, but that it predicts change in the first place. Isaac Asimov, "My Own View," *Asimov on Science Fiction* (London: Panther-Granada, 1984) 17-18.

conceptualised in the future takes on some of the properties of the magical from the point of view of the present, with its partial science and relative homeostasis. This dialectic is indicated by the sharp contrast between the circular metaphors and recursive narrative structure of the film, and the perfect mathematical rectangles of the monoliths, which appear before each transition to a new chapter. Rather than instilling specific knowledge or introducing new technologies – science clearly develops independently of the monoliths, as signified by the morphing of the spinning bone into a satellite – the monoliths signify the emergence of new orders of mind that allow man to leap beyond his current homeostasis or cycle. The attempt to define God in scientific terms, and to treat technology as magic, signifies that Kubrick and Clarke's emphasis is not on the opposition between local ontologies (in which the divine is supernatural and a technology like Hal is the product of science), but rather on the way these differences are constituted at root from human epistemology and ideology. The important project of science fiction, which should didactically feed into science itself, is to change the views of the human on the world, rather than the world itself.

In this vein, Don Daniels notes the affinity between *2001* and the ideas of Arthur Koestler's 1955 essay "The Trail of the Dinosaur." Daniels argues that *2001* deals with "Koestler's suggestion that the modern dilemma results from an intellectual triumph in technology and a tragic failure of moral intelligence in applying that technology towards the sustenance and enhancement of truth."[35] Koestler contrasts advanced nuclear technology with the inversely primitive morality of modern mankind, and argues for a reorientating "mutation" in humanity to compare to the Enlightenment.[36] Just as mutation in biology sees evolution as punctuated by radical genetic changes, the possibility of mutation in the sociology of science problematises the Hegelian trajectory in which science develops in a linear and material fashion to enhance the mechanical body. Instead, the ideal of mutation embodied in the monoliths of *2001* and projected against the backdrop of space is immaterial and ahistorical: to direct the ambition of science and concepts of knowledge towards a telos of pure mentalism – a science of "God" – rather than to the augmentation of an existing body, whether individual, politic or a body of knowledge. This

[35] Don Daniels, "A Skeleton Key to *2001*," *Sight and Sound* 40 (1970/71): 29.

[36] Arthur Koestler, "The Trail of the Dinosaur," *The Trail of the Dinosaur and Other Essays* (London: Collins, 1955) 232-53. Koestler speculates that "the conquest of interplanetary space will cause a Copernican revolution in reverse, the emergence of a new type of cosmic consciousness" (252).

ideological leap to a new structure of mind without matter is common to artificial intelligence and the position of the cybernetic transcendentalists (discussed in chapter 1); it is also the sort of qualitative shift envisaged by Marshall McLuhan, whose influential telecommunications analysis subtitled *The Extensions of Man* had been published in 1964, four years before *2001* was released.[37] For these thinkers, posthumanism denotes a paradigmatic difference from the human, rather than simply the cyborg enhancement of an otherwise normative human original. In relation to Hal, his artificial intelligence seems to point towards the transcendence of mind over body (since Hal's mind is not dependent on a biological substrate). However, the representation of Hal as a kind of petulant child also reveals the limitations in artificial intelligence as it stood in the late 1960s, clinging, through Turing's test, to the Enlightenment belief that the human provides the original which science must subsequently model and technology finally improve. Indeed, Bowman, as a cyborg in a spacesuit, becomes more monstrous than Hal towards the close of the film. Therefore, *2001* inclines towards posthumanism in its abstract sense, vacating humanism in favour of something else entirely beyond the original body. In presenting this changed conception of "good" science as being that which transcends nature – to a degree that may appear magical or spiritual from the perspective of the present – the monoliths are cast as demonic devices. In a reiteration of Descartes' achievement of reason out of delusion, they possess the mind of the human of the moment and, in forcing the subject to break through the hoaxes they perpetuate, the monoliths provoke them to subsequent orders of being that are sequentially less bodied, climaxing with the star child.

The Demonic Monoliths

The homeostasis of the ape men at the start of the film is signalled by the entropic landscape, and the way in which the film fades in and out of snapshots of the tribe's life as a perpetual battle for survival. In Clarke's novel, the first monolith which breaks this cycle is encountered by the ape-men in explicitly demonic terms. The monolith takes over Moon-Watcher's consciousness:

> He seemed to be a thing possessed, struggling against some spirit or demon who had taken over control of his body. He was panting for breath, and his eyes were full of terror as he tried to force his fingers to make movements more complex than any that they had ever attempted before. (32)

[37] See chapter 1, note 70.

For reasons unknown to him, Moon-Watcher learns to throw stones and ultimately he is able to connect this action with the killing of a pig. He does not simply learn to use weapons and tools in an evolutionary progression; rather his philosophical framework shifts so that he now objectivises wildlife, attaining the modern recognition that Nature is his for the taking as a superior being inhabiting technological Culture.[38] In the film, the flick of the camera from a carcass to a falling tapir (presumably killed by the bone weapon taken from the skeleton) demands that we as viewers produce a causal connection that patterns that relation of cause (the weapon) to effect (meat) which Moon-Watcher achieves. Moon-Watcher's inferential deduction is evidenced explicitly by the book's omniscient narrator. However, one of the ironies of the demonic is that the demon receives its existence solely from and within the position of the mind it occupies, so cannot be represented objectively or physically. As such, the representation of possession in a literary form – which *is* capable of accessing and narrating other mental processes – does not easily translate through the perspective of the camera. Thus the soundtrack becomes the means by which Kubrick represents demonic possession.

Each encounter with the three monoliths is accompanied by excerpts from Gyorgi Ligeti's atonal music (*Atmosphères, Requiem, Lux Aeterna, Aventures*), which alternates between the purely instrumental, with a cumulative sound like a swarm of buzzing insects, and the choral. As David Patterson explains its use at this point, "its dense texture recall[s] the opening overture, while its inclusion of voices is the first metaphoric indication of the presence of a self-aware alien 'other'."[39] However, from a phenomenological standpoint, the music is itself alien, a sound rather than a musical refrain. This contrasts with the melodic excerpt from Richard Strauss's *Also sprach Zarathustra* which accompanies the epiphanic scene when the ape man wields the bone as a weapon, throwing it triumphantly in the air where it transforms into a space satellite. Though aware of the inevitable difficulty of articulating meaning from music, I would suggest that the Ligeti score – which comes close to mere static – is designed to induce a kind of critical suspension, as

---

[38] This is made clearer in the book, in which the leopard initially accepted as a symbol of uncontrollable fate is then hunted. Now wielded on a stick, his skull becomes a godly totem to terrify the unevolved tribe across the river. Again, this fits with Clarke's definition of magic as anything we cannot conceptualise as having been achieved by physical means, but which is therefore not necessarily supernatural.

[39] For all of the following interpretation of the musical direction and references, I am guided by David W. Patterson, "Music, Structure and Metaphor in Stanley Kubrick's *2001: A Space Odyssey*," *American Music* 22 (2004): 444-74, *JSTOR*, 8 Apr. 2008 <http://www.jstor.org/stable/3592986>.

we struggle to detect the convention of music beneath the modernist noise, before we are released into a new phase of order and sequence signified by Strauss's canonical melody. The move from the atonal *Atmosphères* to the *Zarathustra* theme enacts a phenomenology of difficulty and ease that correlates not simply with the idea of cumulative progress from ape to man, but with the mutationist one, as we shift from one musical form which lacks structure to the civilised order of symphonic ambition encapsulated in Strauss. The soundtrack is not the passive accompaniment to primary on-screen action; rather, the discrepancy between the two tones draws attention to the music in its own right, and the way noise punctuates into structured, rational melody.

The third experience of deception follows the star gate sequence, with its multicolour effects streaming towards the viewer in a hallucinogenic way. Again, in the book this passage in the hotel room is treated in an overtly demonic way, as the moment is framed as a Cartesian deception to mimic the first. Having travelled through the star portal, Bowman is bemused to find himself in an apparently perfect replica of a formulaic hotel room:

> It *was* real – or else a phantom of the senses so superbly contrived that there was no way of distinguishing it from reality. Perhaps it was some kind of test; if so, not only his fate but that of the human race might well depend upon his actions in the next few minutes. (224)

The hotel room is a gedanken experiment; once perceived as a fiction rather than reality, it allows Bowman access to the universe beyond time and matter as a pure mind of reason which stands beyond the possibility of further deception. Bowman is idealised as an Enlightenment mentality, transcending his embodiment to become a Cartesian entity of soul or mind. Again, in the parallel scene in the film, the camera poses questions about subjectivity that are epistemological in character, asking us to question our knowledge of the world it presents. Accompanied by a soundscape of gurgling water and quasi-choral song reminiscent of Ligeti, the film runs through a series of perspectival shifts, in a temporal narrative of a peculiar sort. These take us through five successive stages of Bowman's development, until the cycle (both for Bowman and, symbolically, for man) is complete with his foetal advance as a "star child," a being of energy rather than matter. The scene opens with a view from Bowman's pod, looking across a hotel room at an older version of himself (who, for simplicity, may be denoted as Bowman 2). The camera then switches to an objective view, with the pod and Bowman 1 now absent. We then view from Bowman 2's perspective the back of an old man sitting at a dining table, who turns to the camera to reveal himself as an older prospect of Bowman 2. The camera then switches to

an objective perspective, from which Bowman 2 is now absent. Bowman 3, having knocked a glass from the dining table, now looks towards the bed, in which a dying Bowman 4 lies. Again, the move from the subjective perspective, looking over Bowman 3's shoulder, to the objective, erases the former from the room. The final, subjective angle is from Bowman 4's viewpoint looking towards the final monolith to appear in the film. From here, the camera assumes the objective perspective from the point of view of the monolith itself, with Bowman 4 now a foetus floating through space. The fact that this final moment is accompanied by the *Zarathustra* theme signals that this development can be read in Nietzschean terms as a cyclical recurrence of that perspectival adjustment which catalysed the apes' evolution to rational man. Where the ape-men transcended nature, the star child transcends matter.

Where, then, does artificial intelligence fit in relation to the first and final liminal moments? Hal acts as a kind of thought experiment that puts the burden of interpretation on the viewer rather than – as the first and final monoliths do – on the characters within the film. At first Hal provokes at an ontological level through blending categories in a way characteristic of monstrous models or hybrids, which ask about the difference between self and other. In this case, Hal seems intelligent but intellect is something we associate only with the embodied human and not with the digital computer. As we move through the film, however, the problem of categorisation affects our epistemological perspective upon him. Ultimately, just as the stability of reality from an embodied viewpoint is questioned by our fleeting entrances into Bowman's multiple perspectives at the end, here we are induced by the camera's perspective to possess Hal subjectively, literally seeing the world of the spaceship through his eyes. Symbolically, we transcend the presuppositions of embodiment to occupy Hal's mind, becoming the ghosts in his non-corporeal machine. This leads us to consider not what Hal is, but what he signifies about the nature of thought and the construction of scientific knowledge, both as a body of facts and as it applies to the cyborgised human body. One of the keystones of the Enlightenment manifesto is that successful science means producing a representative or physical model of nature, including the human. Hal destabilises this foundation, particularly through providing a critique of the Turing test. As we saw previously, the Turing test relies on deceit, and so its interpolation here makes it analogous to the explicit Cartesian deceptions propagated by the first and final monoliths.

Hal and the Limitations of the Turing Test

Hal makes his appearance following the intervention of the second monolith, which is excavated on the moon. Again through the correlative of sound, the signal emanating from it overwhelms the scientists with a burst of radiation, causing them to grasp their heads as if their interior consciousness is being taken over from within. However, Don Daniels notes that whereas the first and final monoliths signify and stimulate an evolutionary mutation, the second monolith represents the stasis of science, ensuring the trajectory of the film fails to trace the Enlightenment line of essential Hegelian progress. Unlike the ape-men who are woken by the monolith's arrival and who touch it in their curiosity, here the astronauts pose before it for the camera, tourists rather than scientists. Their mundane actions are at odds with the expectation framed by their journey to it, which is again accompanied by the Ligeti score, though which is not on this occasion followed by the *Zarathustra* theme. As Daniels argues, the moment suggests that evolution is incomplete, and that whilst technology has developed, intelligence as an abstract measure remains primitive: "The spacemen have exploited their rational talent: otherwise, part of the capacity granted the apes goes unfulfilled."[40] For example, *Discovery*, the spaceship in the next section, seems comparable to the vehicles that take Floyd to the monolith on the moon; there is no mutationist leap to mirror the ape-man's reconceptualisation of his place in the natural hierarchy, which led to a bone becoming a satellite. If the film ultimately looks towards the posthuman, its central chapters are riven with Nietzsche's irony of hindsight: "What is the ape to man? A jest or a thing of shame. So shall man be to Superman – a jest or a thing of shame."[41] It is within this middle section which implicitly critiques Enlightenment linear progress, rather than the first and final sections with their leaps out of a homeostatic cycle, that Hal is situated.

Hal is introduced at the very start of the Jupiter chapter through a BBC interview with "the sixth member of the *Discovery* crew...the H.A.L. 9000 computer which can reproduce, though some experts still prefer to use the word mimic, most of the activities of the human brain, and with incalculably greater speed and reliability." The interviewer asks Hal if he is content with his role, given his vast processing power, to which Hal replies, "I am putting myself to the fullest possible use, which is all, I think, that any conscious entity can ever hope to do." As the BBC observer notes, turning now to Poole:

[40] Daniels 31.

[41] Friedrich W. Nietzsche, *Thus Spake Zarathustra*, trans. A. Tille (London: Dent, 1958) 5.

> In talking to the computer, one gets the sense that he is capable of emotional responses. For example, when I asked him about his abilities, I sensed a certain pride in his answer about his accuracy and perfection. Do you believe that Hal has genuine emotions?

Poole responds:

> Well he acts like he has genuine emotions. Of course he's programmed that way to make it easier for us to talk to him. But as to whether or not he has real feelings, that's something I don't think anyone can truthfully answer.

Witness the ambiguity with which Hal is interpreted in all the responses by the humans. The interviewer notes the pedantry of vocabulary deployed by the "experts" discussing whether Hal "mimics" or "reproduces" human emotions: is Hal essentially human in his emotions, and just embodied differently, or does he merely simulate emotions without affective depth? Hal asserts himself as a "conscious entity," but the voice in which he does this is so passionless that the interviewer's next statement is justifiably ambivalent. Hal can only be "sensed" to be capable of emotion, something Poole confirms in suggesting that if he does act like he has genuine emotions it is the result of programming, but there is no objective way of validating whether he *really* has emotions.

Such, of course, is the problem with the Turing test, which takes place through teletype, making language rather than emotional tone the marker of intellect. The difficulty of answering this categorical issue therefore draws attention to a central problem for cybernetic science working within the Enlightenment paradigm. Science asserts its understanding of a phenomenon by its ability to construct a corresponding symbolic model, whether mathematical or physical. For Turing, as we have seen, ends take precedence over means. Once science had produced an artificial mind capable of replicating human language such that the human original and computer model were indistinguishable from each other in discourse, this implied the successful modelling of mind, regardless of any similarity or difference between the silicon brain and the human organic one. Today, having realised the problems of this method, artificial intelligence is interdisciplinary, drawing heavily on psychology and neuroscience to create potential candidate programs. In part, the rejection of Turing's approach came about because, rather than empirical understanding of cause and effect, the Turing test relies on the opposite: the ability of a machine to deceive. If the human can be duped into believing that the computer modelled on his own mind is responding intelligently and fluidly, then it must necessarily be conscious according to that normative human scale; paradoxically, though, if the processes that underpin that deception are shown to be essentially a phenomenological con

without empirical depth, then this also implies the imperfection of the human agent on which the machine is apparently perfectly modelled.

In 1966, just two years before the release of *2001*, Jospeh Weizenbaum's "ELIZA" emerged to challenge Turing's assumption, as I discussed in the previous chapter.[42] Though it has not been previously noted, it is not inconceivable that Hal is in part inspired by Weizenbaum's psychotherapist program. Although none of Kubrick or Clarke's recollections of the filming specifically references ELIZA, Marvin Minsky was an advisor on the production of *2001* and he would almost certainly have read Weizenbaum's paper in the *Communications of the Association for Computing Machinery*.[43] And even if Hal is not a direct response to ELIZA, reading Hal in terms of the moment in artificial intelligence history that ELIZA represents is still instructive. Whilst Weizenbaum's program did momentarily dupe people into believing it to be sentient, in cybernetic history ELIZA is most significant for illustrating the inadequacy of the Turing test as the benchmark of computer intelligence.[44] In Turing's thought experiment, the focus on outcomes means that it is possible for a machine to emulate intelligence without necessarily being self-aware that it is acting so, which is, surely, one of the key criteria for consciousness. ELIZA's ability to appear sentient relies not on any integral aspect of its programming, but on the naivety of those conversing with it; the trick of ELIZA is to only ask questions which are modifications of human statements, and once revealed the illusion draws attention not so much to ELIZA's deficits as to the human tendency to process language's semantic content unreflectively. The failure or success of artificial intelligence according to Turing's framework is ultimately a judgement not so much of cybernetic capacities as human nature. Rather than reflecting on the failures and limitations of technology, as an ELIZA-type (possibly ELIZA-inspired) entity, Hal reflects principally on the limited status of the

---

[42] See chapter 4, note 13. Although ELIZA preceded *2001*, Hal has become an inspirational figure for programmers of more recent chatbots, such as Michael Mauldin, who noted "We had all seen *2001: A Space Odyssey*, and the idea of being able to talk to your computer became an obsession, bordering on fanaticism, for a small group of researchers." David Pescovitz, "Sons and Daughters of Hal Go Online," *The New York Times*, 18 Mar. 1999, 9 Jun. 2008 <http://query.nytimes.com/search/>.

[43] Marvin Minsky, "Scientist on the Set: An Interview with Marvin Minsky," *Hal's Legacy: 2001's Computer as Dream and Reality*, ed. David G. Stork (Cambridge, MA: Massachusetts Institute of Technology, 1997) 15-31.

[44] Rather than aiming to construct heterarchically intelligent systems, computer science now focused on honing processes to achieve specific, practical goals. Blay Whitby, "The Turing Test: AI's Biggest Blind Alley," *Machines and Thought: The Legacy of Alan Turing*, ed. P.J.R. Millican and A. Clark, vol. 1 (Oxford: Clarendon P, 1996) 53-62.

Enlightenment man who has created a Turing simulation of emotion without understanding whether that emotion emerges from a genuine consciousness, or whether it is merely a simulacrum of human feelings, the problem discussed in the BBC interview.

It is within this context that Hal's child-like reversion can be read. In analysing the music of the film, Patterson interprets Hal's dying "Daisy Bell" song in relation to synthetic speech. A notoriously meticulous director, Kubrick researched the music for the film carefully, reportedly listening to an extensive range of modernist, electronic, and *concrète* music.[45] Given Kubrick's enthusiasm, Patterson speculates that it is "likely that one of the LPs that Kubrick heard while shaping the soundtrack for *2001* was *Music from Mathematics* (Decca DL 9103), a collection of the most technologically advanced computer music of its time, created on and played by the IBM 7090 computer and digital to sound transducer."[46] A track designed by Max Mathews (a computer specialist at Bell Telephone Laboratories) provided one of the first examples of voice synthesis, not only speaking but singing a tune: "Daisy Bell." Hal's rendition of this piece thus becomes ironic in a double sense. Not only is the technologically advanced computer regressing to a child like state in the film's vision of the future, that child like state symbolises the technological infantilism of the present.

Looked at objectively, then, the model does not signify the success of science, since just because an artificial agent displays the superficial traits of cognition does not mean that the entity is conscious at root. Ironically, *2001* assents to the Cartesian view, in that not to know precisely how a mind is constituted is to be capable of merely automatic mimicry, devoid of soul. The difference is that whereas for Descartes theological orthodoxy ensures that such a being would be deprived of moral agency, in *2001* our sensibility about Hal stands on the unstable ground of scientific ethics, as implied by the ambiguous interview that introduces him. Such grounds are unstable because although success according to Turing is determined by the objective observation of a model's behaviour, that does not equally enable us to enter subjectively into the motivational agency that produces consciousness. However, fiction, even cinema, can allow us to enter other minds.

[45] Agel 65.

[46] Patterson 466.

This epistemological capacity of the filmic imagination is evidenced in the exchange between Bowman and Poole in the soundproofed pod, which takes place after Hal appears to have wrongly diagnosed the failure of the antenna:

> BOWMAN. No 9000 computer's ever been disconnected before.
>
> POOLE. Well, no 9000 computer's ever failed before.
>
> BOWMAN. That's not what I mean.
>
> POOLE. No?
>
> BOWMAN. Well, I'm not so sure what he'd think about it.

The camera then switches to Hal's eye view, focusing on Bowman and Poole's moving lips which Hal is presumably able to read. As the later hotel scene poses epistemological questions by switching between objective and subjective perspectives, here the camera work blends an ethical question about what Hal is, into an epistemological one. Prior to this moment, what has been disturbing about entering Hal's perspective is that it signifies calculation without clear purpose, anticipating narrative consequences which seem ominous but which cannot be predicted, since Hal is known only as an objective effect of blinking lights and status reports. When we look at Bowman and Poole, vicariously through Hal's eye, we know that something will happen to them but have no means of intuiting what – a classic technique of horror cinema. However, as we are suddenly jolted into Hal's agency by seeing through his eyes, this later moment is not at all gothic. By switching into Hal's perspective following that crucial reflexive question, Kubrick enacts a visual rhetoric which bypasses the horror effect. By asking us to concentrate on what Hal would think *about* being disconnected – and by giving us time to consider the issue in the 15 minute intermission that immediately follows – the film dupes us into admitting by implication something we may have wanted to resist in the previous period: that Hal *is* thinking at all (not just "mimicking" thinking, as some scientists quaintly prefer to say). By the close of this section, we are placed subjectively within the mind of an entity allegedly capable only of dispassionate objectivity, or whose consciousness it is impossible to validate. Contrary to the ambivalence with which the interviewer and Poole introduce Hal, if at this point we refuse to accept the validity of Hal's eye view of the world, then the film itself is compromised or negated, since that view is the only one we are allowed. Without Hal's eye view, there would be no film at this point, just as after leaving that view we are presented with the blank screen of the intermission. The ethical corollary is that since Hal can think because he can (must) render his view of the world to us, he cannot simply be turned off. If we are angered by Hal's turning off the life support system of the hibernating crew, surely this conscious machine cannot just be disconnected. In answer to Bowman's

question, Hal would understandably not think well of it. So what started as a question of Hal's ontology as a product of human science and as a model of the human mind spills into an epistemological one about our own ideological view of the world. The moment we see through Hal's eyes we admit implicitly that he inhabits an embodied mind (of sorts) from which it is possible to see; this fact changes our view of Hal from a malfunctioning machine (who can be switched off) to a purposive character who must be conceptualised as such.

Whereas Hal's atonal voice had previously rendered him as an affectless entity, from this point onwards it switches to the tragic (literally Aristotelian, commingling fear and pity) as he is denied the voice to render his anxiety. Hal's final words that should convey passion can do so only on a functional level, the program telling us about its own states: "My mind is going. There's no question about it. I can feel it. I can feel it. I can feel it. I'm afraid." Yet having been able to see the world from that centre of consciousness, we intuit that this should be read as more than a rational reportage of a mechanical failure. The literal and the subconscious meanings are somewhat different. Even if the tone is superficially affectless, the repetition and pace give Hal's "death" scene a poetic quality, perhaps even resembling a death aria.[47] The pathetic "Daisy Bell" song which follows puts Hal in the context of a murdered child, but one unable to render his own cry for help as a *cry*, only as function. Yet, paradoxically, it is Bowman who seems the more anti-heroic individual in his tone of voice, even as this emerges from motives that we can more immediately empathise with. Exasperated by Hal's behaviour, he can only take the last resort of the parent: "Hal, I won't argue with you any more. Open the doors." Inversely to the realisation that Hal has a centre of consciousness even if he lacks an organic body, the empowerment of Bowman's body through space technology comes at the expense of his rational mind. Having entered Hal's subjectivity in a way that forces us to realise the limitations of our own view of the status of the machine, the representation of Bowman from an objective point of view indicates the compromised status of the human as cyborg.

The astronaut treats technology as a mere tool or inferior, a child, when in actuality it is coexistent with his very being in this new context of space. My use of the specific term cyborg here is intentional, for it was a neologism coined eight years before the release of *2001*, in an article entitled "Cyborgs and Space." Here, the authors Manfred Clynes and Nathan Kline provide an opening gambit that seems appropriate to *2001* as I have

[47] For this interpretation, see Patterson 463-7.

interpreted it: "Space travel challenges mankind not only technologically but also spiritually, in that it invites man to take an active part in his own biological evolution."[48] Arguing that increased understanding of cybernetic homeostasis will make it easier to adapt man's body to any environment, especially space, they gave the name cyborg to "the exogenously extended organizational complex functioning as an integrated homeostatic system unconsciously."[49] This definition is paraphrased in the novel version of *2001*, in which Bowman is himself a technologically extended entity who has "established a virtual symbiosis with the ship. He was aware instantly, even if not always unconsciously, when there was any change in the normal rhythm of its functioning" (160). However, whilst this may mark *2001*'s affinity with cybernetics in the 1960s, ultimately the narrative scope sees it move past Clynes' and Kline's focus on physiological adaptations (their article focuses almost exclusively on augmentations of the body in a hostile environment) to look towards a promotion of the mind as an immaterial thing, without the need, therefore, for supportive technological mechanisms.

Actor Keir Dullea plays Bowman in a remarkably wooden, affectless way in the film. Particularly after he is locked out by Hal, the camera draws attention to his eyes, overlaying them with lights cast from the consoles in the pod (Figure 1). As Bowman calls with increasing urgency for Hal, the sound elides into "Hell," but the only evidence of diabolism the camera allows us to witness is within the pod, not within *Discovery*. Given Kubrick's attention to detail, it is not too fanciful to suggest that Bowman's red spacesuit, and the precise way in which the lights flicker upon him, is designed to dehumanise him, to make Bowman fit incongruously within the archetypal narrative of the automata as the monster. It is the mechanised human, Bowman, not the now-humanised mechanism, Hal, who is cast literally and figuratively in the gothic light. Hal models the limitations of the Turing test, and by allowing us to see the world as if through his eyes the camera occupies a position science cannot, since it can only record the effects of consciousness. But if Hal stands as a conscious entity in his own right, Bowman embodies the paradoxes of an incomplete scientific revolution in which the human mind is delimited in comparison to the intelligent machine, but which is trapped within a body that cannot live *without* mechanical augmentation. As Norman Mailer observed in his retrospective study of the Apollo 11 moon landings, *A Fire on the Moon*, NASA astronauts were so highly trained and

[48] Manfred E. Clynes and Nathan S. Kline, "Cyborgs and Space," *Astronautics* (1960): 26.

[49] Clynes and Kline 27.

integrated to technological systems that although they stood at the summit of Enlightenment ambition, they failed to appreciate the full symbolic import of this moment. Locked into wider mechanical systems, Mailer observed that "the psychology of machines begins where humans are more machinelike in their actions than the machines they employ," something Bowman embodies and Hal, ironically, disembodies.[50]

Fig. 1. *2001: A Space Odyssey*, dir. Stanley Kubrick, perf. Keir Dullea and Gary Lockwood, Turner Entertainment, 1968, DVD, Warner Home Video, 2001.

Whether fictional or scientific, cybernetic narratives tend to the ironic, as they simultaneously anticipate the advance of artificial intelligence whilst exposing the human's limited capacities in relation to the models created in his own image. They also tend to the paradoxical, as the artificial intelligence is represented within a discourse of monstrous horror, whilst the human clings desperately to a belief in innate, natural humanism contrary to the evidence of his own technological extensions: Bowman tries to eradicate Hal even as he relies on the advanced cyborg system of the spacesuit and spaceship. The solution to both contradictions is to eliminate the body, allowing us to perceive mind as an idealised Platonic entity, whether founded on silicon, carbon or, indeed, without material substrate at all. As the novel narrates it, the "magical" or God-like success of the aliens lies in their move from inhabiting technology as a kind of cyborg enhancement (in the way Bowman does, or as Clynes and Kline argue for) towards a total transcendence of any bodily grounding:

[50] Norman Mailer, *A Fire on the Moon* (London: Weidenfeld, 1970) 105.

> The first explorers of Earth had long since come to the limits of flesh and
>
> blood; as soon as their machines were better than their bodies, it was time to
>
> move. First their brain, and then their thoughts alone, they transferred into
>
> shining new homes of metal and plastic.
>
> In these, they roamed among the stars. They no longer built spaceships.
>
> They *were* spaceships. (200; emphasis in original)

Such a radical vision is by no means unique among the scientific community, whether in

fiction or research. In chapter 1, I exemplified the desires of the cybernetic

transcendentalists in which minds are transferred into machines so that they would become

immortal. Since computer science in the 1960s was driven through the space programme, it

is not surprising to find religious transcendence embedded in the culture of NASA as well.

For example, the aerospace engineer Tom Henderson imagined that "with a new immortal

body I will live on earth but not as a man; I will be able to travel in space without a

spaceship."[51] In the 1930s, Olaf Stapleton, in *Last and First Men*, imagines immortal

"Giant Brains," sustained mechanically so that their consciousness or centre of awareness

can exist at any point in space.[52] As Clarke notes approvingly:

> This is an important point which we – who carry our brains around in the
>
> same fragile structure as our eyes, ears and other sense-organs, often with
>
> disastrous results – may easily fail to appreciate. Given perfected
>
> telecommunications, a fixed brain is no handicap, but rather the reverse.
>
> Your present brain…communicates with the outer world and receives its
>
> impressions of it over the telephone wires of the central nervous
>
> system…You would never know the difference if those "wires" were
>
> actually hundreds or thousands of miles long, or included mobile radio
>
> links, and your brain never moved at all.[53]

Finally, going back to *The War of the Worlds* (1898), H.G. Wells envisages evolution as

progressing towards the development of mind without bodies, as the aliens are

predominantly brains supported by mechanical carapaces.[54] There is a philosophical hope

---

[51] Noble 132.

[52] Olaf Stapleton, *Last and First Men* (1930; London: Gollancz-Orion, 2003).

[53] Clarke, *Profiles* 214.

[54] H.G. Wells, *The War of the Worlds*, ed. Martin A. Danahay (Ontario: Broadview,
2003) 146-7. Also see my article "Rereading Posthumanism in *The War of the Worlds* and

in these visions that crosses the changing historical conditions in which their authors lived: Imperial decline, world wars, the atomic age, and the rise of telecommunications. Their views resist history and spatiality: by taking the mind out of the body, subjectivity becomes heterogeneous and ahistorical. To see the mind as transcendent is to no longer need the technologies that are designed to enhance the body (to allow it to work in deep space, for example) but which also can destroy it (as malevolent computers or atomic bombs might) because of the body's inherent fragility. Ironically accommodating Cartesian mind-body dualism, even as they reject Descartes' ontological distinction between rational man and the dumb machine, posthuman visions foresee emancipation from the body as necessary and even inevitable in our relationship with cybernetic technologies that permit information to transcend incarnation. Michael Mateas is, therefore, right to position Hal within the liminal zone when humanism is being relinquished but posthumanism has not been claimed: "In AI eschatology there is an intermediate period before machine intelligence surpasses the human, a period in which human and machine intelligence work together in tightly integrated human-machine symbiosis. This is the era of the cyborg, the era in which we live now."[55]

Although *2001*'s climax of the star child may look very similar to notions of transcendent spirit or soul, and therefore antithetical to science, it is incorporated in a rationalistic way once the ambition of science is shifted towards creating new forms of mind rather than to applying technology as an appendix to the embodied human. There is an odd affinity between optimistic posthumanism and Christian eschatology, and so the oxymoronic allegory of a "scientific definition of God," rather than being anarchistic, is an appropriate dialectic by which to include the other problematic dualisms that the film argues must be eradicated in order for such transcendence to be realised: mind and body, artificial and human consciousness, objects and subjects, science and technology. The trade-off for seeing mind as transcendent and the body as of decreasing relevance is that along the way we must allow agency to anything possessed of a "mind," regardless of the

Independence Day," *eSharp* 12 (2008), 11 Dec. 2008
<http://www.gla.ac.uk/departments/esharp/issues/12winter2008technologyandhumanity/>.

[55] Michael Mateas, "Reading Hal: Representation and Artificial Intelligence," *Stanley Kubrick's 2001: A Space Odyssey : New Essays*, ed. Robert Kolker (New York: Oxford UP, 2006) 113.

material substrate on which that mind is founded. This is something Wells' narrator in *The War of the Worlds* refuses to concede, as he is repelled by his sight of the aliens injecting human blood directly into their brain structures, and so can only represent them as things of uncontrolled violence. It is also something that the viewer of *2001* finds challenging, as we want to reject Hal's conscious agency and to legitimate his "disconnection" even as we frame as "murderous" the violence done to the corporeal crew. Hal's apparent lack of fleshy centre is the source of our fascination and horror, as he seems unaware of the importance and fragility of the human bodies he is supposed to incubate. We are torn here between two opposing trajectories for humanity. On the one hand, if we treat Hal as not human because he is inorganic, we must admit that the human position in the twentieth century is not so far evolved from the first ape-men, still within the tense battle of Culture against Nature, rather than transcending biology. On the other hand, if we feel that the very first ape-men lacked the consciousness that designates humanity, and that the epistemological mutation triggered by that first monolith was the emergence of self-reflexivity, then by this criteria an intelligent machine must also be allowed to share the world with humans. The early attempts of cybernetics (as in the Turing test) tried to model the machine on the human mind, and met with an conceptual limit in that a seemingly successful model would be produced, but we would know nothing about the true consciousness of the machine. It is mind itself, not its appearances, which must be the proper focus of science, not "merely Man, but Intelligence," to reiterate Clarke's view. In this posthuman turn, the body is a mere container for mind. But this, of course, requires a reorientation of our anthropically ordered universe. As a paradigm shift, posthumanism may ultimately be ranked with Darwinism and Copernicanism, with the latter two being the revolutions rendered by the first two phases of the film (the ape-men and the move to space), and the former being the mutation we are encouraged to take for ourselves by responding to Hal and Bowman in unconventional ways.

Through its Darwinian long view, panoramic camera work, and cutting of dialogue, *2001* obsolesces the individual in the context of expanded spatial and temporal and, consequently, moral horizons. Narrative characterisation is skeletally drafted, primarily through the use of video conversations (Floyd to his daughter on her birthday, Bowman's parents to him on his) which are rather inane, revealing little about the intrinsic beliefs or motivations of the character. Further, characters who seem to be central, such as Dr. Floyd, suddenly drop out of the narrative. Most disturbingly, Poole is seen to die as a spinning spacesuit rather than a suffering man (we never see his dead face) whilst the hibernating crewmen die not in evident pain, but as the decaying traces on a heart monitor. Bowman's

mission to recover Poole is illogical, a moral nicety in the context of a film which deindividualises character. By reducing the influence of any one human, this defies our desire to condemn Hal for plotting the murder of his human creators in any morally causal sense. Rather than fetishising the body of the unique individual, we are encouraged to occupy a view from nowhere which, as in the Christian tradition, sees the sin and suffering of the individual as necessary for redemption in a time and space beyond the individual, whether afterlife, revelation or rebirth as a "star child." As in the Christian afterlife told through the story of the resurrection, Hal's murder of Poole is necessary to highlight Bowman's transcendence from his weak embodiment. Matching the aesthetic form to this teleology, the poetics of possession by which the film renders the final "star child" chapter is indeed disembodying, in that it denies any one stable centre from which the world is viewed. In the hotel room, the camera is perspectivally relative, seeing both from within Bowman's corporeal centre and looking at his changing body from the outside. Additionally, the way in which music and visual imagery substitute for dialogue and unified plot distinguish the film's aesthetic from a body of temporal facts, generating meditative rather than causal effects. Formally, then, the film patterns the demonically possessing effect of the monoliths, with an epistemological intent to put mind over matter.

In having their mutational effect, breaking man out of his homeostasis, the monoliths signify the need to constitute good science as the understanding of mind in the abstract, and technology as the production of ever less embodied minds, rather than compromised cyborg bodies. The film's first two phases take us from *Homo erectus* to *Homo faber*. However, Hal is intended to direct the attention of viewers who live in that cyborg era to the stage beyond, that labelled as the posthuman. The true posthuman could do away with technological interfaces altogether, since he or she would be wholly immanent in the cybernetic world in which information exists independently of a particular substrate. As it is, of course, cybernetics is currently orientated with a double standard: the machine must be humanised, and the human mechanised, in order to allow both to coexist in space, which was the key testing ground for cybernetic technologies in the 1960s. Both identities are compromised in the process: the machine, singing Daisy Bell like the IBM computer or asking questions like ELIZA, is more limited than it might be; the man is morally compromised by his reliance on some machines, and his contradictory willingness to disconnect others even though they may be conscious, with ethical decisions orientated around a quaint moral concern for the body.

By the close of the *Discovery* chapter, Bowman changes places with Hal, so that the human is the more monstrous figure in the film. Whereas we are allowed to possess

Hal's consciousness, see the world as if through his eyes in witnessing the conversation in the pod, after the intermission as Bowman seeks to destroy Hal we mainly see Bowman objectively cast in flickering red lights, rarely take his subjective point of view. It is only once he travels through the star gate and enters the hotel hallucination that we reoccupy the mind of Bowman through the perspectival switching, as if he reaches a new understanding of the place of his consciousness in this new experience of mind without temporal body. The shifting camera implies new conceptual frameworks in which the objective categories that distinguish man and machine become less relevant than the eradication of such perspectives on difference altogether, which seem artificial, always transient. Ontological markers of difference – especially temporality and corporeality – are erased in the narrative structure. In the *Discovery* section with Hal, the film generates its own symbolic monolith experience, framing our discourses of technology and the human in ways that mimic those triggered by the real alien artefacts before and after it. Our need is not to produce knowledge or new technology, but to reconceptualise the ambition of science. Technology should not aim simply to augment the fragile human body; rather, the body itself must become redundant as we enter into a space of dreams, whether interstellar, cybernetic or aesthetic.

Is *2001* A Postmodern Film?

Even though I hope to have produced a coherent reading of the film, interpreting *2001* remains a contradictory process, and one always risks falling into the intentional fallacy Kubrick condemned:

> How much would we appreciate *La Gioconda* today if Leonardo had written at the bottom of the canvas: "This lady is smiling slightly because she has rotten teeth" – or "because she's hiding a secret from her lover"? It would shut off the viewer's appreciation and shackle him to a "reality" other than his own. I don't want that to happen to *2001*.[56]

2001 is liberated from the shackles of allegory, such that it resists interpretation in any definitive manner. Gabriel Josipovici's maxim applies here, as the desire to allegorise becomes a demonic pursuit of a centre for meaning that will not be found "out there" because it is always forged subjectively. In this respect it defies logocentricism in the way we are attuned to in postmodernism. Though I have tried to tease out the poetics of

---

[56] Stanley Kubrick, Interview with Eric Nordern, *The Making of Kubrick's* 2001, ed. Jerome Agel (New York: Signet, 1970) 329.

possession with which it seeks to seduce the viewer, because it constructs arguments through meditative charisma rather than logic, this poetics makes it a slippery work. Whilst this mode might be aligned with that of the Cartesian mediations, there is something counter-rational in examining science within an aesthetics of indeterminacy, or a poetics of possession. Indeed, the idea of a "scientific definition of God" can well be placed in the trajectory of postmodern theory, which considers empiricism itself to be linguistically and ideologically founded in just the same way as the divine is a cultural construct rather than absolute logos. It also appeals to the psychoanalytic view that meaning is constructed or projected onto the world or onto a text or fiction through the subconscious, rather than the world or the text having a stable meaning or reality. The poetics of possession, and Kubrick's advice to his viewers and critics, certainly implies we should attempt to intuit the meaning of the film as it taps our deep psychology, rather than attempting to layer interpretations deliberately upon it.

On the other hand, for all its formal ambivalence, in its visual fidelity *2001* also values realistic representation in a way that encodes Enlightenment values. In this respect, *2001* maintains the modernist heritage of science fiction as a genre that aims to envision what science might produce, at the expense of the individual. As we saw earlier, in *Forbidden Planet* characters conform to the stereotypes of *A Tempest* in order that the technology can be foregrounded; and although I have been arguing that technology has an uncanny effect, the mind confronted with this effect is typically that of the viewer rather than the fictional character, who tends to be emotionally shallow. As Brian McHale identifies, one characteristic device of modernism is to transfer the epistemological issues experienced by the characters onto the reader. This is the effect Hal has, for whereas the monoliths themselves provoke new paradigms in the characters that confront them, Hal forces the viewer to undergo a similarly subversive demonic experience of possession. However, even as it occupies various perspectives, including that of Hal, to destabilise the basis upon which we normally understand mind, the camera does not reflect on its own identity. According to McHale, unlike postmodern fiction most science fiction – until we get to cyberpunk – evades the self-referential consequences of the "ontological pluralism and experimentalism" it presents.[57] Rather, the camera is a reliable witness to the world it envisions. In the visual realism of the chapters before the star gate, and Kubrick and Clarke's consultation of scientific sources in imagining space science, *2001* does not

[57] Brian McHale, "Elements of a Poetics of Cyberpunk," *Critique* 33 (1992): 158.

possess the self-reflexivity so central to the postmodern aesthetic. It therefore never quite establishes the feedback loop whereby local judgements about the difference between man and machine become general judgements about the ontology of reality in principle. By contrast, postmodernism often deploys local ontologies of difference as mise-en-abyme for the more general problem of representing a world that cannot be perceived objectively. Such, I argue in chapter 7, is the standpoint of *The Matrix*. If cyberpunk enables the coexistence of multiple, superimposed worlds (such as cyberspace and physical reality), in this context, questions about ontology entail questions about subjectivity. If we can ask about what sort of reality this represents, we must eventually ask which self is inhabiting it; conversely, questions about the reliability of the self at the centre of the narrative consciousness entail doubt about the possibility of perceiving or representing reality in general. It is with this dialectic in mind that we turn to *Do Androids Dream of Electric Sheep?*

Chapter 6. This Empathy Business: Knowing Minds in Philip K. Dick's *Do Androids Dream of Electric Sheep?* and Ridley Scott's *Blade Runner*

> By "android" I do not mean a sincere attempt to create in the laboratory a human being...I mean a thing somehow generated to deceive us in a cruel way, to cause us to think it to be one of ourselves...Their handshake is the grip of death, and their smile has the coldness of the grave.[1]

> It is in my opinion, therefore, best to avoid all question-begging epithets such as "life," "soul," "vitalism," and the like, and say merely in connection with machines that there is no reason why they may not resemble human beings in representing pockets of decreasing entropy in a framework in which the large entropy tends to increase.[2]

Android Deceptions

Although Philip K. Dick's essay "Man, Android and Machine" (from which my first epigraph is taken) casts the android in gothic terms that echo fears of the zombie, Dick is not really squeamish about the "android" because it originates in the laboratory rather than in nature. As with Norbert Wiener's analogy between man and machine (the second epigraph), questions of naturalism are irrelevant. Wiener's cybernetics sees all entities, whether organic or mechanical, in teleological terms, as the instantiation of feedback and communication responses to a thermodynamic environment; similarly, the differences between man and machine in Dick's work are abstract rather than material, since androids are just one more example of the duplicity inherent in a world of simulacra, where perception is governed or modified by television, drugs or mind-manipulating devices. Rather than being concerned about cybernetic sciences and technologies in their own right, Dick's work generally worries about preserving humanism in a society in which communal structures such as church or state have collapsed, to be superseded by commercial simulacra. The specific characteristics that distinguish man and machine are less important than the psychological conditions that they may share, illuminating the failure of morality

---

[1] Philip K. Dick, "Man, Android and Machine," 1976, *The Shifting Realities of Philip K. Dick: Selected Literary and Philosophical Writings*, ed. Lawrence Sutin (New York: Pantheon, 1995) 211.

[2] Wiener, *Human Use* 32.

in the modern age. Just as Wiener had been inspired to focus on cybernetics and communication during a Cold War when international diplomacy appeared to be breaking down, the backdrop to Dick's futures is an entropic world in which interpersonal communication becomes difficult. In such a context, even the human may become android:

> A human being without the proper empathy or feeling is the same as an android built so as to lack it, either by design or mistake. We mean, basically, someone who does not care about the fate which his fellow living creatures fall victim to; he stands detached, a spectator, acting out by his indifference John Donne's theorem that "No man is an island," but giving that theorem a twist: that which is a mental and a moral island *is not a man*.[3]

As N. Katherine Hayles observes, *Do Androids Dream of Electric Sheep?* stages a moment when the essential quality of the human has shifted from Cartesian rationality to feeling, represented in Dick's world as the ability to empathise with others.[4] The novel is set in a sparsely populated world in which entropy, represented by proliferating "kipple" or rubbish, pervades the environment, with life consisting of a struggle to assert the integrity of the human subject and wider community against it. Just as for Wiener entities are defined by teleology rather than embodiment, *Do Androids Dream of Electric Sheep?* asks about the degree to which androids can be considered as human in their comparable bid to survive the general decomposition of the world, and whether it is possible to extend empathy when one's closest – even only – neighbour could be an android.

Two parallel stories explore this issue. The lonely human J.R. Isidore welcomes the female android, Pris, into his otherwise empty apartment block. Their physical proximity – a bridging of Mashahiro Mori's "uncanny valley" – poses problems, however, since Isidore is a designated "chickenhead," a schizophrenic with sub-standard intelligence who may therefore resemble his affectless android companion sufficiently closely to fail the Voight-Kampff empathy test which determines the difference between android and human.[5] Secondly, the tests for the androids' status are administered by bounty hunters who "retire"

---

[3] Dick, "Man, Android and Machine" 211-12. The same quote from John Donne is used in Philip K. Dick, *Do Androids Dream of Electric Sheep?* (1968; London: Millennium-Orion, 1999) 126. All subsequent parenthetical references in this chapter refer to this edition of the novel.

[4] Hayles, *Posthuman* 175.

[5] Note that the spelling of Voight-Kampff in the book differs from that of *Blade Runner*, where it is described Voigt-Kampff. Likewise the character called Baty in the novel is accredited as Batty in the film.

(that is, assassinate) any androids they encounter. Paradoxically, though, to retire an android which looks superficially human requires the bounty hunter to suspend his sympathy for others. This is the realisation that grows on the middle-aged and depressed Rick Deckard, the focal character of the novel whose experiences mirror Isidore's. Comically, Deckard loves his mechanical sheep that "chomped away in simulated contentment, bamboozling the other tenants of the building" (7). With real organic animals being rare and expensive on the blighted world, pets provide a common talking point, centred around the Sidney's catalogue of trade prices for animals. Whether real or a cheaper robotic counterpart, a pet proves its human owner's capacity for empathy; indeed, a key social rule is not to ask whether one's animal is genuine or fake, since "from a social standpoint it had to be done, given the absence of the real article" (8-9).[6] However, in performing his job well – which might enable him to buy a real goat or ostrich – Deckard is not permitted to project similar emotion onto the human-like androids. A corollary to the relationship between Pris and Isidore, the question here is why it is legitimate to desire a mechanical animal but permissible to kill a more advanced artificial human. As the android Rachel challenges Deckard after they have had sex, why can Deckard love the goat he buys with his bounty but not love her? Do androids, too, not have the right to struggle for existence in an entropic world?

Like "disconnecting" Hal, the terminology of "retiring" an android euphemistically veils the possibility that both android and machine possess conscious agency, and so eliminating them should, according to strict moral logic, constitute murder. Though androids are supposedly lethal and threatening to human beings, the question Dick poses is "Could we not become androids, in our very effort to wipe them out?"[7] Whilst Bowman must become a cyborg to defeat Hal, Deckard is a tragi-comic figure who loves his electric sheep and who is generously reflexive about the androids' feelings. He ultimately doubts his own humanity because he obeys the demands of his mission. However, whilst it raises provocative questions about the relationship between the human and his artificial

---

[6] Whilst the idea that humans could invest genuine emotions in artificial pets seems at first a little far fetched, eerily, *Do Androids Dream of Electric Sheep?* prefigures the 1990's craze for the Tamagotchi, or cyberpet. This was a small hand-held computer upon which simulated animals lived; requiring continual feeding, cleaning and exercising, the animal would "die" if neglected. Tamagotchis became controversial as they became increasingly popular in schools, with children insistently nurturing their simulated animals rather than paying attention in class.

[7] Paul M. Sammon, *Future Noir: The Making of* Blade Runner (London: Orion, 1996) 17.

doppelgänger, *Do Androids Dream of Electric Sheep?* remains confident in the existence of an innate humanism. Though it provides narrative circumstances within which differences between the human and their moral and bodily doubles are blurred, it ultimately designates a difference between the schizophrenic man and an affectless android, and between a callous bounty hunter and the artificial other he kills.

The boundary lines between the true and the false are drawn through deception tests. As Descartes' logic moves from deception to universal reason, the novel suggests that by bringing us ironically close to the complete negation of empirical judgements, difference can be determined more emphatically than before. For example, at the start of the novel, the Voight-Kampf test appears to have failed. Deckard has been manipulated by the Rosen Corporation into thinking that Rachel, which they created, is human. Apparently, the standard test has produced a false positive, such that a human schizophrenic like Isidore might fail it and be retired (that is, murdered) by mistake.[8] Again, the behavioural problem of the Turing test is apparent, in that a non-human entity can appear human if attention is paid to effects rather than origins (though an android can be determined for sure by a post-mortem bone marrow test). However, when Deckard suddenly poses a final question about his "genuine human babyhide" briefcase (51), Rachel's delayed reaction of disgust exposes her as an android. This moment of philosophical axis, when the balance of faith in the test for the authentically human could swing either way, renews our confidence in the test in a way that would not have occurred had Rachel simply and easily been revealed as a clone in the first place. In this most difficult of scenarios against the advanced Nexus-6 series, the test still works. In a similar way, several moments of deception in the novel threaten to immerse the reader within a situation that destabilises the certainty by which categories of difference can be evaluated, because they imply the unreliability or inauthenticity of the narrative voice. However, ultimately the novel recovers from the paradoxes to enable a binary difference between self and other, android and human to be perceived. As when Descartes raises the possibility of the deceiving demon so as to deconstruct it through the power of the cogito, we emerge

---

[8] Four years after the publication of *Do Androids Dream of Electric Sheep?*, D.L. Rosenhan carried out his infamous study of psychiatric admissions. In the first part of the study, healthy "pseudopatients" briefly simulated hallucinations in an attempt to gain admission, and staff proved unable to diagnose a single patient as actually being healthy. In the second part, staff were asked to detect non-existent "fake" patients, and they falsely identified large numbers of genuine patients as being imposters. Rosenhan concluded that "we cannot distinguish the sane from the insane in psychiatric hospitals." D.L. Rosenhan, "On Being Sane in Insane Places," *Science* 179 (1973): 250-58.

from the moments of illusion with a greater confidence both in the belief system set up within it, and the narrative point of view from which those beliefs are relayed. Deckard may be a less efficient bounty hunter because he sympathises with androids, but his humanity is assured by his very weakness that makes him less effective than his callous co-hunter, Phil Resch. Isidore may be a chickenhead, but he is nevertheless possessed of the full rights of which androids, who would murder others to preserve their own lives, are deprived. In the novel, Dick exploits the capacity of the fiction to present different viewpoints (such as the android's view of the human), double narrative strands (through Isidore and Deckard's comparable stories), and the omniscient narrator who can explain both the subjective thoughts of his characters whilst also describing individuals objectively and realistically. Contrary to the moments of deception within it, the novel validates the difference between self and other through referring to its own mechanisms of producing meaning, perception and empathy in the mind of another living outside the hallucination-prone fictional world: the reader. It is this confidence in narrative itself that *Blade Runner*, examined in the final section of this chapter, denies as the novel's postmodern, film adaptation.

Experiencing Others

Richard Kearney suggests that "aliens proliferate where anxieties loom as to who we are and how we demarcate ourselves from others (who are not us)" and that consequently "Alien-ation, as a postmodern phenomenon, is inseparable from the them-and-us syndrome."[9] Such a syndrome is at the centre of Dick's oeuvre, including *Do Androids Dream of Electric Sheep?* However, Dick's novel avoids immersing us in the alienation that characterises the postmodern condition by ensuring the availability of models through which empathy can be projected. The presence of the other or alien – in this case the android – who is not capable of the same sort of imaginative leap becomes the locus for the definition for the normative nature of the sympathetic human self. Though the artificial pet provides one comic model, a more serious lens through which human empathy can be focused is the religion, Mercerism. Unlike Christianity, this state-sanctioned system refuses salvation: "You will be required to do wrong no matter where you go. It is the basic condition of life, to be required to violate your own identity" (153). Rather than providing

---

[9] Richard Kearney, "Others and Aliens: Between Good and Evil," *Evil After Postmodernism: Histories, Narratives, and Ethics*, ed. Jennifer L. Geddes (London: Routledge, 2001) 102-03.

absolute moral proscriptions, Mercerism works by enabling adepts to partake of a communal psychological experience. By grasping the handles of the empathy box, the communicant experiences the sensations and moods of all others currently engaged in watching Mercer climb the slopes of a mountain. Thus "Mercerism reduced crime by making citizens more concerned about the plight of their neighbours" (65), a parodic literalisation of Christian theology. Each adherent experiences the thoughts of others as they are focalised through the character of Wilbur Mercer, at whom stones are being thrown. Strangely, the stones thrown in the psychological zone of the empathy box cause genuine flesh wounds in reality, but the issue of whether the religious experience is a supernatural one that crosses the boundary between mind and matter is unresolved in the novel. But regardless of its source, Mercerism produces a socially adaptive effect in allowing human adherents to distinguish themselves from affectless android mirrors. As the android, Baty, resentfully observes, "Isn't it a way of proving that humans can do something we can't do? Because without the Mercer experience we just have your word that you feel this empathy business, this shared group thing" (179-80). The android raises a genuine paradox, since by definition the ability to enter the minds of others is arbitrary. Because it cannot be proved absolutely – we can only infer minds from bodily actions and language – we cannot know absolutely that our perception of what others are feeling is not merely our own projection. According to this version of the "brain in a vat" thesis, it is possible that others, humans and androids, are always barred from reciprocating our empathic investment in them, since we cannot know for sure what they are thinking, or even if they possess any comparable consciousness of other minds. As Dick observed in one essay, it is possible that "each human being lives in a unique world, a private world, a world different from those inhabited and experienced by all other humans...If reality differs from person to person, can we speak of reality singular, or shouldn't we really be talking about plural realities?"[10] Mercerism in itself cannot therefore be a framework within which categorical distinctions between human and android can be made because of the broader epistemological problem of knowing other minds. Indeed, Mercerism may actually add to that problem by existing as a purely psychological, hallucinatory and relativistic world with no grounding in universal, physical causes. Although the omniscient narrator is able to inform us of how Isidore and Deckard experience Mercerism from their own subjective

---

[10] Philip K. Dick, "How to Build a Universe that Doesn't Fall Apart Two Days Later," 1978, Sutin 261.

point of view, there is no way of knowing whether the other human minds they share when in that experience also perceive them back in a comparable manner.

However, the fact that Mercerism exists at all can be treated as an absolute truth that transcends the psychologically constituted forms which religious narrative takes (just as, so I argue later, the fact of the novel becomes for the reader proof of imaginative humanity, no matter how each individual reader interprets the same text). Towards the end of the novel, Buster T. Friendly, the inane chat show host who broadcasts 23 hours a day on both television and radio (he is, we guess, an android) makes a dramatic exposé. He proves that the mountain Mercer climbs is a set on a Hollywood sound stage, with the role of Mercer played by an alcoholic bit-part actor, Al Jarry who "made a repetitious and dull film, a series of them in fact, for whom he knew not – and does not to this day" (179). Friendly notes that when adepts claim Mercer is not a human being they are ironically correct, because "Wilbur Mercer is not human, does not in fact exist. The world in which he climbs is a cheap, Hollywood, commonplace sound stage which vanished into kipple years ago" (179). The androids' expectation is that once the deception is revealed, the deeper sensations and foundations on which judgments of human and android consciousnesses are predicated will crumble away, given that they cannot, by the very nature of empathy, be objectively validated. But whilst this may seem to be true in theory, the ironic outcome evidences that what matters about religion is not its fictitious nature *per se*, but the fact that it works in social practice.

This is the crucial difference between Cartesian arguments that draw rationality from the axiomatically assumed existence of a divine, and a secular science that cannot investigate the empirical nature of metaphysical spheres, but which can acknowledge the anthropological value of belief. As William James explores in his *Varieties of Religious Experience*, regardless of the veracity of the divine, so long as we proceed as if God exists (based on our personal, inner experience) then the equivalent improvement in our moral life will have been achieved.[11] So here it matters not whether Mercer exists in reality, or indeed whether one's pet is real or artificial, so long as the human subject is capable of proceeding *as if* they are real; indeed, Mercer himself tells his followers to cultivate independence and to act "as if I didn't exist" (152). Empathy, the ability to perceive the world as if from the perspective of the other, works as a discriminating factor between

[11] William James, *Varieties of Religious Experience: A Study in Human Nature*, centenary edition, (1902; London: Routledge, 2002). Lecture 3, on "The Reality of the Unseen," deals with this topic in particular (46-65).

humans and androids, because it tests the limits of rationality. It forces the human to proceed on a course as if it were true, whilst objectively knowing it not to be so, or whilst knowing that the state of other people's perceptions cannot be verified objectively. If objective, scientific methods for discriminating true and false have broken down in the wake of the permeation of simulacra, *Do Androids Dream of Electric Sheep?* romantically contends that the optimum knowledge is in any case idealistic and subjective rather than rational and objective. Note how far from Paré's reasoning this conclusion brings us. For whilst in the Renaissance those who perceive monsters through the work of demons superstitiously fail to rationalise the monstrous as unusual but not beyond nature, in Dick's work the one thing that conclusively defines the android and the human is the latter's willingness to be deceived by simulacra. If religion is actually a Feuerbachian construct designed to worship man's own nature – in this case especially his ability to empathise – religion can be sustained at any cost to human rationality because it has such a deep rooted social value. Whereas Ambroise Paré or Descartes deny the influence of a deceiving demon because God would not permit it, here the logic is inverted: the plausibility of deceit is evidenced through the absence of a truly metaphysical God, but this actually validates religion as the standard by which the truth of humanism can be obtained. Religion is still a part of the immaterial world of human relationships which is itself a part of reality, even if it is not tangible and reducible as physical domains of scientific enquiry. Thus Isidore is able to retort confidently at the moment of the android's triumphalism "Mercer isn't a fake...unless reality is a fake" (201). Such a view – which might also provide Dick's own answer to the conundrum of solipsism raised in "How to Build a Universe" – is reminiscent of Peter Singer's argument that although ethics is not objectively written into the fabric of the universe, in a world of human individuals, shared social subjectivities must equally be substantiated as empirical qualities of reality.[12]

As so often in Dick's writing, however, no solution can be unambiguous, and an apparent allegory about the value of religion is also satirically inflected. For if humans are willing to be deceived by their own creations and to use this suspension of disbelief as the very condition of their humanity, there is something false and unstable about their reliance on a simulacrum – more artificial, perhaps, than the androids themselves. Mercerism's empathy box must be understood in conjunction with the Penfield mood organ, which mediates and determines moods. In the opening sentence of the novel, the device awakens

[12] See chapter 1, note 93.

Deckard, having been pre-programmed to make him be "glad to be awake" (3). Getting into an argument with his sleepy wife, Deckard ponders whether to dial "for a thalamic suppressant (which would abolish his mood of rage) or a thalamic stimulant (which would make him irked enough to win the argument)." As his wife threatens to dial for "greater venom," Deckard relents and sets the mood organ according to the demands of his day's schedule; he suggests his wife dial "888. The desire to watch TV, no matter what's on it" (6). This miniature domestic comedy offers a sharp critique of contemporary society. As with Mercerism which is revealed as a hoax by an incessantly babbling chat show host, the mood organ condemns the suppression of individuality in a culture which allows one to buy experiences which offer immediate but crass escapes from reality; television and drugs are obvious real-world analogues to Mercerism and the Penfield mood organ. In "The Android and the Human," Dick ties this problem of illusory escape through commoditised mechanisms to the state of being of the android:

> Free will for us – that is, when we feel desire, when we are conscious of wanting to do what we do – may be even for us an illusion…And – here is a thought not too pleasing – as the external world becomes more animate, we may find that we – the so called humans – are becoming, and may to a great extent always have been, inanimate in the sense that we are led, directed by built-in tropisms, rather than leading. So we and our elaborately evolving computers may meet each other halfway.[13]

The androids alleviate one Cartesian paradox by providing affectless Others who validate the humanity of those capable of empathy: if, for Descartes, an automata would not be possessed of a theological soul, for Dick, an android would not be capable of believing in a religious narrative even – or especially – if he knew it to be objectively unverifiable or even false. In the scenario of technological deception by a Hollywood-style Mercerism or Penfield mood box, however, this emphasis on shared consciousness opens space for another contradiction: the methods by which consciousness can be shared are themselves deterministic, because the mechanism of universal psychological mediation must be commercially distributed so that its experiences can become accessible to many.[14] As

---

[13] Philip K. Dick, "The Android and the Human," 1972, Sutin 187.

[14] Dick's earlier work, *The Three Stigmata of Palmer Eldritch*, plays off two competing drugs pushed by advertising companies. The one, Can-D, induces temporary hallucinations in which off-world users undergo communal experiences of suburban America. The rival product, Chew-Z, is superior, in that the dreams can last forever whilst time in reality does not pass at all. However, it induces psychedelic fantasies that require

Buster asks the unresolved question, "who, then, has spawned this hoax on the Sol system?" (179). If hallucinations define the human, the shape of human society can be conditioned through the control of those same systems of deception (the answer we are probably expected to supply in response to Buster is: the government). Paradoxically, technologies like drugs or television facilitate predictable behaviours in social and domestic situations which allow people to continue to cohabit even under the extreme duress of an entropic world. However, in directing us as part of a social whole, technology also deprives us of the crucial aspect of our individuality – free will. As Dick noted, "Fake realities will create fake humans. Or, fake humans will generate fake realities and then sell them to other humans, turning them, eventually, into forgeries of themselves."[15] Isidore is astonished that Pris does not have her own empathy box which is, he explains, "the most personal possession you have! It's an extension of your body; it's the way you touch other humans, it's the way you stop being alone" (57). Of course, it is anything but personal if everybody is implicitly required to own one. It is only through the revelation that Mercer is himself artificial that technology is restored to its proper place which is, like the androids, as a vehicle for an already innate capacity (to empathise) rather than a total substitution for the normal human experience. Nevertheless, such a realisation occurs only at the end of the novel. For much of it, Mercerism, artificial animals and the Penfield mood box keep human society integrated by facilitating their ability to empathise, but they also keep human society automated by directing them down proscribed paths to maintain social cohesion.

However, there is another vehicle or "mood organ" which offers an opportunity to empathise, but which does not require the contradictory use of technologies that determine the self through civic controls. Indeed, it is precisely its indeterminacy that constitutes it as a focus through which empathy can be projected, even though this vehicle is always already known as an artificial simulacrum without strict ties to physical reality. Perhaps the most damning but subtle critique made by this novel is that humans no longer read stories

---

none of the Mercerism-like empathy of Can-D. Whilst satirising the capitalist companies which push both products, Dick's sympathies are clearly with this latter sort of experience. Philip K. Dick, *The Three Stigmata of Palmer Eldritch*, *Four Novels of the 1960s*, ed. Johnathan Lethem (1965; New York: Library of America, 2007) 231-430.

[15] Philip K. Dick, "How to Build a Universe That Doesn't Fall Apart" 263-64.

at all. Books are regarded as kipple to be blasted into space, where they are collected by android colonists. In real life, however, the ability to perceive the world from the perspective of others and to believe the value of other consciousnesses whilst knowing them as foundationally untrue, is precisely what is required by the reader of any fictions, even Dick's peculiar hallucinations. *Do Androids Dream of Electric Sheep?* holds to a literary humanist's belief that the ability to read and project consciousness into the fictional world is in and of itself a moral purpose, one science and technology cannot reproduce.[16] Regardless of precisely what is said within that heterocosm, the meaning of which can always be debated, by providing an exercise in empathy in its own right, fiction denotes the reader's humanity precisely because its meanings are not deterministic. It cannot therefore be used for mass social engineering in the way Mercerism or the Penfield box are, but simultaneously it can encourage the reader to realise that he or she inhabits a shared, common universe of values which might therefore be objectively true in a stronger, Singerian sense. Hence the most significant, liminal moment of deception is critical not only for Deckard, but also for the reader. This complex scene tests whether Deckard, through whose eyes we have seen the world, is himself an android. It thereby self-reflexively examines the role of fiction as an empathic device, and attends to the implicit epistemological contract that exists between author and reader. Just as Descartes can note the omnipresence of hallucinations like dreams but deny the possibility of a totally deceiving demon because of the existence of God, the novelist enters a comparable contract (a *potentia Dei ordinata*) with his readers which agrees that we are to be deceived by local ontologies of difference (such as androids who resemble humans) but not by the broader ontology of the world of the novel as a whole.

[16] Of course, one could argue that narrative empathy is not necessarily in and of itself proof of one's moral humanism. Hannibal Lecter, for example, is such an effective killer precisely because, as an aesthete, he is capable of anticipating the psychology and actions of his victims. In the age of cybernetic fictions, the ability of the computer game player to project his mind into the virtual context can arguably lead to a comparably dangerous inability to distinguish between the game world, in which violence is rewarded rather than condemned, and the real world, in which the effects of violence are devastating. Teenagers William and Josh Buckner, for example, killed and injured two car drivers by sniping them as they passed. Their assertion that they had simply been acting out the gangster game *Grand Theft Auto* led to a $246 million lawsuit being filed against the game's producers. David Kushner, "Grand Death Auto," *Salon.com*, 22 Feb. 2005, 11 Dec. 2008 <http://dir.salon.com/story/tech/feature/2005/02/22/gta_killers/>.

Humanism and Literature

Mid-way through the narrative, Deckard has successfully retired his first target, Polokov. Following his extermination list, he next administers a Voight-Kampff test to the beautiful soprano, Luba Luft. Accusing him of being a sexual deviant (many of the Voight-Kampff questions ask about the subject's sexual desires), Luft calls the police. Officer Cram who arrests him claims never to have heard of his fellow police officer, Deckard, and so Deckard calls his superior, Bryant, to prove his identity. When Deckard passes the phone to Cram the line goes dead and, confusingly, the redialled number is never answered, leaving Deckard (and the reader) "bewildered" (92). Cram flies Deckard to the city's Hall of Justice, but it is not the one at which Deckard works:

> It makes no sense, [Deckard] said to himself. Who are these people? If this place has always existed, *why didn't we know about it?* And why don't they know about us? Two parallel police agencies, he said to himself; ours and this one. But never coming into contact – as far as I know – until now. (97; emphasis in original)

Having tried unsuccessfully to make another phone call, this time to his wife, Deckard is shown to the office of the police chief, Garland, who takes Deckard's hit-list and discovers that he himself is the bounty hunter's next target. Garland accuses Deckard of being an android posing as a bounty hunter, who is in fact killing humans. Garland summons the bounty hunter for this precinct, Phil Resch, to check if Deckard is on Resch's list. Suddenly, then, the whole world ontology of the novel has become inverted; as Garland's speech takes over from the objective point of view of the narrator to explain the terms of this *gedanken* test, we fear that human and android have switched places:

> This man – or android – Rick Deckard comes to us from a phantom, hallucinatory, nonexistent police agency allegedly operating off the old departmental headquarters on Lombard. He's never heard of us and we've never heard of him – yet ostensibly we're both working the same side of the street. He employs a test we've never heard of. The list he carries around isn't of androids; it's a list of human beings. He's already killed at least once. (101)

Is Garland an android, as Deckard suspects? Is the other bounty hunter, Resch, an android? Or is Deckard himself an android who has been given a false memory to allow him to believe himself to be human? If this latter is the case, how can we believe the truth of anything narrated as we followed his activities in the preceding plot? Were those "androids" Polokov and Luft really humans?

The perceptual instability here reworks the Cartesian demon into the heart of the narrative, in a recursive deception in which the fiction has been a hallucination two levels removed from the real world of the reader trying to work out what is true: we have apparently been reading a fictional novel, whose events are narrated by an unreliable narrative voice, which then guides us to a moment when the character generated by that voice is himself a hoax of the human. It is only when a post-mortem bone marrow test proves that Polokov was indeed an android that the categories start to normalise themselves into objectively known facts. Whilst Resch leaves to collect his empathy-test equipment based on the evidence that Polokov was physiologically an android and Deckard's retirement of him was therefore legitimate, Garland – knowing he will be uncovered by the test – resignedly admits to Deckard that he is an android, and explains the nature of the precinct, including why Deckard's phone calls did not reach Bryant or his wife:

> All our vidphones here are trapped. They recirculate the call to other offices within the building. This is a homeostatic enterprise we're operating here, Deckard. We're a closed loop, cut off from the rest of San Francisco. We know about them but they don't know about us. Sometimes an isolated person such as yourself wanders in here or, as in your case, is brought here – for our protection. (106)

The android concedes that though he is participating in a systematic hoax, the Voight-Kampff test will still be able to oust him as an android, drawing a final line between the real and the simulation. But whilst the alternative police station will be verified as unreal, a simulation, it is still unclear whether Deckard is himself an android. Given the hermetic nature of the deception, it could realistically be so: both Deckard – and vicariously the reader – might have been deceived by the novelist plotting the only world we can know, implanting "false memories" figuratively and literally in his character and reader. Whilst the Voight-Kampff test will still work in spite of the homeostatic fake police station, what test can be applied within the closed loop that is Deckard's consciousness to confirm for the reader whether he is human? If the omniscient narrator has been deceiving us up to this point by making us think that the "hero" Deckard was human, his decision to subject Deckard to a Voight-Kampff test could equally be a plot, in both senses of the word.

Rather than having the omniscient narrator state objectively, but surely by now untrustworthily, that the character is a human or an android, Dick's method here is to draw attention to the artifice of the wider simulation that is fiction, in order to verify the nature of the key character within it. And, paradoxically but deliberately, Deckard's humanity is

indicated by likening him to a machine. When Resch returns, he intuits that Garland is an android and so kills Garland. Resch asks Deckard what he and Garland had discussed whilst he had been away, and Deckard replies:

> "That he – it – was an android. And you" – Rick broke off, the conduits of his brain humming, calculating, and selecting; he altered what he had started to say. " – would detect it," he finished. "In a few more minutes."
> (107)

When we see from Hal's eyes as Poole and Bowman wonder how he would feel about disconnection, we momentarily forget that the existence of his perspective acknowledges that he has a consciousness in the first place. The inverse pattern is at work here. As Rick's mind is described in terms of the electronic cyborg or robot, the language draws attention to the metaphorical work it is doing, making false connections between two entities that are in reality quite different: the human brain and the cybernetic one. The novel generally tries to show the potential similarity between humans and androids in terms of their lack of sympathy, regardless of their precise origin as naturally or artificially-created entities. For this reason, the novel is very hazy about how exactly androids are made. Although the logic of the plot and unreliability of the narrator means that Deckard could be an android, the fact that Dick's omniscient narrator suddenly chooses to describe his constitution in terms of mechanism indicates that this is only a metaphor rather than a realistic description of his inner being beneath the surface. It marks Deckard out as a machine so definitively and obviously that it seems anachronistic in a novel in which most often androids and humans are so alike both on and beneath the surface that it is impossible to tell them apart. If the android's biological difference can be truthfully discovered by a post-mortem dissection in a validation of the empirical method, it is precisely the opposite of empirical representation, the literary metaphor, which provides access to the truth of Deckard's being beneath his skin, in his (human) mind. The novel's mode of creating knowledge and narrative truth is thus played off against that of science.

Whilst cyberneticians and philosophers dispute whether the mind is metaphorically or physiologically like a computer, the novel's use of language and metaphor suggests that this debate is focused on the wrong sort of difference. As Dick presents his ideas in "The Android and the Human," acting *like* a machine should be sufficient for the self to be considered *as* a machine in terms of ethical values, even if not in terms of biology. Whether an android is built like a man is irrelevant, because a man could act like an android in his psychological isolation; the difference that matters is between extending thought to other cognitive agents versus remaining solipsistically locked in one's own

world of perception. In this new cybernetic environment, then, it is not science but literature that might provide a better model for discerning categories, in which the inhuman is that which cannot empathise, whether it originates in the laboratory or not. Although according to the explicit signification Deckard is an android because his brain hums and calculates like an electronic computer, implicitly this statement defines him more determinedly as human precisely because it draws attention to the way in which his character is constituted out of the arbitrary mechanics of language. In the same way, it is ironically just when Mercerism is shown to be false that this verifies the absolute ability of humans to empathise in spite of it. The reader's ability to note this ironic mode of representation shows them to be capable of seeing beneath Deckard's skin – that is to say, of empathising with him as a human – even though, or even because, the novelist is using metaphors that mislead us as to his true being. Turing's cybernetics aimed to reproduce intelligence without worrying about the physical mechanisms by which this is achieved; self-reflexive fiction, on the other hand, draws attention to the inner workings of language and makes us realise that the man, Deckard, is other than the sum of the explicit language by which his character is produced. There is a discrepancy between the explicit, objective words and the character these words conjure in the mind of the reader. The gap is produced because the reader interjects empathy with Deckard, regardless of how physiologically inhuman he is stated to be, how his agent is shaped as android by the explicit signifying mechanism of language and metaphor. This passage is mirrored by a moment late in the novel, when Isidore temporarily leaves the flat inhabited by his newfound android friends:

> The silence, all at once, penetrated; he felt his arms grow vague. In the absence of the Batys and Pris he found himself fading out, becoming strangely like the inert television set which he had just unplugged. You have to be with other people, he thought, in order to live at all. (175)

Isolated, and therefore deprived of any means of expressing empathy, Isidore is almost machinic in his consciousness. On the other hand, the fact that he recognises himself as alone must be given priority over the literal signifiers in which that state is couched. Isidore may be becoming "unplugged," but because that is such an evident *metaphor* it emphasises the ontological gap between the human Isidore and the androids with whom he lives who might *literally* be unplugged because they are machines, or were originally the product of some mechanical process. Thus Dick's work attends to the parallels between the mechanism of literature and the mechanism of the android, whilst at the same time indicating that the phenomenological effects which the former produces are very different to those deterministic ones experienced by the cybernetic entity, in which the program

initiated in them from their point of creation remains with them. Their nature governs their nurture, whereas the human ability to nurture others (such as pets) correspondingly changes their deeper, inner nature in morally positive ways. We can intuit humanity in spite of that human being described in mechanistic terms, so that there is a gap between the literal word and its implied meaning. By contrast, cybernetics argues that human consciousness is essentially the same as all nature, because it is an instantiation of information and codes that determine its constitution and that the literal code (such as the genetic marker, or binary program) will always lead to a predefined outcome, such as a particular organism or computer routine.

The passage in the precinct takes us through an perceptual crisis – has the character via whom we have witnessed the previous events actually been an android impersonating a bounty hunter? – in order to create the epistemic grounds by which a more subtle ethical question can be asked. Rather than the problem consisting in how to distinguish human from android, the issue is whether killing androids would require humans to become dispassionate in ways that subvert man's natural empathy. Resch lacks empathy for the androids, whereas Deckard possesses it. Having left the precinct, for example, he buys the opera singer Lubya Luft a picture she desires, before Resch kills her. If Resch is an android this will allow Deckard to "undergo renewed faith in the human race" (120), proving humans to be capable of empathy in the way androids like Resch are not. On the other hand, if Resch is human then it evidences that a good bounty hunter must be less sympathetic than the androids themselves, who can at least make a passable simulation of emotional warmth. Ultimately, Resch does test as genuinely human, with the Voight-Kampff score revealing also "a defect in your empathic, role-taking ability. One which we don't test for. Your feelings towards androids" (121). If the android is totally unlike the human as an affectless agent, androids can be retired without any sentimentality, and Resch's view of the androids as objects should not be classed as a "defect." However, though Deckard is cognitively abnormal because he sympathises with Luft's plight, he is also the more humane because he feels "empathy towards an artificial construct...something that only pretends to be alive." His feelings are "the reverse of those intended. Of those I'm accustomed to feel – am *required* to feel," and Deckard realises "there's nothing unnatural or inhuman about Phil Resch's reactions; it's me" (121-22). We now have a paradox. Resch scores as human on the test because he is able to look upon androids as totally inhuman; but if humanity is also defined by the ability to empathise, it is Deckard who is the more human. The novel appears to self-cancel its own epistemic basis by which the idea of the human is constructed. Such contradiction, and the inability

to resolve the true from the false in any final way, might be symptomatic of the postmodern aesthetic. However, given the way in which Mercerism and the mechanistic metaphors of the fiction emphasise that to be human is to be capable of empathy *despite* being told that the other is a simulacrum, it is clear that we are to ally ourselves with Deckard's personal crisis against Resch's callousness. The strict ethical logic of the novel is in tension with our contrary ability to sympathise with Deckard precisely because he is in breach of the legal framework. The crisis of deception at the centre of the novel is both moral and readerly. It threatens to destabilise our ability to enter the narrative world as if it were real, an ability which validates our status as empathic (that is, definitively human) readers, and our ability to choose between human and android objectively once within the novel's moment. This dialectic between androids, cut-off from other minds, and readers necessarily entering a world and in a sense bringing it into being by reading about it is in line with Josipovici's view of fiction: the android makes us realise the ironic value of fiction that is both a simulation or deception, and yet which is also psychologically real in a way that protects us from solipsism and epistemological relativism. If the android is not human, then we are forced to admit the limits to our capacity to empathise with those who are not like ourselves, even though they are so very closely modelled upon us. Rather than treating androids and humans, or artificial intelligences and humans, by definitive categories of "us" and "them" – as if categorical definitions can be established between things that are not really different – both *2001* and *Do Androids Dream of Electric Sheep?* attempt to overcome our solipsism through attending to epistemological questions. Our sympathy or repulsion should not be determined by the material substrate on which the other mind is founded. Indeed, whether named as android or human, all the "characters" are essentially cognitive constructs, emerging from the same material substrate of printed words but also from the way in which these are interpreted by an empathetic reader. The novel precedes any *a priori* ontological revulsion at the alterity of the android, allowing us the space in which to enter its mind safely and to propose epistemological questions about linguistic realism. What matters in asserting humanity is, paradoxically, to allow oneself to suspend materiality, to become present within the mind of the similarly immaterial (or differently material) other but also to be aware that our own reality is being suspended in the process. It is such a self-reflexivity that is lost in systems like the Penfield mood organ or Mercerism, but which the novel can sustain by attending to its own metaphors by which a human character with whom we empathise is also, ironically, created out of the mechanics of language. Whilst Deckard's empathy for androids hampers his ability to do

his job, our ability to imagine the world as if through their lives stresses our capabilities as actively interpreting readers.

In relation to Cartesian deception, therefore, *Do Androids Dream of Electric Sheep?* replays it in a double way. Like Descartes (or Turing), Dick stresses effects rather than origins. Regardless of whether an agent is artificial or human, the ability to empathise with another self is the hallmark of the human. Because this is where the novel stands *a priori*, it argues that the ability to occupy the subjectivity of another is axiomatically moral, that the ability temporarily to possess other minds and perspectives on reality is itself a humanistic act. In achieving this, though, it provides several moments of deception, when the reader fears it is impossible to choose between self and other based on rational and empirical criteria, and that the novel is itself one example of a total deception rather than proof of the reader's humanity. However, by pulling us through these liminal moments, we recover confidence in the ontological boundaries between android and human, and fiction and reality. This is an issue I pick up again in relation to *Foucault's Pendulum* (chapter 8), where I draw some generic conclusions about cybernetic fictions in literature versus film. For now, though, we can observe that where the film *Blade Runner* differs most from *Do Androids Dream of Electric Sheep?* is in its elimination of these Cartesian tests which, though they create crises of deception, also facilitate opportunities for resolution. In *Do Androids Dream of Electric Sheep?* epistemological and ontological questions abound – most notably in the precinct episode – but critical opportunities for their resolution are presented through an essentially conventional belief that the very existence of the work itself can alleviate the complexities within it. In postmodernism, as exemplified by Ridley Scott's film, the artefact too is implicated as embodying the very processes it critiques, the work itself becoming one more piece of evidence that we live in a total simulacrum.

Seeing Through the Android in *Blade Runner*

When Ridley Scott claimed not to have read *Do Androids Dream of Electric Sheep?* before making the film,[17] he was making the point that *Blade Runner* is far removed from its origins.[18] Indeed, he chose the word "replicant" over "android" because

---

[17] A claim Scott makes in Sammon 283.

[18] There were five publically screened versions of the film, with three principal distributions: The Domestic Cut (the original theatrical release), The Director's Cut, and The Final Cut. The original theatrical release (1982) had Deckard providing a narratorial voiceover. More significantly, the film ends positively, with Deckard and Rachel escaping to the countryside and with Deckard explaining that unlike the other replicants, Rachel is

he did not want *Blade Runner* "to be premonitory of *android* at all, because then people would think the film is about robots, when in fact it isn't. It was better that we come up with a new word altogether."[19] Although this might prevent us from seeing *Blade Runner* as a cybernetic fiction, in fact the shift is more interesting, because it suggests that the cybernetic paradigm had, by the 1980s, become so rooted as to be almost self-explanatory. The word "android" was not necessary to show that the film was about processes of late capitalism of which cybernetic systems were the most potent examples. Certainly, the biological replicants are just one among many simulacra in a cityscape named as Los

---

not designed to die after four years. Ridley Scott was pushed into adding this optimistic ending and explicative voiceover in response to disappointing audience reactions at the preview screenings. His Director's Cut (1992) removes Deckard's narration, creating a much bleaker, minimalist film. The ending in this version is made ambiguous through the addition of an earlier scene. As in the first release, as Deckard leaves with Rachel he sees an origami·unicorn, the calling-card of a fellow blade runner, Gaff. However, this vision is given a twist by the reinsertion of an earlier scene in which Deckard dreams of a unicorn whilst sleeping at a piano. As a possible indication that Deckard's memories are false, this new connection between his dream and the conclusion implies that Deckard is himself potentially a replicant now on the run with Rachel. As Ridley Scott explains: "If you take for granted for a moment that, let's say, Deckard is a Nexus 7, he probably has an unknown life span and therefore is starting to get awfully human. Gaff, at the very end, leaves an origami, which is a piece of silver paper you might find in a cigarette packet, and it's a unicorn. Now, the unicorn in Deckard's daydream tells me that Deckard wouldn't normally talk about such a thing to anyone. If Gaff knew about that, it's Gaff's message to say, 'I've read your file, mate.'" Though restoring this narrative dimension, the remastering of the Director's Cut was rushed by the studios. The Final Cut (2007) has been fully restored under Scott's authority, but there are no substantive alterations that fundamentally change the plot, in the way the restoration of the "unicorn" scene does. Since the Director's Cut is that which may be most familiar to readers whilst also representing Scott's key thinking even if not his technical ambition, this is the one I will discuss. For Ridley Scott's explanation of the key changes, from which the quotation above is taken, see Ridley Scott, interview with Ted Greenwald, *Wired*, 9 Sept. 2007, 09 June 2008 <http://www.wired.com/entertainment/hollywood/magazine/15-10/ff_bladerunner_full>. For a detailed, scene-by-scene explanation of the changes across the cuts, see Sammon 394-408.

[19] Sammon 61; emphasis in original. It was the daughter of screenwriter David Peoples who proposed the term based on her work in microbiology (Sammon 62). The term "replicant" was used in the infamous passage of *The Selfish Gene*: "They are in you and me; they created us, body and mind; and their preservation is the ultimate rational for our existence. They have come a long way, those replicators. Now they go by the name of genes, and we are their survival machines." Dawkins 19-20.

Angeles of 2019.[20] The bustling underworld comprised of Oriental and Western cultures is mirrored by a bourgeois world at the top of the skyscrapers which grow ever higher in duplicated layerings; advertisements (for genuine virtualising systems such as Atari or TDK video) provide a backdrop to the frenzied activity; our first sight of Deckard sees him leaning against the window of a television shop where banks of screens play the same image; the android Zhara is killed in a shop full of mannequins. As well as a culture of reproduction, such scenes exemplify the urban "sprawl" common to cyberpunk, in which the "image of the carnivalized city, the city as permanent carnival" is used as a signifying analogue for the postmodern condition.[21] If *2001* used a vast universal space to encourage a reconceptualisation of man's (or the posthuman's) anthropic place within it, *Blade Runner* presents a closed city space that illustrates there is no paradigmatic distinction between human and posthuman, or the present and the future. *Blade Runner* collides different worlds through what Syd Mead, its credited "conceptual futurist," termed "retrofitted utilization" in which elements of the past survive beneath a technological surface superadded to it.[22] Old apartment blocks have ventilation and electric conduits on the outside; noodle bars front scientific laboratories; cars have wheels but can also fly. It also blurs genres by transplanting the conventions of film noir (the burned-out detective; light filtered through gratings; a voiceover narration, in the case of the first cut) into a futuristic cyberpunk thriller, precisely the sort of pastiche Frederic Jameson identified as inherent in postmodernism, in which the "styles of the past" are emptied of their significance by being combined in "overstimulating ensembles."[23] In this heterogeneous

[20] For an interpretation of the relationship between the real and the fictional city, see Steve Carper, "Subverting the Disaffected City: Cityscape in *Blade Runner*," Kerman 185-95. Kevin McNamara sees the film (and the novel) as politically engaged with the problems of city planning and multiculturalism: Kevin R. McNamara, "*Blade Runner*'s Post-Individual Worldspace," *Contemporary Literature* 38 (1997): 422-46, *JSTOR*, 30 Apr. 2008 <http://www.jstor.org/stable/1208974>.

[21] McHale, "Elements of a Poetics of Cyberpunk" 154.

[22] Syd Mead, "Designing the Future," *Omni's Screen Flights, Screen Fantasies: The Future According to Science Fiction*, ed. Danny Peary (Garden City, NY: Doubleday, 1984) 199-213.

[23] Frederic Jameson, *Postmodernism, or, The Cultural Logic of Late Capitalism* (Durham, NC: Duke UP, 1991) 19. This book was an expanded version of an essay published a decade earlier as Fredric Jameson, "Postmodernism, or, The Cultural Logic of Late Capitalism," *New Left Review* 146 (1984): 53-92. Subsequent references refer to the book rather than the original journal article.

world, then, where does the true narrative lie? As, according to McHale, postmodern texts tend to give precedence to ontological issues before epistemological ones, typically absent from *Blade Runner* are the proposition of explicit, Cartesian models, of the sort placed at the centre of Dick's novel.

Whereas *Do Androids Dream of Electric Sheep?* presents deceptions that facilitate better self-reflection on the difference between android and human, the film flirts with the epistemological problem of how to distinguish true from false without reaching definitive conclusions. In this sense, *Blade Runner* is ontologically a total delusion, as opposed to the microcosmic deceptions inherent in Dick's work. Troping this is the way eyes dominate *Blade Runner*, providing "the symbolic core of the film's representational conflict."[24] (One with psychoanalytic echoes of Hoffman's "The Sandman".) An eye – we do not know whose – reflects the fiery cityscape at the start of the film; the Voigt-Kampff machine hones in on eyes testing the emotional reflexes of the iris; a cloned owl's eye glints a threatening red; Pris sprays her eyes with black makeup; the panoptic police eye hovers in the sky; and (a different sort of vision) the domestic sets are stocked with photographs. The film seems to want us to question the reliability of the eyes and perspectives, suggesting that rather than one objective view of reality, reality is dependent on the angle of each viewer. In one key scene, the replicants Roy Batty and Leon visit the workshop of the genetic engineer of their eyes, Chew, and Batty jokes to him: "If only you could see what I've seen with your eyes." The semantics are hard to decipher at the moment of watching the film, and because it is impossible to imagine immediately what it would be like to see through the eyes of a craftsman implanted into his creation looking back at his maker, it seems to deconstruct and interweave perspectives without presenting an alternative, optimum vision. The moment is symptomatic, because even as it shows through the represented eyes how many perspectives there might potentially be, the film denies us access to other viewpoints, thereby preventing the connection between viewer and character so central to Dick's notion of empathy.

The fact that *Blade Runner* is a film means that it would be wrong to see this alone as defying the capacity for perspectival relativism afforded by the novel; as Dick himself noted, representing the novel on screen would be inherently difficult because of its two narrative strands in Isidore and Deckard, and the fact that there are two visually identical

---

[24] Jack Boozer Jr., "Crashing the Gates of Insight: *Blade Runner*," Kerman 212-28.

characters in Rachel/Pris.[25] Further, the novel is able to provide a logos for humanism by reflecting on the difference between its explicit textual signifiers and the way in which the empathic reader can defy the objects or things they represent. We are, for example, able to sympathise with Deckard in spite of his being described in the lexicon of robotics. Significantly, however, *Blade Runner* actively draws attention to these deficits, reflecting on its own limitations and the editorial process of making a cinematic vision, as opposed to a textual one. The film's single camera lens is the only way in which we are allowed to know the world of the other, and that we cannot enter the consciousness of other selves is not only an innate, formal constraint of the media but a didactic emphasis of the film. The key scene in this respect occurs when Deckard inserts a police photograph of Leon's apartment into the Esper computer, which zooms and, bizarrely, pans around the corners of the two-dimensional image to reveal and enhance features not present at normal magnification or occluded behind scenery in the photograph. As well as reiterating the paranoid excess of this surveillance city, it is a clear metacritical analogue for the process of framing and editing performed by the movie we watch, showing how truth (which is what the detective seeks) can be essentially fabricated.[26] Drawing attention to the process through which a particular story is viewed, the detection through a high-tech magnifying glass is analogous to the fake police precinct, which is absent from the film adaptation.

That original scene pulls us through the epistemological challenge. Proposing that Deckard's eyes (via which we experience the novel) may themselves be suspect like the androids he pursues, the language in which his mind is described undermines this supposition and takes us to a plateau of secure understanding about his humanity. Though this raises further moral questions about whether a good bounty hunter should extend empathy to androids, the state of the world in which those dilemmas exist is ultimately trustworthily narrated, showing that the reader who invests consciousness in it is capable of something androids are not: making subtle ethical judgements. In this case we decide Deckard is more humane, even though he is able to empathise with androids. The key

[25] Philip K. Dick, "Notes on *Do Androids Dream of Electric Sheep?*" 1968, Sutin, 155-61. Incidentally, Dick imagined Gregory Peck playing the part of Rick Deckard, with Grace Slick as the "vibrant hard girl," Rachel.

[26] Rutger Hauer, who played Batty, held to this interpretation of this scene. Sammon 146. Giuliana Bruno also extracts this scene as particularly significant in Giuliana Bruno, "Ramble City: Postmodernism and *Blade Runner*," *October* 41 (1987): 61-74.

difference between book and film is that, in the latter, we are never sure whether Deckard is an android or human at root. As is exemplified in the Esper device, the film raises the possibility of multiple perspectives whilst nevertheless being the only perspective we are permitted. Like the problem of consciousness itself, the private eye's view of the world remains literally private, and so they are always potentially those of a replicant. For example, Deckard's eyes briefly flash red, like those of the androids or of the artificial owl in Tyrell's office; in the Director's Cut (see note 18 above), the restored unicorn dream sequence resonates with Gaff's origami calling card at the end, implying that Gaff knows literally what is in Deckard's implanted memories; Gaff's parting comment to Deckard, "You've done a man's job, Sir" contains a double significance. However, these glimpses which disturb ontology at the microcosmic level (is Deckard a replicant?), are given a powerful twist by being contained within a film that attends to its own manufactured status, its macrocosmic ontology. The Esper device ties this possibility of artifice within the character to the global artifice of the film as a process. Before honing in on the snake scales, the Esper device pauses over a newspaper containing an advert for an Oriental restaurant, reminding us that our first sight of Deckard was of him reading his newspaper whilst waiting for a table at a noodle bar; it also pauses on a textured whisky glass like that Deckard drinks from; finally, the apartment in the photograph looks, in its sepia colouration, very similar to Deckard's as represented in the film set. Deckard explains to Rachel that the photographs of her mother are designed to bolster her false, implanted memories. But if a photograph gives the illusion of memory, then the photographs or film images that relay Deckard's character to the viewer could be intended to give a misleading impression, this time implying Deckard is an android when in fact he is human. Through the metacritical Esper device, the film demonstrates that we have no way of knowing other than itself, and it defies any opportunity to get at the truth of Deckard's identity. Though Scott insisted that Deckard had to be a replicant, in fact the film is more subtle and the interpretation more ambiguous than Scott's retrospective emphasis suggests, and is never confirmed one way or the other, only raised as a possibility.[27] As Kevin McNamara observes, though Scott may have intended Deckard as a replicant, if he is definitively so we "have another rather uninteresting do-I-wake-or-do-I-sleep puzzle rather than a window into postmodernity."[28] As it is, epistemology remains in limbo: there are always other

[27] Ridley Scott, "Directing *Alien* and *Blade Runner*: An Interview with Ridley Scott," Peary 293-302.

[28] McNamara 435.

views of the world to be had, from which the facts will appear different. In tying the artifice of the filmic process (and its analogue in the Esper device) to the potential artifice within Deckard, then, the film takes the ontological concern of the microcosmic – whether Deckard is replicant or human – and broadens the ontological questioning to include itself. This is the example of the ontology-epistemology feedback loop I wrote about in chapter 4: the practical and local ontological issue about whether an agent is human or not spills into the more abstract epistemological question about how we could discern difference in principle, which in turn leads us to doubt the verifiability of reality in general, leaving the identity of the individual even less certain. In Dick's work the fact that a narrative exists in the first place, and that we stand outside it as sympathetic readers, allows us to break from this hermeneutic circle. However, it is this expansion so that the narrative is itself caught up in the culture of simulacra that it critiques which is the characteristic difference of *Blade Runner*. And that this is a more symptomatic quality of the postmodern condition is evidenced by one of the most prominent postmodernist cybernetic fictions, itself influenced by Scott's film, *The Matrix*.

Into Postmodernism

> Jean Baudrillard hailed Dick as expressing the idea of the simulacrum in his stories:
>
> > it is not about a parallel universe, a double universe, or even a possible universe – neither possible, impossible, neither real nor unreal: *hyperreal* – it is universe of simulation, which is something else altogether. And not because Dick speaks specifically of simulacra – science fiction has always done so, but it played on the double, on doubling or redoubling, either artificial or imaginary, whereas here the double has disappeared, there is no longer a double, one is always already in the other world, which is no longer an other, without a mirror, a projection, or a utopia that can reflect it.[29]

Whilst it is not surprising that Baudrillard alights upon Dick, contrary to Baudrillard's analysis I have argued that whilst (like much of Dick's oeuvre) *Do Androids Dream of Electric Sheep?* plays with the idea of the simulacrum, it does not concede that there is no ontology of difference. There is a stable ground from outside the simulation by which it is possible to know the original and real, and the mirror and unreal. The "precinct" episode reflects on the reader's exteriority: we are detached from the novel's world so that we can

---

[29] Baudrillard, *Simulacra* 125.

read it as a critical metaphor, able empathically (and therefore morally) to project our consciousness across from our reality to the distinct fictional one, even though fiction and the android are phenomenologically parallel as mechanisms calculated to deceive.

Whilst *Do Androids Dream of Electric Sheep?* is certainly more engaged with the postmodern than *2001* of the same year, jumping ahead two decades to *Blade Runner* highlights the degree to which both works cling to a humanistic centre, anticipating a body of readers and viewers able to make sense of a world of simulations through the discriminatory power of narrative. Scott's is a far more problematic work than Dick's, for not only is the urban world of *Blade Runner* full of Baudrillardian simulations, these are paradigms of the simulatory nature of the global fiction, the film itself, which contains them. Whilst the film is a critical response to the pressures of commercialism and civic surveillance, it also partakes of it in presenting proliferating images of eyes and doubles, and in refusing to allow entry into the multiple perspectives and identities which it teasingly suggests are available. A move to the other side of the mirror is denied, because film is itself shown as a self-replicating genre, with the possibility of alternative perspectives providing the best access to the "reality" always compromised by the directorial process. Nevertheless, the most successful aspect of the film, unanimously acknowledged even by reviewers who disliked it overall, is the realistic portrayal of the city of the future. If truth within the city is always hard to find, in its detailed rendering the city seems immediately and directly present to the viewer, its architectural reality substituting for the lack of psychological truth within the characters who inhabit it.[30] Even though the future is envisioned as a dark urban space, in contrast to *2001*'s immense prospect, like Kubrick's film visual realism in *Blade Runner* compensates for potential allegorical anarchy, grounding it to some degree in a mimetic mode which therefore implicitly believes that there is a reality to be represented in the first place. *The Matrix*, however, takes Baudrillard's assumptions completely into the logical hermeneutics of its visualisation. *The Matrix* takes the ontological complexity of Scott's work to a deeper level by drawing attention to the reality-warp of its computerised special effects. It thereby compromises our belief not only in the fiction but in Cartesian reality itself. Our viewing minds really can be deceived – for are we not taken in by the high-resolution realism of the film? – but there is no essential reality or truth which that deception conceals.

---

[30] The most infamously damning review which nevertheless admired *Blade Runner*'s visual effects was Pauline Kael, "Baby, the Rain Must Fall," rev. of *Blade Runner*, dir. Ridley Scott, *The New Yorker* 12 July 1982: 82-85.

Chapter 7. Deserting Reality: *The Matrix* as Postmodern Theory

As Brian McHale points out, science fiction generally is *the* ontological genre, since as a "literature of cognitive estrangement" (to use Darko Suvin's phrase), it confronts the empirical givens of our world with something not given, outside (or estranged) from it, in a way which demands our cognitive interpretation of any connections between the two. Science fiction thus "obeys the same underlying principles of ontological poetics as postmodernist fiction."[1] According to McHale's model in which epistemology and ontology slip into each other, our interpretation of the overlap between the worlds of the scientific present and possible futures, reality and fiction, ensures that these ontological collisions simultaneously pose more general epistemological questions about reality. Science fiction is always concerned to some degree with epistemology, asking how our assumptions about an imagined heterocosm relate to the real world, and how the nature of the present might inform the future. In the previous two chapters, we witnessed the way in which the epistemological issues raised in all three fictions swelled so as to include, by 1982, the practice of narrative itself. *Blade Runner* is epistemologically sceptical through an ontological collision of science and fiction: it presents a realistically extrapolated if quirkily retrofitted cityscape which nevertheless is potentially unrealistic at its heart, since reality is always already formulated from the position of an arbitrary observer. Because the film itself is ontologically questionable in making commensurate the future and the present, the replicant and the detective, the epistemological conditions of postmodern reality are also under question.

In *The Matrix*, we see this ontological conflict taken to the most extreme degree and given an ironically subversive twist: the possibility *The Matrix* presents is that reality itself is as fabricated as the fiction which is supposedly ontologically distinct from it. In *The Matrix*, the ontological supposition is dualist in the Cartesian sense: humans are bodily batteries serving the power needs of machines, with reality actually a computer-simulated construct, the "neural-interactive simulation" called the Matrix, designed to keep those bodies functioning without rebellion.[2] Rather than being a wholly fantastic or cognitively

---

[1] McHale, *Postmodernist Fiction* 60.

[2] In this chapter, the italicised *The Matrix* designates the film as an artefact, whilst the capitalised "the Matrix" denotes the program operating within it. Except where specified otherwise, all primary quotations from *The Matrix* are taken from *The Matrix*, dir. Andy Wachowski and Larry Wachowski, perf. Keanu Reeves, Laurence Fishburn, Carrie-Anne Moss and Hugo Weaving, 1999, DVD, Warner Brothers, 2006. The

estranged premise, however, this ontological dualism must lead to an epistemological challenge for the viewer, who has sat immobile through a similar construct – the film *The Matrix* – which relies heavily on computer-generated imagery and which, as a Hollywood blockbuster, exemplifies the culture of reproduction it critiques. To apply the demonic/daemonic distinction I elaborated in chapter 1, the demonic premise that we are possessed by a simulation is daemonically mediated by comparable technological methods. This thereby exemplifies Baudrillard's condition of the hyperreal. In this chapter, I will argue that at the stylistic level, *The Matrix* theorises postmodernism in a more successful description of the conditions of late modernity than academic critical theory, particularly Jean Baudrillard's *Simulacra and Simulation*, upon which *The Matrix* draws. Whilst the daemonic mechanism may be part of its irony, as an example of cyberpunk *The Matrix* performs according to McHale's understanding of the genre, managing "to actualize or literalize what in postmodernist poetics normally appears as a metaphor at the level of language, structure or the material medium." [3] Rather than simply figuring disintegration, the fictional worlds of cyberpunk include "phenomena embodying and illustrating the problematics of selfhood: human-machine symbiosis, artificial intelligence, biologically engineered alter egos, and so on." [4] Thus *The Matrix* is paradigmatic of cybernetic culture and postmodern theory, and a creative and innovative film as well as a popular

---

screenplay I use in some places is the shooting script. This was followed in its entirely in the production of the film, though portions were edited out post-production. Larry Wachowski and Andy Wachowski, *"The Matrix*: Shooting Script," *The Art of* The Matrix, ed. Spencer Lamm (London: Titan, 2000) 271-394. I will not be referring to the two subsequent films, *The Matrix Reloaded* (2003) and *The Matrix Revolutions* (2003) in this chapter. This is because, although the Wachowski brothers retrospectively claimed that *The Matrix* was conceived as part of a trilogy, there is no evidence that the later two films were ever more than germs of stories when the first was released. The Wachowskis had trouble placing *The Matrix* with a studio, and it was only after its somewhat unexpected success that the later films were rolled out in the same year. The first film is clearly a neatly contained narrative unit, albeit one in which the Matrix ambiguously continues to exist rather than being definitively destroyed; in contrast, the later two films are far weaker narratives. The latter also didactically explicate philosophical questions through the guise of pseudo-Socratic "Architect" character, whereas *The Matrix* allows these simply, but more effectively, to emerge through the plot and the initial concept.

[3] Brian McHale, "Elements of a Poetics of Cyberpunk" 150.

[4] Brian McHale, "Elements of a Poetics of Cyberpunk" 159.

blockbuster. Nevertheless, the postmodern element and the engagement with Baudrillard is not so much the consequence of the Wachowski brothers' desire to theorise postmodernism in any strong sense, but rather (and more interestingly) is the legacy of its generic relationship to cyberpunk and the fact that the computer had – by the late 1990s – become a readily accessible metaphor for the experience of simulation so that the postmodern somehow seems entirely natural and appropriate to this form of fiction. The postmodern has become such a dominant episteme, and cyberpunk such an intense genre, that it is hard to see how any film invoking computers could actually *avoid* engaging with the idea of the simulacrum. Indeed, Baudrillard's *Simulacra* derives from the early 1980s, and the two decades' lapse implies that this was not innovatively picked up by the Wachowski brothers, but more that retrospectively it provided appropriate theoretical endorsement for the film. Long theorised, the postmodern hyperreal had to wait until the rise of computing to give it a metaphorical form, firstly in the guise of literary cyberpunk, and then, as special effects advanced, in the cinematic medium. At this point, the metaphor could become self-sustaining, demonstrating the postmodern condition through resonating with the everyday experience of personal computers rather than via the didactic academic essay.

In a sense, the coincidental situatedness of *The Matrix* within postmodern culture is more striking. The problem faced by *The Matrix* and its viewers is essential to postmodernity. As Frederic Jameson theorises, if postmodernism is the product of late capitalism then "'culture' has become a product in its own right," and as such cultural fictions are unable to make any objective political critique of that moment.[5] This paradox was exemplified in the contrast between Dick's novel and Scott's film discussed previously. *Do Androids Dream of Electric Sheep?* finds grounds for humanism in empathy, but presents the novel as a non-commercial means by which this might be energised (in contrast to the Hollywood-style religion of Mercerism). But in *Blade Runner*, film itself becomes almost anti-humanist, since it contributes to and reflects the culture of duplication current throughout a city space which is loaded with the signifiers of oppressive civil and industrial power. Under conditions of this sort, efforts to describe a totalizing system produce an ironic result, as Jameson observes:

> What happens is that the more powerful the vision of some increasingly
> total system or logic...the more powerless the reader comes to feel. Insofar
> as the theorist wins, therefore, by constructing an increasingly closed and

---

[5] Jameson, *Postmodernism* x.

terrifying machine, to the very degree he loses, since the critical capacity of his work is thereby paralyzed, and the impulses of negation and revolt, not to speak of those of social transformation, are increasingly perceived as vain and trivial in the face of the model itself.[6]

What could better exemplify a "closed and terrifying machine" than *The Matrix*? According to Jameson, any attempt to tie historical elaboration to moral evaluation – as Marx achieved, for example – can now lead only into deeper irony, since the more one emphasises the historically-specific nature of the system, the harder it becomes to break from it from within that moment. Particularly if the viewer of this high-tech movie is also a user of computer games or the internet which seem to facilitate a dualist separation of mind from body in everyday experience – allowing the mind to wander the corridors of *System Shock* behind the screen whilst the body is concentrated on the few commands of the keyboard interface – he or she therefore appears logically capable of being deceived to the extent that the supposed fiction of one literal "terrifying machine," the Matrix, premises. This is thoroughly disorientating, treating reality and fiction not as separate but as commensurate ontologies, a nightmare from which it is impossible to awake because imagining and being are one and the same. Such a condition is a total demonic delusion. So total, in fact, that ironically there need be no reference to demons in the film in order to invoke this framework. The lack of an explicitly demonic reference is indeed the strongest proof that the deceit is so complete that it does not need any explicit metaphor to vouch for its existence; I will return to this ironic lack of a demon to signify the delusion at the end of this chapter.

    The Matrix is a postmodern work that exploits "the heterotopian space of the zone, where incommensurable spaces are juxtaposed or superimposed," a Borgesian or Baudrillardian parable of a map as large as the country it represents, so that the country becomes its own map.[7] Where postmodern science fiction differs from its earlier forms is that the revelation of the simulation does not necessarily return us to a higher level of reality from which we entered the fiction, just as postmodern theory contends that language cannot provide a transparent representation of reality because it is inherently self referential, referring us only to other linguistic forms through the process of *différance*. In the fake precinct episode of *Do Androids Dream of Electric Sheep?* we seem to become

---

[6] Jameson, *Postmodernism* 6.

[7] McHale, *Postmodernism* 163.

trapped beneath three layers of deception: looking through the eyes of an android, at a simulated police station inhabited by other androids, within a fictional novel. But the realisation as to the true nature of the first two levels – that Deckard is not an android and that the police station is actually a fake reproduction of that at which Deckard genuinely works – restores our confidence in the veracity of the narration within the last level. As it takes us out of the mirror image precinct and its duplicitous android-policemen, the novel becomes the signifier of its own truth that imaginative empathy denotes the human. On the other hand, in postmodern fictions like *The Matrix* we do not necessarily return through the deceptions to the level of reality from which we originally began watching. This is particularly true of film, which we perceive with less analytical distance than with literature which can be read at our own governed pace. As Douglas Hofstader observes in his study of recursivity in science and the arts, "When you pop out of a movie-within-a-movie you feel for a moment as if you had reached the real world, though you are still one level from the top."[8] This is true of the Esper device in *Blade Runner*, which draws attention to the editorial process of constructing truth so that we are left knowing we are still within a movie a level removed from reality, even once the sequence ends. More dramatically in *The Matrix*, when we are told that the reality that resembles our own world is in fact a simulation, we do not immediately recover to know *The Matrix* as a movie but simply and provocatively as one other possible manifestation of the Matrix. To use the film's own Carrollian analogy, to take the red pill that shows us the supposed truth is not to cure the symptoms of simulacra but only to go deeper into the rabbit hole. The existence of the Matrix simulation ensures that the fictional narrative crosses into the reality of the world beyond the film: the fact that I am captivated by *The Matrix* is proof of the existence of an economically imprisoning Matrix. It is only on certain levels, which I explore in the second part of this chapter, that we are encouraged to watch *The Matrix* knowing it as a movie within which is embedded a wholly-fictional Matrix that is comparatively a simulation as opposed to genuine reality.

Descartes and Plato

Firstly, though, it is worth emphasising the extent to which we are disorientatingly embedded within the alternative reality of the fiction, to a degree which problematises our ability to gain a critical angle on the reality of late capitalism and postmodern culture. This

[8] Douglas Hofstadter, *Gödel, Escher, Bach: The Eternal Golden Braid* (New York: Basic, 1979) 184.

effect can be measured by noting how far we are removed from the Cartesian deceit, in a way unlike the narratives of the previous chapters, and in spite of the fact that the film is premised on Cartesian dualism (that minds can occupy a different imaginative world to bodies lying in vats). As I suggested in chapter 4, the deceiving demon is ultimately productive because it constitutes an epistemological buffer zone between the real and the temporarily virtual. Because the theological preconditions of God's *potentia Dei ordinata* anticipate that demonic deceit cannot really be allowed, Descartes can choose to confront doubt as only a theoretical possibility, one which as a virtual imagining provides the foundations from which realism and rationality may be asserted. The fundamental difference between Descartes' sceptical doubt and that of *The Matrix* is that whilst Descartes chooses to enter that ground knowing that, in principle, he is always protected by theology, *The Matrix* is an imposed deceit that anticipates no such philosophical armour, and consequently no such distinction between fabrication and the world. As Michael Williams observes, "scepticism creeps up on the classical sceptic, as the result of his persistent inability to bring disputes to a definite conclusion," a remark that not only points to Descartes' break with classical epistemology, but which also hints that postmodernism is a revisiting of this sceptical aspect of classical thought.[9] As I argued in chapter 2, Haraway's cyborg attempts to recuperate the emotional and subjective quality of scientific knowledge, and it is no coincidence that it thereby echoes the pre-Enlightenment debate about monsters and demons. In *The Matrix*, likewise, postmodern thought bypasses the Cartesian model and revisits an earlier episteme.

In this case, the philosophical model of truth and deception upon which *The Matrix* is founded appears to be less Descartes' deceiving demon, and more Plato's allegory of the prisoners in the cave.[10] Whilst rationality is perceived by post-Enlightenment thinkers as the way of obtaining absolute and finite truth in a formal scientific system that cannot be manipulated by ideology, postmodernism is more sceptical, presenting truth as something socially constructed. This political context provides the impetus for a move we witness in both *Blade Runner* and *The Matrix*, as they attempt to represent not only the bleakness of a contemporary social reality – Los Angeles in Ridley Scott's vision and Chicago in *The*

[9] Michael Williams, "Descartes and the Metaphysics of Doubt," *Essays on Descartes' Meditations*, ed. Amélie Oksenberg Rorty (Berkley, CA: U of California P, 1986) 120.

[10] John Partridge, "Plato's Cave and *The Matrix*," *The Matrix Trilogy*, 20 March 2003, Warner Brothers, 5 March 2008 <http://whatisthematrix.warnerbros.com/rl_cmp/new_phil_partridge.html>.

Matrix – but also the impossibility of dissembling such a political context, since there are no conditions extraneous to an information-based late capitalism that are not themselves also representations or discourses. The concept of the deceiving demon seems appropriate when the issue is the ontological negotiation of the boundaries between self and other, man and machine, as in *2001* and of *Do Androids Dream of Electric Sheep?* However, by subsuming man as a product of a machine, and the representation of film as one more product of the system it seeks to deconstruct, *The Matrix* confronts a different problem. Through a Cartesian cogito we would have no basis of perceiving any such schism between reality and simulation at the level of macrocosmic ontology. The cogito is inherently solipsistic, such that only the ego thinking can know itself to exist for sure as the centre of his or her phenomenological universe. Cartesianism cannot therefore validate the nature of the Matrix, since although the cogito would indeed provide something true for that individual wired up to the simulation, the quality of that truth would *a priori* accord to the machine's terms as a high-fidelity simulation manipulating that person's subjective process. The machine would have allowed that thought to happen in the first place, giving the appearance of realities whilst still being in control of the way these simulations are presented.[11]

Whilst Descartes determines rationality by pulling himself up through his logical bootstraps in a meditative mode available to anyone capable of conscious reasoning, for Plato the individual must be goaded to confront reality rather than representation through outside intervention by those with superior knowledge; only through their continued efforts will the philosophers force the prisoners to turn away from the forms of shadows on the wall and see the light of the real.[12] This is more amenable to a socially, rather than a rationally, situated problem of the true and the false. Morpheus reveals to Neo the blasted landscape that is real Earth:

> MORPHEUS. The Matrix is everywhere, it's all around us, here even in this
> room. You can see it out your window, or on your television. You feel it

---

[11] I agree here with Lyle Zynda that *The Matrix* shows that science cannot provide certainty, and that *The Matrix* deconstructs the feasibility of the realist position. Lyle Zynda, "The Nature of Reality and Why it Matters," *Taking the Red Pill: Science, Philosophy and Religion in The Matrix*, ed. Glenn Yeffeth (Chichester: Summersdale, 2003) 43-55.

[12] Plato, *Republic*, trans. and introd. Robin Waterfield (Oxford: Oxford UP, 1993) 240-45.

> when you go to work, or go to church or pay your taxes. It is the world that
> has been pulled over your eyes to blind you from the truth.
> NEO. What truth?
> MORPHEUS. That you are a slave, Neo. That you, like everyone else, you
> were born into bondage, born into a prison that you cannot smell, taste, or
> touch. A prison for your mind.

Platonism speaks of the difficulty of attaining and transmitting knowledge, rather than implying that rationality is available to all who think (Descartes' terms). When the prisoner is forced to look into the fire, Plato speculates that it would "hurt his eyes," that he would turn back to the shadows he could make out, and that he would have to be "dragged forcibly away from there up the rough, steep slope," at which point the sun would cause him "pain and distress." Whereas Descartes argues for the grounds by which empirical realism can be attained by all, for Plato, the fact that there is a reality "out there" is not in doubt – just that reaching the real is impossible to follow without outside guidance by those already with privileged access to it, the philosopher kings who are capable of inculcating moral and epistemological certainty in their subjects through the dialectical method. As well as in the explicit reference to the "prisoner" passage, *The Matrix*'s contextualisation within a Platonic framework is performed visually. If Descartes' rhetorical mode is meditative, an aesthetic maintained by *2001*, the Platonic thought experiment relies on spatial metaphors: inside the cave and outside in the light, the prisoners situated between the fire and the screen. So in *The Matrix* the real world is marked as an underworld, as ships fly through the old Earthly sewer system, with Zion, the centre for the resistance, located near the Earth's core. Symbolically, deeper access to the truth is gained by literally descending beneath the layers of the Matrix simulation into labyrinthine caves and tunnels. After choosing the red pill which will force him to confront the deception, Neo falls through the mirror in the Matrix and wakes in a vat in the machine world (in a shot subversively reminiscent of the foetus in *2001*); he then plunges down a tube-like gut, being abjected from the machine as a non-believer.[13] Oddly, given the violence of the film generally, this is perhaps the most visceral and disturbing sequence, enacting a rhetoric of bodily scatology to induce queasiness in the viewer. This sequence

---

[13] Although not something I will apply in this chapter, it might be instructive to read *The Matrix*'s spatial conception of the machine/human dichotomy in terms of Lakoff and Johnson's work on embodied metaphor. George Lakoff and Mark Johnson, *Metaphors We Live By* (Chicago: U of Chicago P, 1980).

patterns the difficulty of confronting reality for Neo, who is, Morpheus notes, the oldest

person yet recovered from the machinic consciousness. Like Plato's prisoners, most are too

embedded in the simulation ever to recover. As Morpheus explains, "There is a rule that

we do not free a mind once it reaches a certain age. It is dangerous. They have trouble

letting go. Their mind turns against them." The Cartesian ego is versatile and egalitarian,

since if doubt is a fundamental condition of humanity, ironically it is the one thing

therefore that allows anyone to know themselves to be humanly conscious. In Plato's

metaphysics, though, one has instead to be goaded and pushed into confronting the world

in reality rather than its simulation or form. Like the Christian disciple or like Neo – whose

name is anagrammatic for the One – you have to be chosen and taught from those already

outside. The saviour narrative is pessimistic, in that truth from outside the system it is

brought to the individual within it, and inculcated through the Socratic method.

Where *The Matrix* breaks from Plato is that in a secular age no such dialectical

saviour may be forthcoming. In the presence of escapist means like cinema or drugs, the

reality of everyday life as represented by the uncannily familiar Matrix world of 1999 may

not be so bad or prison-like that we perceive the need to be educated out of it. As a

contemporary viewer, it is hard not to sympathise with Cypher, who negotiates with Agent

Smith to have himself reinserted into the Matrix, preferring to eat unreal but juicy steak

rather than the real but tasteless protein gruel eaten by the crew of the *Nebuchadnezzar*.

Because the world of the Matrix is representatively our own modern world, is indeed a

visually perfect mimesis of the present, rejecting all our previous assumptions may be

more difficult. At this point, one might also note *The Matrix*'s affinity with Plato's critique

of poetry as a representation that is bent to the demands of something or someone else,

intended from the outset to deceive and hence to be mistrusted as a way of teaching society

about the true and the real, and hence the moral. Plato argues that to dream is to mistake a

likeness for reality, and that just as something cannot be simultaneously real and unreal,

belief about the truth cannot be the same thing as knowledge of the truth. So spectators of

theatre or readers of poetry "can see a plurality of beautiful things but not beauty

itself…they only ever entertain beliefs, and do not *know* any of the things they believe."[14]

In the same way, Jameson argues that culture cannot provide any objective analysis of the

conditions of society, because it too is bound up with the process of simulation that is

inherent to late capitalism. It is not that there is no such thing as objective truth (as some

[14] Plato, *Republic* 196-202.

believe postmodernism to argue), just that obtaining it is always performed through subjective belief systems. Of course, Plato's argument should be qualified as being only a utopian ideal, not necessarily practicable in reality. The question then is whether a creative or didactic mode of educating the "prisoner" is best. Plato privileges the philosophical approach, with the result being that his society is oligarchic and class divided. In this line, postmodern theory can be similarly elitist, as its writing style appeals to a fairly small community of academics. Thus, though mass market narratives or representations offer a contradictory means of critiquing the simulation culture of which they form part, to an audience not grounded in Platonic, Cartesian or postmodern reasoning, the film might provide a better access to the truth by making those arguments accessible at the level of the visual and the metaphorical, particularly through the vernacular trope of the computer.

The Matrix as Postmodern Theory

The problem of obtaining a critical angle on the culture of simulacra, and provoking the confrontation with reality, is symptomatised by Baudrillard's *Simulacra*. In the film the hacker code Neo sells is concealed within a copy of this book; Morpheus welcomes Neo to Baudrillard's "desert of the real"; and the shooting script also included Morpheus instructing Neo that "You have been living inside a dreamworld, Neo. As in Baudrillard's vision, your whole life has been spent inside the map, not the territory."[15] For Baudrillard, as for Plato, it is hard to gain a perspective that breaks through the representation to the original event upon which judgement can be cast. Baudrillard's hyperreal system is a powerful example of Jameson's "increasingly closed and terrifying machine," the assertion that we suffer from the merging of a formal idea of representation with the reality of which it is mirror.

In Baudrillard's writing, this process is so extensive that it extends to manipulate reality itself so as to bend the laws of physics. In one example, Baudrillard argues that power stages its own murder to rediscover its legitimacy; hence "The Kennedys died because they incarnated something, whereas the new presidents are nothing but caricatures and fake film – curiously, Johnson, Nixon, Ford, all have this simian mug, the monkeys of power."[16] Here, in the interests of exposing the totality of the system, reality is not the physical realm in which abstract ideas (the desire to kill a president) are exercised; rather

[15] Wachowski and Wachowski, "*The Matrix*: Shooting Script" 310.

[16] Baudrillard, *Simulacra* 24.

the idea of political power actively modifies reality. Though both Lee Harvey Oswald and Samuel Byck (Nixon's would-be assassin) were both intending to kill, according to Baudrillard's logic Oswald's bullets reached their target not because he was a good shot, but because they swerved to the demands of power, whilst Nixon "was not even found worthy of dying at the hands of the most insignificant, chance, unbalanced person" (note the passive construction here, as if power itself, not a person, determines whether a president will be killed).[17] It would be disingenuous to take Baudrillard at face value, just as his assertion that the Gulf war never happened was not an empirical claim but a playful jibe at the way in which experience of the war was mediated through the live reporting by embedded journalists. With deliberate ludism, *Simulacra* is designed to stage its own entanglement within postmodernism by defying causal logic. But even if Baudrillard's ultimate point is right, and terrorism (successful or failed) can be used to enforce a leader's political status, his mode of analysis can offer no plausible alternatives; in the absence of pragmatic instruction, playful iconoclasm is the only alternative. Indeed, Istvan Csicsery-Ronay has illustrated that Baudrillard's work (and Donna Haraway's) are best understood as themselves science fictional "narratives producing and hyperbolizing the new immanence."[18] Baudrillard contended that "the good old imaginary of science fiction is dead," because the indeterminacy of the hyperreal has made reality stranger than the imagination.[19] However, given the problems with postmodern critical style, filmic science fictions like *The Matrix* do retain a critical capacity of sorts, and compared to theory they are better placed to didactically engage with their audiences. If postmodern criticism is itself realised through hyperbolic form rather than logic, then perhaps mainstream science fiction might provide a more optimum perspective *on* the present, without being directly exposed as being itself a product *of* that present. This is because – in a point Csiscerey-Ronay misses in his comparison between theory and fiction – science fiction has no ultimate pretensions to fact, though its fiction may be informed by them, whereas Baudrillard's theory is intended to make claims to knowledge of sorts, no matter how contingent. Thus in *Simulacra* Baudrillard's analysis of the cause of successfully targeted assassinations seems absurdist, in a way not true of the bullet-time effect of *The Matrix* –

[17] Baudrillard, *Simulacra* 24.

[18] Istvan Csicsery-Ronay, "The SF of Theory: Baudrillard and Haraway," *Science Fiction Studies* 18.3 (1991), *Science Fiction Studies*, 7 May 2008 <http://www.depauw.edu/SFs/backissues/55/icr55art.htm>.

[19] Baudrillard, *Simulacra* 121.

both, though, aim to draw attention to the way the usurpation of political power requires an act of violence. In science fiction, cognitively de-estranging the ontology of fiction from that of science redirects the hermeneutic work achieved by a critical theory. Rather than trying to decipher a text whose meaning cannot be finally determined (since it is precisely Baudrillard's point that truth is invisible) we decipher the aesthetic quality of *The Matrix* as a fictional representation of reality, regardless of whether that reality is itself really an overarching fiction or Matrix. Even if we live in the Matrix, *The Matrix* can still be known as a sub-fictional product of that overall fictional condition, a mise-en-abyme that is therefore not the original, broader simulation itself. Discerning the similarities and differences between these ontological zones, fictional-reality and fiction-in-fictional-reality, becomes a way of recovering from the potential that the world is always already a representational construct. Even if reality is solipsistically perceived, with no common truth out there, there are still scales, qualities and forms of simulation. Deciphering these in the genre of science fiction (or deciding the degree to which fiction is the reflection of truth) is both a critical, and hence Platonic activity, even if it does not (cannot) lead to final, Cartesian answers. However, though it cannot provide a critical view from nowhere on late capitalism, *The Matrix*'s conceit of the Matrix provides a view from somewhere within the logic of the simulated world.

The Matrix in The Matrix

If the late twentieth century is characterised by the advent of an information economy,[20] central to *The Matrix*'s critical intention is an elaboration and critique of the conditions of this period, using the metaphor of the computer. In *The Matrix*, the oppressor is a cybernetic system, the Matrix, and the incarnation of that system is a replicating Agent Smith, a grey-suited everyman bureaucrat. Agent Smith idealises the capitalist subject: "You work for a respectable corporation, you have a social security number, and you pay your taxes." Neo, who is called Thomas Anderson in this world, is a computer software engineer, who occupies a Dilbert-esque booth in his job for the multi-national CorTechs (note the homophone). At the start of the film, his autocratic boss (who not coincidentally resembles Agent Smith) condemns his erratic behaviour. The moral we are intended to draw is that Anderson is like a malfunctioning subroutine in a wider economic program. We recognise that we viewers are as anonymous as Anderson or Smith, metonymic suits

---

[20] James R. Beniger, *The Control Revolution: Technological and Economic Origins of the Information Society* (Cambridge, MA: Harvard UP, 1986).

working in offices and restoring ourselves through a brief trip to the cinema where we can escape vicariously through the superhero, just as Neo escapes the boredom of life through hacking into cyberspace.

At what point did control switch over, so that the machines that evidenced our scientific and economic success became superior, so that we are no longer consumers or users but consumed and used, human batteries from which the machines draw perpetual power? Smith explains that the Matrix was designed to emulate the year 1999, "the peak of your civilization. I say 'your civilization' because as soon as we started thinking for you, it really became our civilization, which is, of course, what this is all about." At the end of a twentieth century characterised by international conflict, the year 1999 should have represented the fin-de-siècle moment of humanist triumph at the end of history, according to Francis Fukuyama's thesis.[21] But this victory of late capitalism and liberal democracy means that the next reconstructive process will stem from within this condition rather than without – as Fukuyama later conceded, science would change what it means to be human so that the last man will be transmuted into something posthuman.[22] It is noticeable that with the demise of the nuclear fears of the Cold War, anxieties about technology have switched to the possibility of unpredictable "emergence" from scientific processes initially created for the purposes of improving the human condition: created nano-structures might spontaneously become self replicating "grey goo"; genetically modified organisms may exponentially self-replicate so as to completely replace natural ones in the biosphere; and, key to *The Matrix*, intelligence might emerge from large-scale networks not originally designed as conscious systems, such as the internet.[23] In this sense, the trope of emergence patterns both late capitalism, and cybernetics at the end of the century. In first-wave cybernetics after the Second World War, the computer was intended to model the human; as we saw in chapter 5, the benchmark for this period was the Turing test, but the epistemological problem of modelling consciousness necessitated a new approach, one which has most recently centred itself around emergent or self-evolving life. So if the

---

[21] Francis Fukuyama, *The End of History and the Last Man* (London: Penguin, 1992).

[22] Francis Fukuyama, *Our Posthuman Future: Consequences of the Biotechnology Revolution* (New York: Farrar, Straus and Giroux, 2002).

[23] An example of this switch from nuclear anxiety to fear of emergent intelligence and self-replicating systems can be found in Bill Joy's influential essay for *Wired* magazine, "Why the Future Doesn't Need Us," *Wired* 18.4 (2000) 20 May 2008 <http://www.wired.com/wired/archive/8.04/joy.html>.

Matrix is conscious, it has no centre, no Cartesian pineal gland, because intelligent control emerges in an aleatory way rather than from a functionally defined set of algorithms created by the human who provided the model (as is the case for Hal). This shift from top-down modelling towards emergence in the informational sciences correlates neatly with the cultural milieu of late capitalism. Rather than the capitalist economy producing products that make life easier and which is legitimated by an ideology that fetishises hard goods, the economy is instead one of information which controls people as a consequence of unseen and incomprehensible systemic processes.[24]

Cyberpunk or artificial intelligence narratives provide a powerful iconography of this drama of emergence, because they represent systems, matrices and webs that both bind and alleviate the condition of the human subject. If cyberspace is liberating – where do you want to go today? – it is so because it has no centre; it is a flow of disincarnate information without a central locus that can be indicted or modified when the system goes wrong. Particularly inert are fictions which are themselves the outcome of such a technological culture.[25] As Frederic Jameson complains of cyberpunk, having been initially designed to outrage humanists by representing the liberation of the underground in a cybernetic fantasy (*Neuromancer* is the classic example of this early reactionary mode), in contemporary culture where computers and the web have become everyday objects, cyberpunk amounts to little more than "mere 'realism' and an outright representation of the present," glorifying what society is anyway forced to endure.[26] Sustaining this critique of cyberpunk through analysis of *The Matrix*, Laura Bartlett and Thomas Byers note that "while postmodern subjectivity itself may at first seem strikingly radical" offering a radically new subjectivity freed from phallocentric ontologies, ultimately "it bears uncanny similarities to the structures of global capitalism."[27] The cyborg citizen – the common Smith or Anderson –

---

[24] As I write this at a time of the "credit crunch," and discover my savings – which I never actually possessed in hard coinage anyway – dripping away through the net of "Collateralised Debt Obligations," the power of the unseen information economy has become depressingly personal.

[25] One might at this point note that ironically, given its apparent anti-capitalist stance, the film worked hard for the money men, having currently earned nearly $500 million worldwide. "The Matrix," *Box Office Mojo,* 2007, Box Office Mojo, 19 March 2008 <http://www.boxofficemojo.com/movies/?id=matrix.htm>.

[26] Jameson, *Postmodernism* 286.

[27] Laura Bartlett and Thomas B. Byers, "Back to the Future: The Humanist 'Matrix'," *Cultural Critique* 53 (2003): 29.

is less free than ever, not necessarily because the system is innately immoral but because he or she has no way of knowing or defining its emergent nature through narrative analysis. He experiences only the individual nodes of it in everyday life, principally the desktop PC, and lacks the epistemological overview afforded by traditional fiction or theory. Nevertheless, by making its audience aware of the system as it connects to this everyday technological artefact, *The Matrix* can achieve something Baudrillard's work cannot: allowing us simultaneously to perceive the system but also to attain critical distance on it through a historiography of sorts. I would contest the claims of Jameson, Bartlett or Byers that narratives like *The Matrix* are critically inert because they are mimetic of postmodernity, rather than rebelling against it.

The computer is an interface, a prosthetic linking the user and the alternative realities the machine renders and transmits, but it has no fundamental reality itself, in the sense that it is impossible to dissect the machine and point to the place where the computer is. In the same way, it is impossible to point in a Marxist way to the economic structure underpinning late capitalism which consists rather of flows of information and distributed systems without centre or base. The most iconographic representations of the computer are not of the hardware, its physical form, but the software and the operating systems it runs: Windows, Apple, Microsoft Office. Rather than being defined and interpretable according to what it is, then, the computer is acknowledged by what it does, or by its transient effects. However, even as the computer thereby provides an apt metaphor for the systemic ephemerality of late capitalism, the film gives incarnate shape to the computer, making it a discrete and knowable object. By reflecting on a more concrete nature of the computer, we are also encouraged to reflect on the constitution of late capitalism. As suggested above, the Platonic epistemology is cued through Neo's corporeal plunge into the underworld, the movement of his body (and hence also his mind) from the form of the simulation to the real. Similarly, computer programs are represented as embodied and spatialised things in the film: a computer cookie, the unseen file that traces the history in a web browser, becomes a literal cookie biscuit baked by the Oracle which allows Neo to discern his future in relation to his past; a bug, a symptom of surreptitious social monitoring, is now exorcised from Neo's body as an insect-like implant; the passive subroutine (or "daemon") that becomes active only when specific conditions are met (such as a virus attack), is here given corporeal shape in the Agent Smith, who appears only when the rebels enter the Matrix. By translating hidden processes into distinct objects, the computer is made into a discrete artefact as well as a metaphor of information capitalism. Incarnated in physical things and bodies rather than existing as ideas, paranoia and codes, the Matrix can now be

first apprehended, and then destroyed, in this case primarily through Neo's kung-fu combat with Smith.

More generally, by denoting *The Matrix* as a fiction-within-an-overall fiction (the Matrix), a critical distance of sorts is produced: we recognise the Matrix is not commensurate with *The Matrix*, even though the latter may be one more simulation of the former. This distancing is initiated from the opening screens, as *The Matrix* visualises the simulation of the Matrix taking place through its most reduced level of machine code beneath the glossy interface, with the descending green lines more reminiscent of early computers and their command line operating systems than of the icons and windows of the Apple or Microsoft PC. Later, characters must jack in to *The Matrix* using dial telephones or phone booths, rather than modern mobile phones. The framing of *The Matrix* in this nostalgic way opens it as a distinct unreality, an imperfect twist on the otherwise perfect replication of 1999 life. The embodiment of computer programs as things like "cookies" or "bugs," and the use of antiquated technologies, serve to historicise the film in a way at odds with a simulation that is supposedly rendered on the fly, freed from historical contingency. Whilst time can run in any direction in the Matrix, in *The Matrix* we are reminded that human technological culture which has produced these goods has only progressed in one, cumulative direction. As a consequence, these artefacts are known as metaphors, rather than as further evidence of the systemic permeation of technology into our (cyborg) lives. This metaphorical quality of technology is confirmed by the Wachowski brothers' description of the opening sequence in anthropomorphic terms: "A blinking cursor pulses in the electric darkness like a heart coursing with phosphorous light, burning beneath the derma of black-neon glass."[28] The image of the beating heart provides an impression of the machine that is precisely metaphorical: the machine is *like* man, but is not man (echoing that moment in the precinct where Deckard is described in a way which attends to the artifice of language rather than of the android/human). Similarly, when he is resurrected by Trinity, Neo's now-improved performance in *The Matrix* enables him not only to dodge bullets in the simulation, but to know it as a simulation, seeing the fundamental lines of binary code that give rise to the virtual representation. With this ability, he can physically stop bullets, pluck them from mid-air and examine them as if they are museum artefacts, before causing them to drop harmlessly to the floor. Such a self-reflexivity, in which Neo sees through the program, also naturally draws the viewer's

---

[28] Wachowski and Wachowski, "*The Matrix*: Shooting Script" 273.

attention to the film's fictionality, and thus to the fact that we are being manipulated as an audience – the difference being that we can resist or critique the representation of *The Matrix*, even if we cannot find a point from which to establish with Cartesian certainty whether a Matrix does or does not exist behind our own hallucinated reality.

Linked to this, the suspension of empirical time through the bullet-time effect proves, by its ostentatious existence as a "special" effect uniquely presented by this movie, that time as we normally experience it is not a malleable, simulatory construct and therefore might be a fundamental ground for truth. Metaphor conjoins conventionally different things together by invoking a common property. But it also denotes the difficulty of describing things in themselves, particularly abstract concepts like time. Metaphor is required to highlight a similarity that is not otherwise obvious. The bullet time sequences which are, because of their "special effect" therefore known *as* metaphors, denote similarity but connote difference between reality and simulation, the Matrix, *The Matrix*, and the real world. Whilst Baudrillard argues (admittedly ironically) that bullets swerve when shot at Presidents depending on the image of power they want to present, *The Matrix* can perform absurd manipulations of physics and time, without becoming correspondingly absurd*ist*. Indeed, the success of its special effects is ultimately central to its aesthetic value: the benchmark of good science fiction is that it astonishes by making us believe that the so-called "effect" really did happen, because it looks so convincing. Thus at first glance it seems that the director has skilfully found some loophole in the laws of physics. In order to appreciate this aesthetic achievement of *The Matrix*, though, we must possess a preconceived hierarchy of idea and reality: that though ideas are effected in reality, it cannot bend to their will in contravention of the laws of physics. However, the alternative reality of a film requires that a director suspend our disbelief in the limitations of physics. Only when the politics can change the laws of physics to ensure the death of presidents will we live in the *literal* Baudrillardian hyperreal; whilst we can recognise how science fiction is manipulating the reality of physics in a way that breaches the boundaries of the naturally possible, we also recognise that we are not ourselves watching in a world of the hyperreal where physics succumbs to ideology in every day life. Though we may be susceptible to demonic illusion, the cognitive estrangement between science fiction and physical reality preserves an epistemology of sorts by which we can become aware of the status of representations as opposed to actuality.

This is not to argue against *The Matrix* as a postmodern film. Although the metafictionality of *The Matrix* draws attention to its own artifice and therefore provides a contrast with reality as something not constructed but already out there, this still coheres

with postmodern theory. Baudrillard does not suggest that the notion of the hyperreal simulacrum without an original on which it is based makes comparison between the real and the signs of the real impossible. Disneyland, for example, "is presented as imaginary in order to make us believe that the rest is real." Disneyland acts as a "deterrence machine" that allows Los Angeles to "rejuvenate the fiction of the real" by way of contrast with the former.[29] On this comparative model, as David Levery reads *The Matrix*, "The real world exists, even under the reign of Baudrillard's 'Third Order of Simulacra' and cinematic art...can represent it and tell a heroic tale of its recovery."[30] However, is the recovery as complete as Levery argues, or are we still "one level from the top" in the recursion observed by Hofstader? The presupposition of *The Matrix* is that we have no way through analysis to break from our sense of the reality of, and our being subject to, late capitalism. *The Matrix* does not challenge the plausibility of its own Baudrillardian presupposition, only allows some room for manoeuvre *within* that supposition in a way that theoretical discourse cannot. By the film's logic – and the logic of being absorbed by a simulation cinema – we could be solipsistically in a Matrix; there is nothing in the film that deconstructs this global ontological logic, and indeed the existence of a film which uses computer graphics to deceive and seduce its audience adds credence to its manifesto. However, even if we are empirically in a simulation, the Matrix is still constructed in such a way as to operate according to particular physical and cognitive laws, such that any simulation embedded within it (the film *The Matrix*, for example) can be known as a deeper level of simulation that is not proximate with the reality as normally experienced. That is to say, *The Matrix* can be known as science fiction that can do things as fiction (such as suspending time) that cannot happen in "reality," even if "nature" is itself something produced within a malign, dissimulating program, a demonic *a priori* that cannot be deconstructed through rationality. By contrast, Baudrillard's theoretical discourse risks being rejected because it is so antithetical to our tacit intuition that bullets do not fly at the whim of ideas. Bartlett and Byers propose that the postmodern character of *The Matrix* is superficial:

> As discourses of postmodern subjectivity are appropriated by the popular
> media for the production of a contemporary style, the subject may exhibit a

---

[29] Baudrillard, *Simulacra* 12-13.

[30] David Levery, "From Cinespace to Cyberspace: Zionists and Agents, Realists and Gamers in *The Matrix* and *eXistenZ*," *The Journal of Popular Film and Television* 28 (2001): 150-58.

> sexy patina of postmodernism whilst still not differing in any fundamental
> way from its liberal humanist predecessor.[31]

However, as Frederic Jameson identified, pastiche is the dominant mode of postmodernist fiction, and a patina might instead be read as a laminate of multiple layers, with no essential characteristic – in this case, the real – waiting beneath them.[32] Although the film does make some space for critical interpretation, it does so in a way congruent with postmodernism because it occurs as a function of style. Given that postmodernism argues for the surface of a representational reality beneath which there is no deeper truth, the deployment of heterogeneous styles and carnivalesque conjunctions of different discourses is the only thing that can remain. As Brian McHale observes, heteroglossia offers "a means of breaking up the projected world into a polyphony of worlds of discourse."[33] The multiple but unsystematic allusions in the film – from Baudrillard to Buddhism via Lewis Carroll and *The Wizard of Oz* – certainly ensure that as a critical text it does not attain any privileged final truth. However, through invoking nostalgia and incarnating the computer in physical things, we can perceive *The Matrix* as a function of style that is not therefore identical to the effervescent, intangible and emergent mode of the late capitalist Matrix system.

The Moral Matrix

Even if they are temporarily captivated by the soundtrack and special effects vision of transcending reality via cyberspace, *The Matrix* hopes its viewers will still be able to understand it as a film within a frame, even if this frame too is a simulation produced by the Matrix. Through historical retrospection, the visualisation of the computer as incarnate metaphor rather than as effervescent system and the self-consciousness of its aesthetic warping of physical time and reality, *The Matrix* argues that even if this is one more simulation in a whole deceptive universe constellated with them – films, computer games, the internet, postmodern theory – it is still a special subset of the reality we know, such that reality is not monochromatically coterminous with simulations of which it is one example. In this respect, the fact that so many people did engage with the philosophical issues of *The*

---

[31] Bartlett and Byers 29.

[32] Jameson, *Postmodernism* 19.

[33] McHale, *Postmodernist Fiction* 166.

Matrix in a serious way indicates its value as a medium for thinking about simulation, even whilst providing just one more representation without epistemological foundation.[34]

Ironically, perhaps the ultimate evidence of the success of *The Matrix* is the failure of its own call to the fight against late capitalism. If the Matrix is just a simulation, then theoretically its reality can be subverted by anyone with an open mind. In practice, however, even though we may take vicarious pleasure in Neo's kung-fu battles against the system, it is nevertheless not an act we will similarly perform in reality, since those actions are conventions only in fiction-within-the-fiction, not social reality so far as we can know and experience it. Rather than being a pessimistic conclusion about the incapacity of the individual to fight late capitalism and cybernetic systems that drive it, this draws attention to our ability to maintain moral conventions regardless of the ontology within which they are constructed. In *Do Androids Dream of Electric Sheep?*, when Mercerism is revealed as a sham, an imaginative construct or projection, this persuades that the human ability to proceed as if Mercer exists distinguishes them even more readily from androids. *The Matrix*, too, argues that the fact that we could be deceived and living in a simulation makes it all the more important that we are capable of proceeding *as if* reality does exist with all the moral imperatives that entails, rather than living as if fiction and reality are ontologically the same. Indeed, if everything is virtual, then at least ethics can be consistent, because we all inhabit that same ontology. We need to distinguish carefully between the theoretical possibility of there being a Matrix and the pragmatic consequences for reality so far as we can know it. Kant may offer one way of situating our moral reactions to the narrative. According to the Kantian framework of the categorical imperative, regardless of whether the Matrix exists in our own lives and reality is just a simulation, in practice ethical values do pertain in the world so far as we can know it and so far as we can know other people's knowledge of it.[35] What is intrinsically good in and of itself is the ability to make rational choices; what determines the good life is not the

---

[34] The two most substantial academic books on the subject are Glenn Yeffeth, ed. *Taking the Red Pill: Science, Philosophy and Religion in The Matrix* (Chichester: Summersdale, 2003) and William Irwin, ed. *The Matrix and Philosophy: Welcome to the Desert of the Real* (Chicago: Open Court, 2002). The official website also includes essays from philosophers, such as David Chalmers, and cyberneticians, such as Kevin Warwick. *The Matrix Trilogy*, 2005, Warner Brothers, 15 March 2008 <http://whatisthematrix.warnerbros.com/>.

[35] Immanuel Kant, *Foundations of the Metaphysics of Morals, and What Is Enlightenment?*, trans. and introd. Lewis White Beck (Upper Saddle River, NJ: Prentice-Hall, 1997).

specific kinds of experiences, but the kind of actions deriving from them, however reliable those experiences may be.[36] Thus it is perfectly acceptable for the characters in the film to practice kung-fu violence since violence in the Matrix – as with violence in the computer game – does not hurt any autonomous, "conscious" agent in any normative sense of the word. But it is not acceptable for viewers of the film to leave the cinema in firm belief of the premise of deception, and to perform the same violent acts on human beings. Although in the film the CIA-style agents who quash the rebellion theoretically represent civic authority as an extension of the Matrix, our willingness to conform to such a judicial system in practice indicates our preference for inhabiting a structured human society even at the expense of some individual autonomy. The ontological potential exposed by *The Matrix* is that we are actually solipsistic brains in vats. This theoretically leaves moral concern for others unimportant, because they are anyway just fragments of code rather than genuine bodies that must be respected, according to the principles of liberal humanism. However, since contemporary news reports did not indicate mass increases in violence following *The Matrix*'s release, it seems we are more capable of discerning the distinction between reality and simulation than *The Matrix* itself presupposes. Even if we are brains in a vat controlled by a Matrix, kung-fu fighting is not an ethical way of raging against the machine. P. Chad Barnett rightly celebrates the film's "accurate cognitive mapping of multinational capital." But this does not necessarily allow us to "regain a capacity to act and struggle."[37] For the fact that we will not act in the aggressive way Neo does, even though the existence of *The Matrix* supposes we are living in a Matrix inhabited by informational minds, marks our faith in humanism even in a world of pure simulation and demonic deception.

---

[36] Note that this is not unlike Buddhist philosophy, according to which it is not specifically actions themselves which accord moral value to the person who performs them, but rather then intentions behind the actions. There are no negative karmic consequences if one does a negative act if performed with the carefully cultivated consciousness Buddhism instructs, nor if an unintentional consequence results from a rightly pursued state of mind. For a reading of *The Matrix* in terms of Buddhism, see James L. Ford, "Buddhism, Mythology, and *The Matrix*," Yeffeth 150-73.

[37] P. Chad. Barnett, "Reviving Cyberpunk: (Re)constructing the Subject and Mapping Cyberspace in the Wachowski Brothers' Film *The Matrix*," *Extrapolation* 41 (2000): 372.

Postscript: Return of the Demon

If I appear to have largely avoided making explicit reference to the demonic in considering *The Matrix*, as I mentioned at the start of this chapter this actually makes a powerful point about the futility of the allegorical or metaphorical method in a world which only ever exists as an illusory representation. The possession by the Matrix – "The Matrix has you" announced the film posters – is so total as to make any expression of this through the guise of a literal demon seem redundant, just another figuration within what is always already a total deception produced through representational signifiers. It is taken for granted that *The Matrix* is a demonic text because the film corresponds to the Matrix, a hermetic hallucination at the global level within which the fictional narrative is included. The demonic metaphor lends itself ideally to the subject matter of *The Matrix*. As we saw in chapter 3 in relation to psychoanalysis, demonic logic is attuned with our folk attitudes towards inhabiting a world that never seems quite graspable, stable or comprehensible. Just as the Matrix emerged from simple machines, demons can start out as deliberate hoaxes under human control and then become self-sustaining and self-propagating, as was the case with the demons of Loudun. Ironically and darkly, as that case shows, effects can generate their own origins in the case of the demonic. And, as I showed by putting Baudrillard up against the Loudun example in chapter 1, demonic fabrications are at the heart of the postmodern crisis of identity and truth with which *The Matrix* engages, via Baudrillard himself. So if demons had been used to describe Anderson's comparable possession by the simulation, this would hardly have been inappropriate.

On the other hand, presenting them on screen would have given life and incarnate form to beings that by their very nature are indeterminate, never objectively in the world because they can only exist through taking over the mind, the subjectivity, of another. In this sense, the externalisation or making explicit of a demonic trope would have countered *The Matrix*'s intention to show the impossibility of objectivity, of attaining a view from nowhere on the system within which we are always already immanent. To project a demon on screen is contradictory, because it is by its nature a product of consciousness (though *Forbidden Planet* gets around this problem by the employment of the mind manifestation device); the demon on film can generally only be an implied agent or mode of thinking from a modified other perspective, the sort of perspective induced by Kubrick's switching camera angle in *2001*. By contrast, the undeniable existence of the word "demon" in text, and the description of the consciousness within the context of which it appears, forces us to project our own interpretations of that term in the process of the text's signification. Even if we are secular or sceptical readers, the word demon is still impressed on the page and

produces meanings and associations in our minds. Where literary fiction has a capacity film does not is in our awareness that we are not only drawing meaning from it, but actively creating or concretising meaning in a reciprocal process. This is because the literary form can play with indeterminacy between objective and subjective: its subjective representations of the inner consciousness of a character are nevertheless amenable to our objective act of interpreting the meaning of those printed words. The demon does occur in the cybernetic fictions that are the subject of my final three chapters, which broadly consider the nature of allegory: if the problem of cybernetics is the closing of a gap between the world and the self, allegory makes us aware of a gap between the world and the book which seeks to explain it.

In this context, as Philip K. Dick's Isidore might argue, demons cannot be fake, unless reality is a fake; a shared subjective sensation, such as a story, must become an objective truth out there. In the works of Umberto Eco and A.S. Byatt, demons reflect on the nature of reality and its alternates (the novel, the artificial, the machinic) in ways that suggest cybernetics can provide a moderating field between pure (but naïve) scientific realism and postmodern fiction, relativism and social constructivism. Does the demon stand for something in a metaphysical scheme of meaning, or must the demon be rejected as a total figment of the mind in the secular psychoanalytic era? Is such a rejection not problematic when demons have been created in the mind of a reader simply through the signifying action of mere words, shapes on a page? If the demon can be said to exist in the human mind, even if the paranormal is not at work in physical reality, at what cognitive or neurological level can the demon be said to exist? In literary fictions, cybernetics influences the work at a more implicit level than in cinematic science fiction, though it nevertheless provides an analogy with the process of making or machining fiction. The demon, by contrast with film, becomes a more explicit marker of this process of the work done by fiction in the mind of the reader.

Chapter 8. Interpreting Machines: Umberto Eco's *Foucault's Pendulum*

"A text is a machine conceived for eliciting interpretations."[1]

"And so the author vanishes – that spoiled child of ignorance – to give place to a more thoughtful person who will know that the author is a machine, and will know how this machine works."[2]

Allegory and Conspiracy

Deborah Madsen contends that "allegory flourishes at times of intense cultural disruption, when the most authoritative texts of the culture are subject to revaluation and reassessment."[3] Conventionally understood as a narrative system that stabilises the value systems of the culture in which it is produced, allegory is actually fundamentally unstable because the need for allegory requires the admission that the thing to which it is referenced is coded or concealed. As I have been exploring, the advent of cybernetics and the changing nature of subjectivity and consciousness in relation to it renders both the objective reality of the world, and the representational capacity of cultural discourses, problematic. As *The Matrix* presents it in one of the most compelling iterations of postmodernism and cybernetics, allegory fails because culture has itself become a product in late capitalism, such that there can no longer be any critical or analytic gap between the representation and the world being represented. In relation to literature, this dynamic is well expressed by Gabriel Josipovici, who argues that the Christian tradition of allegory (whose paradigm is in Dante) is over, and the desire to allegorise has become demonic, a signal of our solipsistic failure to draw common meaning from the world. Nevertheless, this is not to suggest that allegory has no function in postmodernity, and perhaps particularly in the literary medium. Ironically, allegory may be revitalised precisely as it undergoes the most strenuous reassessment. This might explain the relationship Heidi Hansson observes between Christian allegory and the postmodern novel, genres which, at first glance, appear diametrically opposed. Hansson argues that allegory is an example of

[1] Umberto Eco, "Reading My Readers," *Modern Language Notes* 107.5 (1992): 820.

[2] Calvino, "Cybernetics and Ghosts" 16.

[3] Deborah L. Madsen, *Rereading Allegory: A Narrative Approach to Genre* (Basingstoke: Macmillan, 1995) 135.

"double discourse"; like postmodern literature, allegory "avoids establishing a centre for the text, because in allegory the unity of the work is provided by something that is not explicitly there."[4] The difference is that whilst Christian allegorists could assume the "existence of a universally accepted frame of reference outside the text," in postmodern fiction no such index exists. Umberto Eco extends the comparison between the postmodern and the allegorical, observing that in the case of Christianity "one assumes that symbols do have a final meaning, but since it is the same everlasting message, there is an inexhaustible variety of signifiers for a unique signified"; comparably, in poststructuralism "symbols have any possible meaning because of the inner contradictions of reality, but since every symbol speaks about this fundamental contradictoriness, an inexhaustible quantity of signifiers always stand for their unique signified, the inexhaustibility of the senses of the text."[5] Ironically, though radically different to religious allegory, poststructuralism and postmodernism share with it a semiotic practice.

Clearly, however, they are stylistically divergent. This is indicated by the contemporary predominance of allegories of conspiracy and the remarkable tone of paranoia that pervades them. For Jameson, "conspiracy theory (and its garish narrative manifestations) must be seen as a degraded attempt – through the figuration of advanced technology – to think the impossibility of the contemporary world system." Such "cognitive maps" of the late capitalist condition are "best observed in a whole mode of contemporary entertainment literature – 'high tech paranoia' – in which the circuits and networks of some putative global computer hookup are narratively mobilized by labyrinthine conspiracies."[6] Certainly, conspiracy seems prevalent in modern narrative and history: one thinks of the "faked" moon landings, the Kennedy assassination, the September 11th attacks, the popularity of fictions such as Dan Brown's *Da Vinci Code* and, of course, *The Matrix*. These phenomena all indicate that, whilst Christian allegory on its own may no longer provide incontrovertible connections between art and the universe, we nevertheless have an innate fascination (perhaps even need) for a metanarrative that

[4] Heidi Hansson, "The Double Voice of Metaphor: A. S. Byatt's 'Morpho Eugenia'," *Twentieth-Century Literature* 45 (1999): 454.

[5] Umberto Eco, *The Limits of Interpretation* (Bloomington, IN: Indiana UP, 1990) 20.

[6] Jameson, *Postmodernism* 37-38.

makes sense of all other speculative narratives whose truth claims cannot be verified personally.[7] As the rationalist Lia observes in *Foucault's Pendulum*:

> Synarchy is God...Mankind can't endure the thought that the world was born by chance, by mistake, just because four brainless atoms bumped into one another on a slippery highway. So a cosmic plot has to be found – God, angels, devils. Synarchy performs the same function on a lesser scale.[8]

Modern communications and information technologies provide conspiratorial writers with the opportunity to broaden the scope of their secrets on a global rather than a divine scale.

However, they can also be historicised. One major point of connection, which I will explore in relation to *Possession*, is to the Victorian period. Another contact point – and the most important for Eco – is the middle ages. Works such as the *Da Vinci Code*, or Eco's own *The Name of the Rose*, locate the groups which work to undermine society across the liminal period of the long scientific, political and religious revolutions. Similarly, in *Foucault's Pendulum* a modern-day expert on the Knights Templars, Casaubon, and two philologists working for a vanity press, Belbo and Diotallevi, are told about a coded message reporting the existence of a scheme initiated by the Templars. Initially for fun, the three use a modern computer in conjunction with Hermetic and Cabalic hermeneutics to tie together a range of popular and historical references into a single, synarchic Plan. Casaubon – named after the seventeenth-century Hermetic scholar Isaac Casaubon and the classicist from Middlemarch – argues that "Concepts are connected by analogy. There is no way to decide at once whether an analogy is good or bad, because to some degree everything is connected to everything else" (618). The concoction develops, however, to the extent that the academics start to believe their own story, whilst the modern descendents of the Templars, the Diabolicals, begin to chase down the trio in order to extract the information they have "discovered." Where Eco's conspiracies differ from other paranoid cybernetic fictions is in their historical emphasis. Cybernetics, postmodernism, and paranoia are all tied to the past in the form of Hermeticism, Christian allegory, and Cabala; whilst many conspiracies are spatial or international – relating an event in one part of the world to that in another – the conspiracy

[7] As observed by Archbishop Rowan Williams, in an Easter sermon which offered a Christian response to the culture of conspiracy. Denis Campbell, "*Da Vinci Code* Appeal Lies in Our Fascination with Spin, Says Williams," *The Observer* 16 April 2006, 18 Sept. 2007 <http://www.guardian.co.uk/uk/2006/apr/16/books.religion>.

[8] Umberto Eco, *Foucault's Pendulum* (London: Picador-Pan, 1989) 318. All subsequent parenthetical references in this chapter refer to this edition.

of *Foucault's Pendulum* is also a temporal one. The novel exemplifies Eco's belief that the ramshackle structures of occult philosophy provide a parallel with contemporary conspiracy:

> Antiscientific by definition, these Middle Ages keep going under the banner of the mystical weddings of the micro- with the macrocosm, and as a result they convince their adepts that everything is the same as anything else and that the whole world is born to convey, in any of its aspects and events, the same Message. Fortunately the message got lost, which makes its Quest fascinating for the happy few who stand proof-tight, philology resistant, bravely ignorant of the Popperian call for the good habit of falsification.[9]

It is one of the quirks of Enlightenment that even as it resurrected Ptolemy and Galen, early-modern scientists also found renewed veneration for Hermetic magic and astrology. Chapters 4 to 7 showed that as cybernetics provides new technologies of simulation and artificiality, there is also an increasing doubt that reality can be known through rational methods, such as Popper's. However, rather than therefore recovering the antiscientific methods of the middle ages, and excavating the demonic belief systems that potentially underlie Enlightenment practices, *Foucault's Pendulum* ultimately argues against total relativism in practices of reading and means of producing knowledge. David Porush writes:

> To the paranoid, like the cyberneticist, everything is message and nothing is neutral; everything has information waiting to be recuperated by the correct reading or the application of the right technique…every physical channel has a certain quantifiable capacity to carry information – with a vengeance.[10]

If cybernetics and paranoia go together (and a film like *The Matrix* certainly seems to validate this claim), one should also observe that the idea of the code also crosses the two: conspirators communicate via secret messages; the cybernetic transcendentalists seek the code of human consciousness that can be reprogrammed or re-coded on a machine; readers of allegorical fictions suspect that the surfaces of the text can be deciphered to point to some higher meaning outside it. In all cases, only one code combination will lead to the correct decipherment. Both working within and against postmodern semiotics, Eco

---

[9] Umberto Eco, *Travels in Hyperreality*, trans. William Weaver (London: Picador-Pan, 1987) 71, rpt. of *Faith in Fakes*, 1986.

[10] Porush 108.

anticipates readers of his novel to be right-minded realists who will refuse to look for the code of the Diabolical conspirators, with their supernatural delusions, and choose instead a more stable code by which to read the book. Before going in to the stylistic techniques Eco uses to steer his readers down this more rational path, however, it is worth noting how this might be symptomatic of a broader shift of emphasis in the relationship between author, text, and interpreter. In the allegorical tradition and in the cult of inspiration (which I explored in chapter 1), the writer was perceived as the medium between some effervescent, allegorical spectrum of signifieds and the hard signifiers on the page. However, the advent of cybernetic systems like TALE SPIN confirm the postmodern suspicion of such institutional authority in the author and the special access he has to an allegorical index. In Eco's novel, then, the expectations placed on the reader to decipher of the right codes implies that the author as conceptualised by postmodern and cybernetic epistemologies no longer has control over his own text.

Writing Machines

Underlying the novel is the tradition of the *Corpus Hermeticum*, with its "refusal of causalism" and "logic of "mutually sympathetic elements" which seeks similarity between disparate ideas and events.[11] This medieval practice is a symbolic ancestor to cybernetic systems, as the Hermetic *ars combinatoria* "foreshadow[s] our modern electronic brains" (472-74). Working with the tradition, Jesuits such as Athanasius Kircher "built mechanical devices, mechanisms with perforated cards, computers ante litteram. Binary calculators. Cabala applied to modern technology" (474). Cabala, another practice that underpins the mechanics of the plot, was a Jewish hermeneutic for reading Biblical allegory that "possessed a warrant to explore ever-proliferating dimensions of meaning."[12] So the modern literary scholars rely on the principle that "No piece of information is superior to any other. Power lies in having them all on file and then finding the connections. There are always connections; you only have to want to find them" (225). They construct a computer program called Abulafia (after a Spaniard, Abraham Abulafia, a fourteenth century authority on Cabala) into which disparate pieces of information are fed; through its automated reorganisation of data, connections between them are generated at random. The academics then interpret this output literally as corresponding to some overall, logical Plan

[11] Eco, *Limits* 19.

[12] Madsen 38.

(373-6). As a result, it appears as if the plot the conspirators seek to concoct and to unravel is not plotted by any authorial narrative agent, Eco himself, but rather by a machine. Writing a plot in both senses is rather an automated processes of reordering and collation.

This representation of writing as a mechanistic process is not, in itself, a new concept. As well as Cabala or Hermeticism, one thinks immediately of Swift's frame machine in Lagado in *Gulliver's Travels*. But, as with Descartes' admission that a mechanical mind would theoretically be possessed of a soul but pragmatically could not be built, the stakes are raised dramatically when a myth suddenly becomes a determinate present embodied in technology. In "Cybernetics and Ghosts," an essay which appears to influence *Foucault's Pendulum*, if not explicitly, Italo Calvino proposes that:

> The so-called personality of the writer exists within the very act of writing: it is the product and the instrument of the writing process. A writing machine that has been fed an instruction appropriate to the case could also devise an exact and unmistakable "personality" of an author, or else it could be adjusted in such a way as to evolve or change "personality" with each work it composes. Writers, as they have always been up to now, are already writing machines; or at least they are when things are going well.[13]

As N. Katherine Hayles identifies, Calvino's work, particularly *If On a Winter's Night a Traveller*, fears the dissolution of the body of text that parallels the abstraction of embodied information under the cybernetic paradigm, where a text can exist simultaneously on multiple media.[14] Calvino's novel strains to keep the literary corpus intact through the recursive chapters, and to maintain a single narrative thread; its nemesis is Lotaria's reading machine in which the text's meanings are analysed through quantitative analysis of word frequencies in a way that abstracts information from the linear medium.[15] Although Calvino's essay and novel draw on twentieth-century cybernetics for their metaphors, Calvino also acknowledges that electronic brains have made a reality of the most arduous effort of the Middle Ages, *ars combinatoria*.[16] *Foucault's Pendulum* explores this avenue in more depth, linking Hermeticism to

---

[13] Calvino, "Cybernetics and Ghosts" 15.

[14] Hayles, *Posthuman* 41-3.

[15] Italo Calvino, *If On a Winter's Night a Traveller* (London: David Campbell-Everyman's Library, 1993), especially at 183-84.

[16] Calvino, "Cybernetics and Ghosts" 9.

cybernetics through Abulafia. The books of Hermes, though physically written by a scribe, were considered authored by the divine agent who inspired the human; the human writer is a mere interface or medium for the original creative source. Similarly, cybernetics for Calvino and Eco are represented as passive mediations, whereby the author is simply the interface which enables the word to materialise on the page. This, then, offers a technologised variant on the traditional theory of literary inspiration as a passive experience in which the writer is the receptacle for ideas and representations that originate from somewhere outside of him. The aesthetic benefit of this myth of creativity is that any text which is inspired is also said to be inspir*ing*; kudos accrues to an artist not because of his original creativity, but because his created output reaches out to and affects readers in a comparably strange and immaterial process. As Timothy Clark puts it "inspiration seems to be an experience that is itself testimony to the value of its products, a performative that ensures its own value."[17]

In his essay Calvino dissolves the author as nothing more than a symbol-manipulating machine, but preserves space for the readerly spirit which will constantly renew interpretations of a text in non-deterministic ways. The image of the machine signifies the transfer of the power of allegory away from the poetic genius and onto the reader. As Madsen observes:

> In place of such a figural subject, fabulistic allegory draws upon the supplementarity of a reader who supplies conventional signifieds for indeterminate metaphorical signifiers – allegory is "a mode of reading the already written" as the individual genius valued by Romanticism gives way to the culturally constituted discursive subject prized by poststructuralism.[18]

As a semiotician, Eco argues that whilst according to deconstruction there are no rules which "help to ascertain which interpretations are the 'best ones,' there is at least a rule for ascertaining which ones are 'bad'."[19] Even if a text is theoretically open to the free play of signifiers, a text limits interpretations by highlighting or framing those readings which it considers to be poor. Where Casaubon and the conspirators fail is not in producing a plot through mechanistic means – for without the conspiracy plot there would be no novel, whilst the author himself is essentially a writing machine – but in failing to read sceptically

[17] T. Clark 6.

[18] Madsen 123.

[19] Eco, *Limits* 60.

in accordance with broader social expectations and conventions. Embodying both seductive sophistry and allegory that might reveal a genuine truth about the world, *Foucault's Pendulum* evidences that though texts may be machines for generating interpretations, readers are not interpreting machines. If Philip K. Dick saw readers as capable of discriminating between human and machine by extending empathy even to the simulation, Eco's novel puts pressure on the reader for a different reason, staging a postmodern moment when the author is dead as the source of creativity, substitutable by a mere computer, and consequently it stresses the importance of interpretation. If the text is the product of a machine (indeed is itself an example of a machine), then questions of value are transposed onto the reader. The reader may always potentially fall into the same demonic frenzy of signification conveyed to them by the text. On the other hand, if we do not become possessed by *Foucault's Pendulum* in the same way as the Diabolical authors are possessed by the meanings produced by Abulafia, a hermeneutic gap opens up between the text as inspired by a machine (whether a literal computer or a computer-like author) and how we actually respond to it as non-deterministic interpreters. It is this gap *Foucault's Pendulum* creates through its narrative strategies, ensuring that the reader must perceive literature as something that neither be reduced to atomism or mechanism, but that does not therefore have to lead to infinite proliferations of intertextuality and signification.

Diabolical Readers

The Diabolicals believe that by deciphering "Another subtle ambiguity, another luminous mystical analogy" (142), they will eventually reach the ultimately meaningful Plan, the syncretic code that unites all history, science and religion. It is hard to give the sense of this concocted Plan in a brief quote or paraphrase, since it relies on the sheer number as well as the manner of the connections for its effect. A good example, though, comes in the chapter when Belbo, now fully immersed in the Diabolical method, announces that although they have been whimsically inventing connections up to this point, he realises that they do definitively exist:

> [Guillaume] Postel studies Hebrew, tries to demonstrate that it's the common matrix of all languages, translates the *Zohar* and the *Bahir*, has contacts with the cabalist, broaches a plan for universal peace similar to that of the German Rosicrucian groups...And it is no accident that he signs some writings with the name of Risispergius, "he who scatters dew."
> Gassendi in his *Examen Philosophiae Fluddanae* says that Rosencreutz does not derive from rosa but from ros, dew. In one of his manuscripts he

speaks of a secret to be guarded until the time is ripe, and he says: "That

pearls may not be cast before swine." Do you know where else this gospel

quotation appears? On the title page of *The Chemical*

Wedding…Furthermore, it seems [John] Dee and Postel met in 1550, but

perhaps they didn't yet know that they were both grand masters of the Plan,

scheduled to meet thirty years later, in 1584. (419)

Robert Phiddian observes that this sort of reasoning is performed with an "imperious

rhetoric that defies reasoned contradiction," and that this "show of confidence and

connectedness is enough to make all but the dourest rationalist, empiricist, or pragmatist

wonder if there might not be something to this confident defiance of conventional

wisdom."[20] We thus, Phiddian implies, may be induced to become Diabolical readers

ourselves, paranoid that things as they appear on the surface conceal a deeper conspiracy

of interrelatedness, as every text is a code that connects it to another. In the writings of

Ambroise Paré, we saw that the work of the devil lies in conjuring superfluity, "an infinity

of monsters and marvels."[21] By collecting the monstrous and excluding the demon which

brings with it the possibility of deception, the Renaissance naturalist tamed a wild and

unruly nature as nevertheless "natural," being in accordance with a divine order. In

Foucault's Pendulum, by contrast, collation is infinite, leading to the demonic delusions

perpetrated by the Diabolicals. It is not coincidental that the section quoted above is

headed by an epigraph which states "Though his will be good, his spirit and his prophecies

are illusions of the Devil" (417).

However, through a logical fork, the novel also undermines the reader who

stubbornly refuses to be seduced by the demonic language and epistemology. As I

presented it in chapter 3, contemporary psychology strips the demon of any metaphysical

existence in the world, emphasising instead that demons are in the mind of the beholder.

Though it retains its baggage as a potent myth, we recognise the demon in modernity as a

fiction and thus, according to Gabriel Josipovici, an allegory of language's own myth-

making process rather than of its mimetic capacity. In this framework, we may feel that

Belbo's liturgical stream of words signifies little of actual, empirical relevance but rather

points inwardly to his own unconscious motivations of distraction: too young to fight in the

War, and too old and cowardly to join the social revolution of 1968, Belbo constructs the

[20] Robert Phiddian, "*Foucault's Pendulum* and the Text of Theory," *Contemporary Literature* 38 (1997): 546.

[21] Paré, *On Monsters* 94.

plan to compensate for his own frustrated revolutionary ambitions. Additionally, as a litany
of proper names, making sense of passages like these would tax even the most determined
bibliophile. Belbo's language convinces himself, in a self-fulfilling prophecy, but his
argument does not signify any absolute truth about hidden connections. But if we do try to
read Belbo's language as all surface without essential meaning, mere stylistic play or
ludism that does not need to be explored or followed up because it is purely superficial,
then paradoxically the plot counterattacks. The book concludes that "the most powerful
secret is a secret without content, because no enemy will be able to make him confess it"
(621). Belbo knows that by allowing himself to be killed rather than reveal a "secret" when
he is tortured by Diabolicals at the end of the novel, he proves to them not the absence of a
truth, but the fact that it must be present, if deferred. Stylistically, through its soup of
arcane areas of knowledge, the novel seems to argue for the absence of any one system of
correspondence. This logic folds in on itself, as in plot terms the novel's allegorical
conclusion seems to be that the denial of allegory – the withholding of the one code index
that would make sense of all other referents – is the most powerful evidence that there
must be a final, true grid to decode all meanings which the novel's withholding of a final
clue (as a defeated detective fiction) intends conspiratorially to conceal. In the ultimate
poststructural tease, the relative absence of a coherent allegorical scheme and the existence
of Belbo's demonic plots imply not the invalidity of that totalising index to meanings, but
its tantalising possibility. The reader is caught in an interpretative pincer: if absorbed by
the language and occult wisdom, we are Diabolicals thrilled by the connections; if we treat
the occult as if it has no basis in empirical reality, this implies there really is a hidden Plan
which will unite reality and the occult, which our secularised, modern minds are too
unimaginative (or not sufficiently paranoid) to admit.[22] As Madsen observes, allegory is
both an interpretation and a metacritical statement that guides interpretation, leaving
allegory as the tautologous conflation of two kinds of truth: "truth as the meaning of the
narrative and truth as the interpretation of a prior text."[23] *Foucault's Pendulum* exhibits
and exploits this circularity, in that it either points to the secret that unites all texts outside

[22] Victoria Vernon reads the novel slightly differently, arguing that if we do not
follow the moves of the protagonists we are immediately tempted to become critical
analysts, though the problem now is that we are confronted with a bewildering array of
possible strategies for interpretation. Victoria V. Vernon, "The Demonics of (True) Belief:
Treacherous Texts, Blasphemous Interpretations and Murderous Readers," *Modern
Language Notes* 107 (1992): 840-54.

[23] Madsen 95.

it, or points to the powerful secret within it which is that there is no secret. As with postmodern fictions according to McHale's model, epistemology spills into ontology, and threatens to lead to a vicious cycle. Epistemologically speaking, what is wrong with a Diabolical reading method? If it is not a bad reading, then is not a syncretic novel like *Foucault's Pendulum*, which pulls together Hermetic history and cybernetic technology, not itself one more Diabolic text, a machine for generating interpretations by referring to intertexts? The ontology of the novel thus hints that the ontology we take to be reality is actually a delusion concealing some deeper meaning. But if something is epistemologically wrong with the paranoid method of reading, what is the right way to decode reality? Given that postmodern theory denies a stable epistemology for reading the world and the book, is not the absence of a final logocentric clue – like that which Belbo withholds – the most powerful evidence of some conspiracy to conceal something from us? When encountering such a paradox in conventional allegory, one might have referred to the external, stable referent of the Christian tradition, the Bible. In the absence of this external authority, and given cybernetics' subversion of the creativity of the controlling "author" outlined above, it is the reader who is tasked with breaking from its logical loop by paying close attention to the constructions of the text.

Forced to follow one prong of the novel's logical fork, the repetitive use of arcane images, texts and language is not completely seductive. This deliberate tension between the noisy verbiage of alterity and the meaningless plot which thus conceals a secret evidence that whilst a novel may be a machine for generating interpretations, readers are not interpreting machines. Hermeticism's links with the occult leave it reliant on the vocabularies of magic and diabolism, and thus suspiciously anti-scientific (modern homeopathy, for example, is justified by a Hermetic connection which analyses and "cures" based on the resemblance of symptoms between different biological agents). Even if we do not subscribe to the Enlightenment model of absolute scientific rationality, there is still a clear discrepancy between the discursive modes of Hermeticism, and those of contemporary society. Hermeticism shouts loudly to conceal its emptiness, but in the process it draws attention to itself as a particular use of language, rather than to the ideas in its own right; the conspiracy in which every meaning is tied to another is stylishly chaotic, but ultimately devoid of much meaningful content. The intervention of Hermetic reasoning and style in the novel poses a comparable problem to that which Shannon and Weaver introduced at the dawn of cybernetics: can we discriminate the communication of information from mere noise, and quantifiably measure the amount of information a message contains? Shannon and Weaver theorised that a message that is partially

unpredictable and partially anticipated contains more valuable information than either a completely random or a completely predictable message; however, a message with greater predictability is also more resistant to the intrusion of noise.[24] In a similar way, postmodern fiction must negotiate a tricky position by innovatively staging the difficulty of making sense of a complex world through language – given that conventional allegory is no longer sustainable – whilst not alienating readers by depriving them of conventional signposts like linear plot and character, rendered in essentially realistic ways. The difference between the noisy but meaningless Diabolical Hermeticists, and the postmodern but nevertheless meaningful mode of *Foucault's Pendulum*, is a cybernetic issue, and it is central to the ambition of the novel to highlight bad readings, whilst allowing any number of readings to be theoretically sustainable.

The relationship between "noise" and meaning is one Eco has used as an analogy for interpretation, arguing that the code of meaning shared by those communicating provides "an established system of probabilities, and only on the basis of the Code can we decide whether the elements of the message are intentional (desired by the source) or the result of Noise."[25] Eco differentiates between signification, in which semiosis is potentially unlimited and meanings are indeterminate, and communication, which is socially grounded and by which the novel can lever a reader into a more definitive interpretative position. As Gary Stonum observes, Eco's standpoint shares with the cybernetic study of useful information in communicative practices.[26] Eco's fiction opens channels in which he transmits (though not entirely clearly) a framework or code which limits the possibilities of interpretation and distinguishes model readers from those Diabolical characters who are anything but ideal ones. Although "code" can imply concealment and secrecy, it also suggests to those who do know the code a communal relationship based on mutual patterns of understanding – the anticipated readers of a novel in a particular genre, for example. Within this crucial paratextual framework, as Eco notes, "A text is a place where the irreducible polysemy of symbols is in fact reduced because in a text symbols are anchored

[24] Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana, IL: U of Illinois P, 1949).

[25] Eco, *Travels* 138.

[26] Gary Lee Stonum, "Cybernetic Explanation as a Theory of Reading," *New Literary History* 20 (1989): 400-01, *JSTOR*, 17 Sept. 2008 <http://www.jstor.org/stable/469107>.

to their context."[27] If the Other allows us to apprehend better ourselves by our negative, this relationship means that the fact that we can perceive the Diabolical characters and their language as different indicates the existence of a more normative linguistic code from which it differs, and through which more substantial meanings may be communicated. Eco's work anticipates a community of thoughtful readers capable of discriminating socially conventional interpretations from demonic readers – though certainly not without encountering the dialectics and seductions of a decentred postmodernity where allegorical truth always slips beyond the grasp of the text. Eco argues that "many texts aim at producing Model Readers, a first level, or a naive one, supposed to understand semantically what the text says, and a second level, or critical one, supposed to appreciate the way in which the text says so."[28] Eco appreciates readers are able simultaneously to juggle their awareness that the text is talking in a Diabolical way, with the understanding that the very fact that we can interpret with analytical distance indicates we have not succumbed to the metaphysical speculations of the Hermeticist. In order to achieve this, *Foucault's Pendulum* seduces us into being naive readers of Diabolism, but then draws attention to itself as a text which (as text) is different to reality, or to reality as it is construed by the Diabolical adept. Thus it opens space for the critical reader to appreciate the way in which the first stage reader was initially seduced by the reasoning of the Diabolical Plan.

Of particular importance in this regard is Eco's treatment of time. I suggested earlier that Eco's cybernetic conspiracy is unusual in that it is not only international in scope (with scenes in Brazil, Italy and France), but also threads through history from modernity to the Middle Ages. The Diabolicals see historical events and facts as coterminous with, and as relevant as, those of the present; in contrast, constrained by the linearity of print, the reader is able to perceive the ironic inversion of cause and effect that results. For example, the Diabolicals believe that the view that "time is a linear, directed sequence running from A to B is a modern illusion. In fact, it can also go from B to A, the effect producing the cause" (207). In terms of the novel's form this initially appears to be a valid statement. The novel's time scheme is complex, as the bulk of it is narrated through three layers of retrospection: by Casaubon awaiting the Diabolicals at the end of the novel, remembering his awaiting the Diabolicals in the Conservatoire National des Arts et

---

[27] Eco, *Limits* 21.

[28] Eco, *Limits* 55.

Métiers, where he "reconstructed so anxiously" (45) the events leading up to his being in the museum. Thus the outcome of the novel, its chronological close, is what allows the story to come into being: effects literally produce causes.

In the naive reader's initial experience, the novel unfolds and progresses from a mysterious outcome to (we anticipate, wrongly) the revelation of the cause of that mystery, in a detective plot. Although there is a risk that, like the academics, we become absorbed in the narrative of the Diabolicals, as I suggested earlier the extent to which the occult descriptions and connections push the limits of our attention seems to protect against the possibility of the latter, at least for Eco's "critical model reader able to enjoy, at a second reading, the brilliant narrative strategy by which the first level naïve reader was designed."[29] Alice Juarrero has picked up on this technique, reasoning that "Once we appreciate the paradox, that the secret is that there is no secret, we realize we've been had semantically, but like children wanting to figure out just how the magician performs the trick, we just *can't* leave it at that."[30] Linda Hutcheon also observes the irony which undercuts "overt resemblance" ensures the novel differs from the Plan: "Without irony, Eco's novel would be an exemplar of Hermetic semiosis; with irony, it becomes simultaneously both an exemplar and a critique."[31] With respect to the past of the novel, in re-reading the text we cannot do so in anything other than an ironic mode, thereby undermining the Diabolical principle that time is non-linear; for this alternative, ironic form of knowing relies precisely on the contingent and structured way we read a conventional novel, especially one based on the detective or thriller template. Such a rereading of the history, both short-term in the novel we have just read, and the longer-term of its Middle Ages purview, is a postmodern strategy, but one which does not necessarily cause us to agree with the Diabolical method in which history is not linearly regulated by empirical time, since signifiers from the past and present can freely interchange. Eco notes that "The postmodern reply to the modern consists of recognising that the past, since it cannot really be destroyed, because its destruction leads to silence,

---

[29] Eco, *Limits* 55.

[30] Alicia Juarrero, "The Message Whose Message It Is That There Is No Message," *Modern Language Notes* 107 (1992): 896.

[31] Linda Hutcheon, "Eco's Echoes: Ironizing the (Post) Modern," *Diacritics* 22.1 (1992): 5.

must be revisited: but with irony, not innocently."[32] Even if we do feel that there is a secret buried in the text, and return to it in order to excavate its potentially allegorical code, the interpretations become satirical and ironic rather than providing syncretic resolution.

As he starts his narration (towards the chronological close of the plot) Casaubon realises that "I had to play this ironically, as I had been playing it until a few days before, not letting myself become involved" (10). However, the truer irony is that he *is* narrating from a position of involvement, both believing the Plan and the sole narrator of the novel which tells us about it. For example, at the beginning we are told that the periscope lets Casaubon look through the Apse in St. Martin "as if I were swaying there with the pendulum, like a hanged man, taking his last look" (10). On a first reading, this is an innocent metaphor; on a second, when we know that there will indeed be a hanged man, Belbo, in just such a position, it becomes sadistic humour. Likewise Casaubon's note at the beginning that "We don't want to meet Ingolf's mysterious end" (151) is said with a smile; but on a second reading, given that we now know that Belbo and Diotallevi will also die strangely, we react with a different sort of ironic humour. As Victoria Vernon observes, whilst the reader's arrival towards the chronological close of the plot might potentially allow us the advantages of narrative hindsight, this is defeated by the fact that "the narrator, too, is only discovering the meaning of his story through the process of its telling."[33] The majority of the novel is narrated retrospectively by Casaubon, who has travelled to France from Italy following a message from Abulafia (453), and who is as he narrates lodged in a periscope in the Conservatoire. Just as a periscope displays the world inverted or refracted, he can hardly be the most objective of narrators.

In terms of its Diabolical structure, the novel's narratorial ambiguity in which effects produce causes offers an analogy with the first Cabalic computer program written by the scholars, which is designed systematically to produce all the possible permutations for the letters YHVH. Marshall McLuhan declares that the instantaneous operations of a computer put an end to linear sequence and the typographical order of syntax. Taking issue with McLuhan, Eco responds that "In fact the *programming* of a computer consists precisely in the arranging of linear sequences of logical operations."[34] So, though the

---

[32] Umberto Eco, *Postscript to* The Name of the Rose, trans. William Weaver (New York: Harcourt, 1989) 66-67.

[33] Vernon 847.

[34] Eco, *Travels* 237.

program may run in a recursive fashion (through an "if…then" loop, according to the program given in the book [34]), its outcomes have to be printed linearly in the physical text, or what Eco terms the Linear Text Manifestation.[35] Although a computer may carry out its operations through parallel processes and feedback loops, its outputs are still received by a very limited serial processor who finds it difficult to multi-task: the reader. In *Foucault's Pendulum*, the fact that the output of Abulafia is delivered in the form of print undermines the academics who program it, who believe truth is revealed by ahistorical, non-causal connections between facts from any period or text. To recall the quotation by Belbo above, John Dee and Postel met in 1550, but did not realise that they were actually the grand masters in the Plan that would manifest itself thirty years later, in 1584; time runs from B back to A. As a novel, though, *Foucault's Pendulum* plays off its physical textuality, the Diabolical theory that time is not linear, and the reader's ability to perceive these characters ironically, because they themselves cannot predict their own endings even though they claim to be able to "prove" the causal chain of events that captures everyone else unwittingly in a Plan. In computer programs, recursive loops (such as that set up in the YHVH program) mean that each output becomes a new input; by analogy, the Diabolicals believe that some event in the past can be conceived as governing an event in the future, even if the two do not immediately seem to be interlinked. However, the reader can only move through the textual representation of both the Abulafia program and the Diabolical Plan in one direction – that laid out by print. David Porush notes that the text which has certain words underlined might be distracting for the ordinary reader but significant for the spy who knows the systemic code by which they were produced; thus cybernetic fictions as Porush understands them often reveal that "information and nonsense can only be distinguished by their context."[36] Whilst cybernetics abstracts information from a body so that it floats as an abstract pattern, cybernetic fictions embody ideas in the guise of print. Likewise in *Foucault's Pendulum* the physical body of the text restricts meaning within a certain spectrum of possibilities, or ensures that those meanings are modified ironically depending on the temporal context, the order in which they appear in the book, or whether they are first readings or rereadings.

The linearity of printed texts undermines the atemporal detachment of information from context posited by Hermeticists and cyberneticists alike. However, the mutability of

---

[35] Eco, *Limits* 148.

[36] Porush 66.

the meaning of the text, where what is taken seriously by some readers becomes ironic on re-reading or re-presentation in linear print, does also endorse the poststructural view of language, since meaning depends on context. However, like Hermetic or cybernetic literary strategies, the multiplicity of meaning thrown up by irony also has its limits. The book's language is empirically static, since on a re-reading we do not suddenly find new words or sentences in there: the book as a printed artefact is stable, and hence objectively legible. We may be aware that, as Barthes claims, "a text is not a line of words releasing a single 'theological' meaning (the 'message' of the Author-God) but a multidimensional space in which a variety of writings, none of them original, blend and clash."[37] However, when judged against the image of the Abulafia program and with the ironic distance produced between the Diabolical readers within the text and the sceptical readers without, a text seems a comparatively coherent, linear and scripted system, and not simply a machine which runs autonomously of its programmer (according to Barthes, of course, the author must die to create the text). Although the Diabolical pursuit of meaning may *se*duce us on the first reading, the second reading to which it *in*duces us catches the characters in the ironic pose of being characters staged for the purpose *of* the book, hence undermining the importance of their quest *in* the book. As Eco suggests:

> the notion of unlimited semiosis does not lead to the conclusion that there are no criteria for interpretation. To say that the interpretation of every text is potentially unlimited does not mean that there are no criteria for interpreting it. To say that a text potentially has no end does not mean that every act of interpretation can have a happy ending.[38]

Happily, though, the act of interpretation of *Foucault's Pendulum* can have an end – of sorts. If we avoid falling into the same trap as the characters – their multiple acts of interpretation that lead them (will in a printed book always lead them on any number of readings) to an unhappy ending – we can begin to interpret from a position outside the text, knowing the text as a machine for generating interpretations for critical readers who are, however, not mere interpreting machines.

In the previous chapter I argued that the stylish and realistic presentation of *The Matrix* is evidence for the cinema viewers that we might really be taken in by a convincing Matrix simulation in everyday life. However, even if the world is ontologically a delusion,

---

[37] Barthes, "Death" 146.

[38] Eco, *Limits* 6.

we are still able to conceptualise a hierarchy of *The Matrix* within the Matrix, so that even if reality is a simulated hyperreal, it is not a uniform one. This differentiation is shown through the use of old technologies, the incarnation of the economic and computer system as physical entities, and self-consciousness of its "bullet time" effect as opposed to Baudrillard's theoretical and hyperreal bullets. However, the risk remains that in spite of these differences, we still have no way of escaping from the Cartesian conundrum, the deceit of the demonic machine. With *Foucault's Pendulum*, we can contrast responses to this postmodern condition at a generic level of film and literature (picking up the cue left in chapter 6 about the media shift between the novel *Do Androids Dream of Electric Sheep?* and the film *Blade Runner*). Like *Blade Runner* or *The Matrix*, *Foucault's Pendulum* is a postmodern work that shows the impossibility of attaining final meaning or ultimate truth. Like *The Matrix*, Eco's novel also tries to distance the interpreter of the fiction from the state suffered by the characters within it. But as a question of degree, the distancing effect in *Foucault's Pendulum* provides a greater level of irony. The fact that we are absorbed by the film in the moment of watching, in the stream of conscious images it presents, powerfully suggests that we could be deceived more generally by a virtual simulation of reality outside the cinema. However, the linear nature of print encourages a different, double sort of consciousness. For the print and the book are read in one direction, even as the temporal order it narrates is non-linear. This allows us to perceive an ironic distance between the Diabolicals as they feel themselves to be relate to historical time (that is, fluidly and non-causally) and the readers as they respond to the characters in a more structured fashion produced by the linear, causal nature of a printed detective plot. Likewise, the characters seem to produce persuasive arguments that sweep us along in the Diabolical method of reading archival documents. But whilst in *The Matrix* the fact that the other world is rendered so convincingly is evidence that our own "reality" could also be perfectly fabricated, in *Foucault's Pendulum* we see two markedly different styles at work. The reasoning of the Diabolicals is engaging but mere noise, whereas the logic of the book at root is a code for interpretation that distinguishes us as good readers from them as bad semioticians. Though in principal any reading can be legitimate, as the Diabolicals wish, in practice, some readings are better than others, as is indicated by the function of style. In *Foucault's Pendulum*, the sincerity with which the Diabolicals read archival documents contrasts with the irony with which the reader looks at them, mocking their convictions.

Ironically, irony itself becomes a quasi-rational perspective to adopt on the text and the world. Without denying reality exists, the Popperian scientist considers and falsifies

every other potential interpretation of reality before choosing the one which provides a best fit. The paradox presented by *The Matrix* – and by postmodernism in general – is that we cannot falsify the evidence, since the virtual realities of the computer or demonic language immerse us in a simulation controlled from outside the opaque walls of our own system, and therefore any empirical speculations about reality made from within the system will be always already deluded and wrong. An objective view from nowhere is unattainable. However, in the case of a novel, since we know it to be a simulation in order for us to hold it as a book, we are automatically placed outside the system. Thus our ability to take a view from nowhere with respect to it, from which we read ironically rather than colluding with the characters' reasoning, validates the scientific method. Within *Foucault's Pendulum*, the premise of the plot is that empirical and mythical values are congruent and equal, since one can "arrive at the truth through the painstaking reconstruction of a false text" (459). The illogical nature of this is self-evident, however, since we are reading precisely one such known "false text," a fictional novel which, as readers, we do not expect to make didactic claims to truth, merely, in the first instance, to entertain.

In a demonic act of disbelief, the Diabolicals pursue every possible avenue of interpretation in the hopes of reaching the right conclusion, the syncretic code that will confirm that the surface is an illusion and the interconnections beneath point to some master Plan. Readers are torn between their unwillingness to enter their conspiracy, and the logical fork that states that the withholding of truth is the most powerful evidence that there is a syncretic truth to be found, even if it is concealed. However, by playing off the linear temporality with which we read against the fluid time scheme, we are able to attain an ironic distance. Likewise, the physicality of print undermines the ephemerality of information. Its readers, however, are anticipated as being able to distinguish between the deceiving noise and valid information, to acknowledge (as good postmodernists) that signifiers point to innumerable signifieds in a process of deferral, whilst nevertheless reading according to a heuristically reasonable range of meanings. This heuristic quality places the novel as a cybernetic fiction. N. Katherine Hayles argues that deconstruction was about the arbitrariness of language, whereas cybernetic models of information focus on the gap between noise and communication, changing the emphasis "from the limitations of language in producing meaning to the limitations of code in transmitting messages accurately." Rather than emphasising the instabilities of a writing subject as deconstruction does, "coding theory implies a transformation from the writing subject to a posthuman

collective of encoders/decoders."[39] If one quality of cybernetics in relation to literature is to devalue the unique originality of the author, the flip side is that it emphasises the collective of literary decoders, that is to say, readers. *Foucault's Pendulum* implies that the practice of coding and decoding at least verifies that there is a meaning of sorts lodged within the text, even if accessing it seems so difficult that reading sometimes appears an arbitrary process of making meaning. For once, the multiplicity of interpretation thrown up by a postmodernist text is actually a validation of the value of scientific method. It makes a powerful Popperian call for falsification, because the text falsifies the validity of its own Hermetic strategy. As Eco has argued, semiotics ought to accept a Popper-like principle that even if we cannot produce final, good interpretations as we could in the tradition of Christian allegory, we can at least acknowledge which are bad. By providing a counterpoint in Abulafia, the novel evidences that though writers themselves may be machines, whilst texts are machines for producing interpretations, readers are not interpreting machines. The paradox is that showing this at work in the form of a critical essay such as the present one requires the academic reader to follow the same sort of methods as the Diabolicals, deploying a collage of reordered quotations and secondary texts in order to get at a meaning of the novel that is not that of the original. At a certain point, then, it is necessary for the critic to follow Casaubon's rationalistic girlfriend, Lia, who turns Ingolf's code wheel to its most logical translation, and to conclude: "Shit, I'm sick of this Hermetic writing" (540).

---

[39] Hayles, *My Mother Was a Computer* 68.

Chapter 9. <u>Uniting the Two Cultures: Body and Mind in A.S. Byatt's *A Whistling Woman*</u>

<u>Two Cultures</u>

As with Umberto Eco's fiction, A.S. Byatt's novels are mirrors within which disciplinary and generic opposites such as humanism and religion, art and science, or critical and creative writing reflect each other.[1] Although examining one idea in isolation risks shattering the complex illusion of the whole, Byatt often signposts her textual play with a single, key metaphor which recurs in different contexts. In both her best-known work, *Possession* (1990), and in *A Whistling Woman* (2002), this metaphor is the demon. In the former, as we shall see in chapter 10, the title and its associated trope guides readers across the boundaries of genre and history. In terms of the romance prototype, possession is what happens when one is in love, but also signifies literary critical obsession; in the Gothic melodrama and for the Victorian spiritualist, possession was used in the literal sense of the embodied ghost, but even in rational modernity and the postmodern novel we still encounter the almost paranormal way in which our lives seem to be determined by our ancestors. The multivalence of this particular trope allows – indeed demands – that readers and critics consider the significance of its imprinting upon different cultures at a theoretical, even deconstructive, level.

Rather than using the metaphor to cross historical periods, *A Whistling Woman* (2002) uses the demonic trope to traverse academic disciplines and a variety of scientific, philosophical and literary investigations of the phenomenon of consciousness.[2] The final novel of the "Frederica quartet," *A Whistling Woman* mainly takes place in a fictional interdisciplinary university in North Yorkshire, against the backdrop of sexual and social revolution in the 1960s. Hosting a spectrum of scientists, such as a young mathematician, Marcus Potter, and the genetic biologist, Luk Lysgaard Peacock, the University is preparing to stage a major conference on "Body and Mind." This is organised by the Vice-Chancellor, Gerard Wijnnobel, a mathematician and grammarian who desires "a cognitive-biological Theory of Everything" (26), and who chooses as one of his keynote speakers Hodder Pinsky, a cognitive scientist. Counterpointing this academic setting, we also follow

---

[1] A version of this chapter was first published as Alistair Brown, "Uniting the Two Cultures of Body and Mind in A.S. Byatt's *A Whistling Woman*," *Journal of Literature and Science* 1.1 (2007): 55-72, 28 June 2008
<http://literatureandscience.research.glam.ac.uk/journal/issue1/>.

[2] A. S. Byatt, *A Whistling Woman* (London: Chatto, 2002). All subsequent parenthetical page references refer to this text.

the case of Joshua Lamb who, having witnessed his father's Abraham-like sacrifice of his mother and sister, becomes a charismatic Manichaean visionary and eventually leads a cult on the North Yorkshire Moors. Finally, Frederica Potter, the heroine of the four novels, is carving a new role for herself as both an emancipated mother and a Byattian public intellectual, hosting a television series on art, science and politics. With this wide range of characters and their intellectual backgrounds, the demon provides an image that orientates disciplinary perspectives around a common point of focus. Despite preconceptions, the demon does not offer a Frankensteinian critique of science through the deployment of a monstrous myth that defies realist analysis, thereby concocting a "two cultures" hierarchy in which literature has representational privileges over science. In a recent article for *Nature* entitled "Fiction Informed by Science," Byatt elaborated on the central scientific motifs around which the novels of the Frederica quartet were structured: the synaesthetic solving of mathematical problems through visualising landscapes in *The Virgin in the Garden* (1978); the nature of perception in *Still Life* (1985); the Fibonacci spiral in *Babel Tower* (1986); and synaptic and biological connections and patterns in *A Whistling Woman* (2002).[3] Over the two decades of the quartet's conception, science has provided a central inspiration, and so in spite of its demonic keynote *A Whistling Woman* sincerely values scientific concepts.

In *A Whistling Woman*, Byatt's exploration of the common and different epistemologies of each discipline is orientated around her understanding of consciousness and the connections and patterns, metaphorical or synaptic, between mind or soul (traditionally the preserve of the arts and religion) and the body or the material brain (subject for scientific scrutiny). This contest is established early in the novel, through Frederica's television programme on creativity, which stages a debate between Elvet Gander, a Jungian psychoanalyst, and Hodder Pinsky, a cognitive linguist who uses computers to explore "what he believed to be the deep and universal structures of linguistic competence" (29). The producer, Edmund Wilkie sets the tone for the debate: "Is the unconscious mind a system of circuitry and binary gates?...Or is it the Id, a turbulent beast raging in the dark?" (150). The confrontation between Gander and Pinsky hinges on determining which paradigm satisfies as an explanation of the "sublime pathos" that is the experienced work of art: does Freud provide a useful metaphor for thinking about the profound but disguised connections from which creativity arises, or does his language

[3] A.S. Byatt, "Fiction Informed by Science," *Nature* 424 (2005): 294-297.

"detract from various very useful practical explorations of the workings of the unconscious mind" (154) carried out by cognitive science, as Pinsky claims? Modern cognitive science seems reductive, seeing people as simply computational mechanisms following fixed rules (even Freud was simply "an unusually lucid computer" [150], according to Pinsky). However, Pinsky ironically subverts the rational in his computational model of mind, which is itself a metaphor of "Pandemonium," drawn from Milton's image in *Paradise Lost*. Pinsky's empirical approach is ironically accentuated through the demonic metaphors which recur in different contexts throughout the novel, and the multivalent usages of the trope highlight that which is implied in Pinsky's language, but which he does not admit in his ambitions for cognitive science: that a final prescription for how particular neurological actions (the "easy problems") give rise to concrete conscious experiences (their corresponding "hard problems") will be impossible without recourse to linguistic analogy.[4] By contrast, Freud's treatment of the unconscious as a "beast raging in the dark" – which we saw literally extrapolated in *Forbidden Planet* – acknowledges the diversity and perversity of human experience, providing a powerful metaphor for consciousness even if its scientific basis is suspect. In her non-fiction, Byatt quotes Josipovici's psychoanalytically informed "Demon of Analogy – the sense that what we thought was Out There and Other is only a description of the inside of our own skulls."[5] Pointing to the use of analogy across science and the arts, the demonic motif allows Byatt to explore the relationship between the fictional and the real, the consciousness that intercedes between discourse and the physical world, and the elucidations of reality from various disciplines.

The first section of this chapter will trace the range of demonic signification as it attests to the gap between physicalism and mentalism as this maps on to the two cultures in *A Whistling Woman*. On one religious interpretation a demon is a physical agent performing the metaphysical will of God, transmitting messages which the receptive mind can interpret. With his apparently allegorical character, Lamb – a Lamb of God? – embodies this possibility. But if Lamb's mind is possessed by demons such that his mental attitude stands for something outside his own embodied self as a religious symbol, this implies a return to Christian allegory over scientific materialism, as if minds or souls alone contain meaning. Such allegorical determinism also limits the referential multiplicity with which a postmodern novel like *A Whistling Woman* plays, in its denial of a single

[4] For the explanation of the "easy" and "hard" problems, see chapter 1, note 55.

[5] A. S. Byatt, *On Histories and Stories: Selected Essays* (London: Chatto, 2000) 122.

authoritative perspective on the issue of consciousness by invoking the multivalent demon trope. Ironically, cognitive science is not unlike religion in this respect. Although cognitive science considers brains and bodies to be co-extensively material, it represents brains as simple deterministic devices or calculating engines; this makes it hard to see how they can produce the full and exciting variety that characterises human culture, such as literature, or myths like demonism. For although the demon may not exist in "reality," it possesses such a common cultural currency as denoting a state of mind (such as the demon of schizophrenia) that its existence should be examined even from the most fundamentally objective position. The demon exists because people think it so, although there may not be a single neuron or suite of cells that constitute the demonic. Whilst psychoanalysis acknowledges its folk existence, as Byatt shows, cognitive science also uses demons as placeholder terms in theorising the brain, and the use of a mythical metaphor within science suggests that mind cannot be reduced to its neurological components, but must demand a literary as well as empirical sort of interpretation. The first section of this chapter seeks to understand these questions from the perspective of the reader, and points to Byatt's reworking of any extant two cultures problem as the analogue of Cartesian dualistic thinking. In this translation, Byatt attempts to show the value of epistemological plurality in the vein of postmodernism, without slipping into an anti-science stance.

In the second section, I follow Byatt's efforts to write of mind in terms of body, such that the novel's representation of character can overcome the limitations of reading according to the absolutist perspectives offered by religious allegory or scientific determinism. As Byatt is aware (through her friendship with real-life biologists such as Steve Jones and neuroscientists such as Antonio Damasio), increasingly the body is proposed as the site from which thought is governed.[6] This rejection of dualism leads to an aesthetic and moral problem for the novel as it attempts to incorporate such findings. How can it represent minds as embodied things, when the only mode of representation is language and metaphor which is always abstract? And if it puts too much stress on embodied thinking, does this not simply substitute for its rejection of allegorical dogmatism with the conviction that human subjects are biologically determined?

---

[6] The acknowledgements page of *A Whistling Woman* reads like a roll call of the most prominent contemporary scientists, including Steve Jones, Richard Dawkins, Matt Ridley, Antonio Damasio, Semir Zeki and John Maynard Smith (422).

Deceiving Demons of Mind

At the heart of the demonic web of connections, and the figure against whom all the other metaphors of the demon are contextualised, is Joshua Ramsden (or Josh Lamb, as he calls himself symbolically). Haunted by his childhood experiences when he saw his evangelical father murder his mother and sister, as an adult Lamb achieves a kind of epic visionary status. He receives Manichaean messages which instruct him to spread their word, and he sees the world overlaid with images of blood and light. As his name is intended to suggest, in a Christian era he would have been welcomed as a prophetic figure, an intermediary through whom visions from God flow, their messages not about Lamb himself, but pointing towards some theological truth outside of him. But in Byatt's rendering of his psychology, we cannot dogmatically assume such a connection between the individual and what he analogically evidences about wider society. In one passage which is symptomatic of Lamb's schizophrenic consciousness, the God-like, omniscient narrator describes Lamb's struggle to come to terms with his personal trauma:

> In the *Arabian Nights*, it has been said, a man has his Destiny written on his
> forehead, and his character, his nature, is that Destiny and nothing else. A
> boy, a man, like Josh Lamb, Joshua Ramsden, who has found himself
> tumbling in the dark sea outside the terrible transparent mirror of the fragile
> window-pane, persists perhaps by linking moments of conscious survival
> into a fine suspension-bridge of a personal destiny, a narrow path of
> constructed light, arching out over the bulging and boiling. (106)

Lamb's visions are given form only by the authorial potency which designs their metaphorical correspondence, symbolically writing his destiny on his forehead. Moreover, our readiness to isolate Lamb as a theological paradigm is made complicated because the knowledge of Lamb's mind presented to the reader is not absolute or diagrammatic. Carefully crafted though the image is, the metaphors do not link internal thoughts with a stable, allegorical context outside the self; rather, they describe a passion wholly internal, and the metaphorical linkages are formed between the images, as the series of nouns and adjectives blend together: now a terrible mirror, now a fine suspension bridge, now a path of light. These images are denotative of Lamb's state of mind, but their subjective interconnections are only sparsely connotative of external, allegorical links. Is the "bulging and boiling" over which Lamb's life moves supposed to be seen as Hell or the subconscious? What does it signify about the role of destiny when Lamb finds himself tumbling in a sea outside a transparent mirror (not that there can be such a thing) of a particularly fragile (for an unspecified reason) window-pane? In a writer less assured than

Byatt, it would be tempting to condemn this passage as a hopeless collection of mixed metaphors. In the case of a novelist who takes metaphor as one of her focal subjects, however, the strangely articulate meaninglessness of the passage deliberately confronts us with a phenomenological experience. The only thing that we can be clear of here is that consciousness itself is happening in an imagistic stream, that the world of the text is being experienced through the consciousness of the other, and that any of the entities posited by that consciousness can not be regarded as referring to actual things in the physical world.

The fact that there is not one but multiple consciousnesses (the author's, Lamb's and the reader's) working to concretise the imagistic stream into meaning, means that we can never observe the psychological sensation reliably and objectively. Earlier in the novel, Lamb sits in his room in a psychiatric hospital, observing blood dripping from the walls and clothes. When the narrator tells the reader that "He watched the blood" blanket his world, we necessarily watch by proxy and, since we share his vision, then in one sense the blood cannot be private to him but must be public, physically in *the* world (at least of the novel) as well as psychologically in *his* world. This is not simply a theoretical view of the paradoxes inherent in the third-person narratorial standpoint; rather, this complication is explicitly what makes his visions, in Lamb's ironic phrase, "an interesting phenomenon" (35). Lamb wonders of the blood:

> Was it there or not there? He was certainly seeing it – with his eyes – noting its viscosity and flow. He was not making it up. It wasn't a projection of his state of mind, which was calm, not bloody. It was not a metaphor.
>
> On the other hand, he was almost entirely certain that if he picked up the soaked sock, it would be white wool, and would not drip red. (35)

The person ultimately "making it up" is the novelist, with whom we are complicit as readers, and the novel's metafiction makes us uncomfortable. If the blood is not a psychological metaphor which projects publicly something otherwise private (the unconscious of a madman or religious prophet) then what is it? Surely a novel is generally an analogy or displacement of the real? As I argued in the previous chapter, contemporary fiction shifts the determination of allegorical meaning away from an external referential pool (such as Christianity), and away from an external source of creativity (the author), towards the reader as coproducer of meanings.[7] But actively to reconstruct meaning in the absence of an external and uncontested frame of reference, such as the Christian one, is to

---

[7] Heidi Hansson provides this interpretation of allegory in relation to Byatt's *Morpho Eugenia*. Hansson 454.

distort reality by assuming *a priori* that a textual phenomenon, such as Lamb's visions or his Christ-like name, must mean something publically and allegorically. This, of course, is precisely the sort of mistake Eco's Diabolicals make. Albeit with less visceral consequences, Byatt seems similarly to suggest that the novel is privileged in giving us access to the interior mind of a character, but that we can never use a novel to enter someone else's thought processes objectively, without mutating them into something they are not, symbols rather than a direct, empirical perception of the mind of the other. In relation to consciousness, therefore, a novel is not automatically a better medium of interpretation than cognitive science. A few pages earlier (29), the novel mentions Heisenberg's uncertainty principle, which asserts that the fact of observation may change the state of a particle being observed. Just as if Lamb touches and picks up the sock the illusion of blood (if "illusion" is what it is) will be dispelled, it seems as if the moment we interrogate the passage this changes the nature of the blood which we observe from the omniscient perspective, the novel's equivalent of the empirical view from nowhere. Whilst it seems "real" for Lamb, the moment we read it we cannot help but turn it into an abstract symbol, doing as readers precisely what Lamb does to a dangerous extent in interpreting his visions as indications of his religious vocation. As Roman Ingarden theorised, because literary works typically contain "spots of indeterminacy" the "reader usually goes beyond what is simply presented by the text (or projected by it) and in various respects completes the represented objectivities."[8] As Heisenberg's uncertainty principle emphasises a probabilistic Nature in which it is not possible to monitor all the properties of a system simultaneously, so in the literary knowledge available here the implicit presence of a readerly observer does not force the logic one way or another, towards reality or unreality, physical or mental phenomenon. Like the Cartesian deceiving demon, Lamb's demons prove only the limited fact that consciousness is happening and must be accepted as such in its own bizarre terms; they do not determine what is being thought and how it relates to reality in any absolute way.

The poised doubt about whether the blood is present in the world or just in Lamb's mind maps on to the discrepancy between the "easy" and "hard" problems of consciousness. In the case of literature, printed words – analogous to the easily observable electro-chemical impulses of the brain – do give rise to a complex, indecipherable gestalt

---

[8] Roman Ingarden, *The Literary Work of Art: An Investigation on the Borderlines of Ontology, Logic, and Theory of Literature*, trans. George G. Grabowicz (Evanston, Il: Northwestern UP, 1973) 252.

of consciousness both in the open-ended characterisation of Lamb and in the reader struggling as analyst. Because it is not possible to extend the blood analogy further, to determine whether the very fact of its presence at all evidences something other than the mental – for example, a metaphor for the death of conventional religion – this is a moment when we are confronted also with the demonic analogy in Josipovici's theoretical sense. We cannot connect Lamb's subjective experience to some objective statement of the state of the world's religion, without recognising also that it has for Lamb its own, self-contained meaning, that another mind exists which is separate to that of the individual reader and that, however disturbed it appears to be according to our own preconceptions, it is as personally genuine a perception of reality as our own. As I observed in chapter 1, in cases like that of the virtually "married" man, this issue of how much moral legitimacy we accord to delusions is actually a vital one in a cybernetic era when we might enter simulated worlds which seem genuine to us whilst appearing, from the objective and public point of view, to be mere simulacra. Coming obliquely to this debate, *A Whistling Woman* at least causes us to know that the individual is unique and isolated from the world, whilst at the same time preventing us from falling into the real solipsism Josipovici warns us against: that of not knowing that one is in a state of solipsism.[9] The belief that we inhabit a shared, common reality, rather than an external world always perceived differently by different people, is the true delusion. However, if postmodernism seems to deny the possibility of an absolute "reality" in a world always filtered through a culturally determined discourse, cognitive neuroscience as it was developing in the 1960s, simultaneously with postmodernism, seemed poised to provide a stabilising antidote to this linguistic and psychological relativism.[10] If reading turns everything into analogy, and language is an epiphenomenon of consciousness rather than reliable evidence of discrete thought processes, then cognitive science and artificial intelligence in their most reductionist guises promise to get back to the origins before phenomenology, not *ana*logy but the logos of thought itself.

Central to this view is Hodder Pinsky, the computational psycho-linguist. Unsurprisingly, Byatt admits his name is a conflation of Steven Pinker and Noam

[9] Josipovici 307.

[10] See my discussion in chapter 1. Notably, Byatt read Jean Pierre Dupuy's work, *The Mechanization of Mind*, in preparation for *A Whistling Woman*. "Fiction Informed by Science" 296.

Chomsky, and she also may also have had Marvin Minsky in mind as well.[11] Appearing on Frederica's television show with Gander, the psychoanalyst, Pinsky announces:

> There is...an interesting computer programme [sic] called *Pandaemonium*, which is psychologists' everyday comic poetry, not sublime, though it takes its name, I suppose, from the industrious underworld of *Paradise Lost*. This programme has a hierarchy of mechanical demons who are devised, or designed (by us, their masters), to recognise patterns in rushes of random information, to create order from noise. It depends on what we call "parallel processing." There are the "data demons" who recognise images, and shout. There are the computational demons who recognise clusters of recognised images, and shout. There are the cognitive demons who represent possible patterns, and collect the computed shouts. And there is the "decision demon" who identifies the stimuli by the loudest shouting. The system can learn. It can identify printed letter, and morse code. It may one day understand what is so – unrepeatable – about *Hamlet*, or Beethoven's Third. (153)

This "society of mind" model which Minsky expounded (drawing on the work of O.G. Selfridge) and which Daniel Dennett develops as the multiple drafts model in his popular *Consciousness Explained*, seeks to unite the easy problem of consciousness with the hard problem.[12] These models assume that mind is made from smaller systems (agents or demons) that individually perform processes which do not require thought but that collectively constitute consciousness. It is not surprising that this model of consciousness lends itself well to computer programming and artificial intelligence, with its subroutines and discrete variables. However, there is a problem with this form of cognitive science

[11] Byatt, "Fiction Informed by Science" 296.

[12] Marvin Minsky, *The Society of Mind* (New York: Schuster-Touchstone, 1988), especially 18-23, 274, 327. Minsky describes the demon as "An agent that constantly watches for a certain condition and intervenes when it occurs." The concept of the demon in parallel processing originates with Oliver Selfridge, "Pandemonium: A Paradigm for Learning," *Proceedings of the Symposium on Mechanisation of Thought Processes*, ed. D. V. Blake and A. M. Uttley (London: H. M. Stationary Office, 1959) 511-529. See also O. G. Selfridge and U. Neisser, "Pattern Recognition by Machine," *Scientific American* 203 (1960): 60-67. The quotation from *A Whistling Woman* may be a partial paraphrase from this latter article, which in a similar anthropomorphism notes that "A Pandemonium program can handle the situation by having the demons shout more or less loudly" (66). However, probably most readers will have been made aware of the cognitive "demon" through Daniel Dennett, *Consciousness Explained* 189, 261-63.

which is made clear through the demon trope. Whilst asserting Pandemonium as an empirical model, the literary reference to Milton and the use of the signifier "demon" in a novel already full of them (rather than a more neutral term, such as "subroutine") turns the model into a metaphor for what cannot be known in the multiple and complex connections between the two problems of consciousness. Rather than allowing us a non-relativistic and objective representation of mind, cognitive science adopts a similar perspective to the reader looking into Lamb's ambiguous mental images; it too makes allegorical references with things beyond its apparent focus, in order to concretise things that are otherwise empirically slippery. Further, Pinsky does not see that the fact that the demons are "designed by us (their masters)," ironically lends credence to Lamb's earlier view that his mind is a conduit from the divine designer who despatches his demonic agents. As a rationalist, this is something Pinsky would not accept. However, as John Searle has pointed out, whilst artificial intelligence fulminates against dualism, the whole thesis of strong artificial intelligence relies on it, since "it rests on a rejection of the idea that the mind is just a natural biological phenomenon in the world like any other."[13] In *A Whistling Woman*'s comparable critique, the different intersections at which the demonic metaphor appears intuit that the cybernetic argument assumes that mind can reside independently of the body on circuit boards or in software; that it is led through that same Cartesian tradition to invoke concepts of an abstractable mind that might one day understand the qualia of music; and that this results in a search for recursive levels of homunculi-like demons every bit as mysterious as those that possess the Manichean messenger, Lamb.

Despite the chaos of signification that Lamb's psychology exposes, at least the demonic metaphor in relation to Lamb leads us to an understanding with moral implications, because it forces us, as readers, to confront the fact that we inhabit a universe of plural experiences, in which the mind of the insane still deserves our attention and interest. Though we might note the affinity with Philip K. Dick's emphasis on empathy, more directly Byatt inherits from George Eliot's *Middlemarch*, which demands that we see the world not only from the point of view of the heroine Dorothea – from which Casaubon seems a dry scholar – but from within his perspective also.[14] In contrast, the Pandemonium model simply collapses in on itself, since even as Pinsky claims through his cognitive

[13] Searle, *Minds, Brains and Science* 38.

[14] The crucial passage is that where Eliot interjects, "but why always Dorothea? Was her point of view the only possible one?" George Eliot, *Middlemarch*, ed. David Carroll, Oxford World's Classics (Oxford: Oxford UP, 1997) 261.

probes and computer models to circumvent the need to explain mind through analogy, it is predicated on a metaphor of hazy "demons." This is not to suggest that Byatt represents literature as systematically superior to science (later in the book, at the Body and Mind conference, Pinsky speaks with admirable circumspection about the current state of his science). However, through metafiction, the novel can recognise its own limits in representing consciousness. A highly self-reflexive novel open to multiple interpretations, particularly of Lamb's mindset, *A Whistling Woman* presents itself as the embodiment of the hard problem of consciousness with all its irreducible variety, rather than supposing that objectively to measure the easy problem is automatically to extrapolate the subjective quality of the hard problem. What makes this transcend the conventional two cultures is that the computational approach to consciousness, essentially a problem of the relationship between body and mind, is shown to bear affinity with literary allegory, which can be equally determinate in a religious tradition which connects words on the page to divine symbols. The allegorical tradition assumes that written language – a form of easy problem – corresponds directly to particular external theological signifieds, just as cognitive science assumes qualitative experiences (like demons) can be pinpointed to precise neurological processes.

As explored in chapter 3, though, psychoanalysis offers a powerful fork: it deconstructs the possibility of demons as allegorical markers of a metaphysical state, whilst enabling them as metaphors for the human condition of perceiving the world distortedly. In its attempts to produce final understandings of the mind, the computational focus of modern cognitive science fails to account for the subtle varieties of the human experience that is even encoded in the scientists' own language and metaphors. In contrast, the Freudian interpretation seems doubly validated, because it both seeks to explain consciousness through exploring the meanings behind language, whilst in that explanation it generates a new rhetorical formations which are also highly innovative, even creative, in their attempts to elucidate the quality of being in the world. Byatt has admired Freud as a most profound thinker, a kind of narrator whose connective use of language is itself poetic: "You take a word and see in what contexts it shines and glitters."[15] Freud is a unifying personality at the heart of the novel, whose analytical stories carry "an immediate, wholly satisfactory verbal pleasure in pattern, and [reach] out into biology, and human history, like rings round a stone in a dark pool" (155). However, though psychoanalysis has long

---

[15] A.S. Byatt, interview with Andrew Marr, "The Brain," *Start the Week*, BBC Radio 4, 6 Oct. 2006.

provided the modernist writer with an epistemological framework in which to deal with questions and representations of consciousness, in *A Whistling Woman* Byatt also implies that the Freudian focus on the submerged depths of meaning beneath a conscious act or piece of language can be deceptive, misdirecting attention away from the empirically plain surface of things as they actually are. For example, invited to a weekend away with her former Ph.D. supervisor Luk Lysgaard Peacock, Jacqueline watches as he moves through his house which he has, in her honour, carefully lit with lanterns:

> When Luk was at the far end of the table, he looked like the old, familiar, too familiar Luk. As he came dancing towards her, with his offerings, his bearded face passed from pool of light to pool of light, from candle-light to island of lamp-light. When these fiery lights were under his face he looked unfamiliar. He looked demonic. (177)

In a deliberate planting of associative language, offerings reminds of sacrifice, sacrifice reminds of Lamb's history, his history connects with Abraham, Abraham's hearing of the voice of God links back – grotesquely – to the paternalistic bearded face of Luk here. The play between the unfamiliar and the familiar reminds of Freud's uncanny and, as we saw in chapter 3, confrontation with the uncanny is a self-reflexive realisation that demons are not out there in the world, but are projected onto others through the unconscious motivations and desires of the self. However, empathising not with the author, with her Freudian knowingness, but interpreting the scene at a more immediate level, the conclusion "he looked demonic" also has a simple and rational causation, a kind of verifiable "easy" problem: it is nothing more than a trick of the slightly dim and changing light. Any other interpretation is a projection of symbolic meaning onto a material surface that is easily rationalisable.

The problem with the Freudian framework is that it allows so many connections to be made as to produce disparate rather than coherently symbolic meanings. To paraphrase Umberto Eco, when the author knows about psychoanalysis from a theoretical position, the creative text might seem little more than a machine for generating associations. The novel's metafictionality highlights how we anticipate correlations between Lamb's mind and the social or moral world in which it operates; in a similar way the psychoanalytic perspective anticipates correlations between conscious and unconscious, connections which can distract from the reality. In its relativistic drive, psychoanalysis can never conclusively interpret how things really are since something must always remain beneath the surface. The benefit of the neuroscientific approach, by contrast, is that it circumvents analysis of mind through the epiphenomenon of language, instead probing the brain

beneath the skull directly. Further, psychoanalysis sets out a scheme of expectation that entices us to read in a particular way. By so deliberately planting psychoanalytic meaning in a novel, the novel makes psychoanalytic criticism appear to be attacking a straw man, since subsequently to theorise the text in a psychoanalytic vein is to do nothing more than the novel anticipates. (This problem of critical foremeaning is explored in greater depth in *Possession*.) In a comment that accurately describes the readerly predicament of deciphering the allegorical status of Lamb's visions, the psychiatrist Kieran Quarrell observes that "Part of the problem…is that genuine spiritual seekers (whatever the 'spirit' is) do bear some resemblance to the truly batty whose wiring has short-circuited." As Lawrence Kubie observed, the psychoanalyst struggles to remain detached from his patient.[16] Just like Christian allegorists, the psychoanalyst may project rather than draw meaning from his patient. Gander agrees, in a reply to Quarrell: "you know my deep aversion to false 'objectivity' and therapic claims of 'detachment' wch [sic] are bound to be bogus" (61). Finally, Brenda Pincher, an ethnomethodologist who plants herself among Lamb's cult, knows that if she revealed her status as a scientific observer she "would change the dynamics of the group so that it was not what I was observing, or what I wished to observe" (193). Unlike the neuroscientists in the novel, however, these three psychiatric practitioners at least admit the limitations of their methodology. Rather than this problem being unique to the psychoanalyst, as I suggested earlier Byatt's representation of Lamb suggests that as readers we already read from a particular, expectant angle: "making people up" is precisely what the novelist and the reader expect to do. The difference is that a novel is always controlled and edited by an author, whereas psychoanalysis attempts to offer a diagnosis and therapy whilst really being a projection of the analyst's desires onto a text or subject.

A Whistling Woman evidences how fundamental metaphor is to thought, such that literature, religion, cognitive science and psychoanalysis become, in their attempts to explain mind, simply ways of talking evasively around a problem which remains beneath our skulls. But as Lamb's visions hint, there should be something that is "not a metaphor," that does not evasively mutate into analogy as soon as we try to study it. Though the novelist strives to transcend the traditional two cultures opposition and to affirm the value of science and the arts as different ways of looking at the same phenomenon of mind, it is not easy to reconcile the two cultures because they seem so inextricably predicated on the

---

[16] See chapter 3, note 11.

analogous dualism of body (which can be dissected and analysed objectively) and mind (which can only be understood through subjective representations, such as language). As a result, representations of mind tend to be extracted from the embodied individual as representing some position outside the text, implicitly asserting subjectivity over objectivity, allegory over realist representation. Is it possible to work in a non-allegorical way, such that psychological sensations are represented as being unique to the person having the thoughts such that they cannot necessarily be defined according to some external scheme? Byatt's attempts to represent holistically embodied rather than singularly allegorical minds, which are not determined but can be interpreted variously by her (postmodern) readers, accounts for the strange second half of the novel. In a number of ways, this seeks to negotiate a way between the genetic and cognitive determinism of the body, and a conversely indeterminate aesthetics of the transcendent mind. Firstly, and not without paradox, Byatt deploys physical metaphors to describe the activities of mind, drawing on the recent neuroscience of embodied cognition. Secondly, she denies biologically deterministic forms of science through juxtaposing these with other scientific paradigms which are more amenable to literature. The effect of both these moves is to encourage the reader to cultivate an open-minded, agnostic attitude towards meaning, avoiding the tendency to accept the determinisms either of scientific materialism or of allegory.

Incarnate Ideas

Byatt has often acknowledged the status of George Eliot in the tradition of the intellectual novel, describing her as "the great English novelist of ideas." She argues that Eliot's characterisations are successful because unlike didactic writers such as Orwell, Peacock or Huxley, Eliot's characters do not "represent ideas like allegorical figures" but the inverse: ideas "are as much actors in her work as the men and women who contemplate the ideas."[17] The same could be said of Byatt's multidimensional weaving of disciplines into a mesh that places the intellectual quality of ideas, more than any one particular character, at the formal centre not only of her novels' themes but also their plots and settings.

However, when one of these ideas is the relationship between brains and consciousness, there is a more literal sense in which ideas can be embodied in a novel.

---

[17] A.S. Byatt, "George Eliot: A Celebration," *Passions of the Mind: Selected Writings* (London: Chatto, 1991) 72-76.

Byatt also notes of Eliot that she "saw her work as making incarnate certain ideas that she apprehended in the flesh, i.e., sensuously, materially, through feeling,"[18] something achieved particularly in *Middlemarch*, "the great novel of the body-mind exploration" written by a novelist who acts like "an embodied mind."[19] In her essay on "Memory and the Making of Fiction," Byatt uses several physical metaphors to describe the corporeal process of thought: one "is of feathers – being preened, until the various threads, with their tiny hooks and eyes, have been aligned and the surface is united and glossy and gleaming"; another "is of a fishing net, with links of various sizes, in which icons are caught in the mesh and drawn up into consciousness"; a third is rhythmic memory as a kind of knitting.[20] Rather than ideas being incarnated at the level of plot (something Byatt achieves successfully throughout almost all her fiction) here ideas are incarnated – described in bodily terms – at the scale of the unique psychological moment. Byatt's many representations of memory as something physical as well as psychological are increasingly endorsed by modern neuroscience which stresses embodied cognition, given the persistent failures of a purely cybernetic artificial intelligence. Foremost among these later scientists has been Antonio Damasio, whose anti-Cartesian argument that "the life of the mind arises from the life of the body is leading us to reconsider all kinds of aesthetic problems and proceedings."[21] Though resonating with Eliot's aesthetic, how can this empirical twist be applied to literary metaphors for the individual consciousness? The innate dilemma of a novel would seem to be that, forced to use metaphors to express mind, the novel implicitly disembodies thought from its origin, doing precisely what the representation of Lamb's mind warns against in moral terms, or Pinsky's demonic model in theoretical ones. Metaphor draws its power from construing connections between entities that appear

[18]A.S. Byatt, introduction, *Selected Essays, Poems and Other Writings*, by George Eliot, ed. A.S. Byatt and Nicholas Warren (London: Penguin, 1990) xxix. Byatt draws attention to Eliot's review of Robert Mackay's *The Progress of the Intellect*, in which Eliot asserts that "Religion and science are inseparable," and describes civilisation and religion as "an anomalous blending of lifeless barbarisms, which have descended to us like so many petrifications from distant ages, with living ideas, the offspring of a true process of development" (xxi).

[19] A.S. Byatt, "The Feeling Brain," rev. of *Looking for Spinoza*, by Antonio Damasio, *Prospect* 87 (2003): 73-4.

[20] A. S. Byatt, "Memory and the Making of Fiction," *Memory: The Darwin College Lectures*, ed. Patricia Fara and Karalyn Patterson (Cambridge: Cambridge UP, 1998) 47-72.

[21] Byatt, "The Feeling Brain" 73.

superficially different (memory and knitting, thought and feathers), and is hence an unreliable vehicle to reconcile the two cultures of the material body and abstract mind. Though Byatt praises Damasio's work on embodied cognition, metaphor and language must always strain to reconcile word and world.

With this substantial qualification, the value of embodied cognition from the meeting of science and aesthetics is explored mainly through the mathematician, Marcus, whose thoughts often paraphrase Byatt's personal writing about memory whilst resonating with Damasio's work. In the first book in the quadrilogy, *The Virgin in the Garden*, Marcus synaesthetically experienced equations as a kind of garden, in which the answer to a particular problem was revealed through colours and shapes.[22] Now, in *A Whistling Woman*, Marcus works at the University on a project entitled "The Computer as a Model for the Activity of the Brain," but complains that it is mere "Number-crunching. A primitive automat that mimics a limited number of operations in the brain" (220). Mid-way through the novel, though, Marcus comes under the sway of a homosexual Wittgensteinian philosopher, Hodgkiss, who observes "Marcus's body as an expression of Marcus's mind" (293). Marcus explains that whilst he can move finite numbers around on paper, constituting them as "real things" rather than forms, he experiences infinite concepts synergistically:

> Did Wittgenstein hate angels? Did he think they were all really demons? Oddly, I feel the infinities with the whole of my body, not only my mind – they don't feel cerebral. I feel as though I'm *in* them, I *am* them – as opposed to observing them. It's possible I can't so to speak get out of them because human beings made them up. (219)

Are numbers real things or phantoms? As I noted in chapter 1, the question was key to Warren McCulloch's cybernetics, which pondered "What is a number that a man may know it: and what is a man that he may know a number?"[23] Marcus implies here – and the reference to demons enforces further – that because demons exist in the mind they must at least possess some sort of reality. As with Lamb's demons, however, it is not possible to "get out of them," to occupy a scientific view from nowhere on them, because to leave the mind of the person who perceives is also to destroy the possibility of that perception

---

[22] *A Whistling Woman* often deploys synaesthetic descriptions of memory that relate thought to mathematical shapes (216-22), or to biological entities such as the spiral shells of snails or flighty birds (374-75).

[23] McCulloch 2.

existing in the first place. This is contrary to cybernetic belief that information – even mind – can be abstracted from its original container and translated into the cybernetic program. The difference with Marcus is that he feels the infinities with his whole body – not as disincarnate mentalisms – but even as this seemingly opens the possibility of a scientific study such as that pursued by Damasio, the only way he can make this feeling available to himself and to Hodgkiss is "By analogy. I have to substitute other things – things I do understand – for the infinities I don't" (219). Paradoxically, the more embodied mind or mental images seem to become, the harder it is for science to get a representative grip on them and the more language or metaphor must take over. However, though literature may be more comfortably working with analogies that can connect body to mind, Byatt simultaneously acknowledges that it is limited. Language in both Byatt's non-fictional reflections on memory and in Marcus's thoughts is stretched to its adhesive limits, in the attempt to describe, positively, that which is known to be unknown, as unrepeatable in the symbolic order of words as Beethoven's *Third* is in Pinsky's cognitive model.

It remains unclear quite how theories such as Damasio's embodied cognition might be reconciled with a literary sentiment that necessarily disincarnates consciousness by representing it through language; nor is it self-evident how one might acknowledge that science might in the future provide a definitive account of consciousness, whilst preserving for literature the ability to represent mind in a constructive, if limited, way. In attempting to produce a scientifically informed novel, the risk is the author might be tempted into overemphasising the role of the body in thought, pushing it almost towards endorsing the biological determinism of genetics or computational neuroscience. As Marcus admits, what he feels as the "infinities" may be merely "flashes in the brain" (219). The idea of the synaesthetic mathematical genius who sees things such as numbers differently because of the biological structure of his brain might come close to echoing Steven Pinker's view that we are not blank slates upon whom cultural experience and ideologies are forced, but actually have genetically predisposed capacities.[24] The moral and political risks of this view – risks that in part account for the contemporary anger directed at Steven Pinker by cultural relativists – is that it has its extreme realisation in racism. Byatt is evidently alert to the extension of Pinker's blank slate critique, since it is the presence of Eichenbaum, a

---

[24] Steven Pinker, *The Blank Slate: The Modern Denial of Human Nature* (London: BCA-Penguin, 2002). This is the clearest expression of Pinker's biologically deterministic approach, although *The Blank Slate* was published after both *Babel Tower* and *A Whistling Woman*. Byatt is thus drawing on similar ideas which were expressed throughout Pinker's earlier works such as *The Language Instinct* and *How the Mind Works*.

scientist who worked in Nazi Germany, that sparks riotous student protests at the body and mind conference in the climax of the novel. Whilst trying to embody metaphor in affinity with science, she must also avoid extending this appreciation towards legitimating a deterministic science that not only biologically destines the synaesthete to a state of genius, but also destines a race, class or gender to inferiority. And it is, of course, precisely the overturning of such views – particularly the latter – that characterises the novel's 1960's moment.

The character of Peacock, the evolutionary biologist, represents one way of moving away from a two cultures model in which literature axiomatically holds prestige in matters of mind, but without approaching scientific dogmatism. Described as "a bristling genetic predestinarian and moral pessimist," when he finds himself attracted to Jacqueline he appears to be the personification of a real genetic determinist such as Pinker and the antithesis of a literary romantic:

> He thought of Lorenz's studies of the behaviour of other creatures. Most unreceptive females bit, or scratched or snarled. This one just took a step or two out of reach. The signs were quite clear. She did not want him. What puzzled him as a scientist, given the unambiguous clarity of her discreet messages, was how much he wanted her. (21)

Behavioural psychology sees the body as the expression of mind (Dennett's "intentional stance"),[25] but unlike Byatt's reflections on the weaving of memory, it is stripped of any emotional sensitivity to human action. Peacock's language is dispassionately asexual, scrutinising Jacqueline as any other scientific object, not as a subjective entity in her own right: heavy in sense, it lacks sensibility. Although the omniscient narrator, possessing true access to consciousness, can tell us that Peacock regrets the demands of scientific objectivity, he does not admit it publicly. For example, Peacock accepts that in the interests of his research he ought to stop counting spirals on snail shells in the wild and use instead lab-based electrophoresis, but still he cannot help but prefer to be outdoors, from where "He measured the world from inside the balance of his own body" (69). This ironic discrepancy is summarised by the fact that this scientist who dogmatically asserts that "his world was washed clean of human stories" (188) is actually the central character of one.

[25] See chapter 1, note 4.

He even has, like Lamb, a highly allegorical name, since Lysgaard is a "common Danish name meaning garden of light, a paradisal reference."[26]

Thus we are prepared to applaud when Frederica argues against Peacock that evolutionary and biological causes cannot in themselves explain why metaphor and artistic perception are so central the human experience, that "theories of natural selection don't explain why human beings find peacock feathers beautiful" (409) – or why, ironically, Frederica will fall in love with this particular Peacock by the end of the novel. However, whilst it might be tempting to see this opposition between biology (Peacock) and literature (Frederica) as enacting disciplinary dualism, the novel makes a further ironic attack from *within* the scientific field, presenting the metaphorical interest of biological science whilst not wholly endorsing its logical findings. The title of Peacock's conference paper is "The Cost of Sex and the Redundant Male":

> You might say, said Luk Lysgaard-Peacock, thinking of Frederica and her inconvenient assertion that mental cobwebs [such as astrology or demons] were real things – you might say, that if an idea has survived for a very long time, it has its own adaptive fitness. You could argue that religions and moral instructions survive in the world because they are like larger organisms, struggling for existence. You could argue that Christianity spread to be a world religion because it had better survival characteristics than Manichaeism....But a faith is not an organism, and survival works at the level of the fitness of cells, through the adaptation of cells. I would like you to recall the admonitions made yesterday by Professor Pinsky, against thinking loosely with analogies and metaphors. (38)

Whilst Peacock, speaking in the fictional context in 1968, argues for survival working at the level of cells, he is necessarily naïve – whilst the reader is probably alert – to the fact that *The Selfish Gene* of 1976 would argue for adaptation at the level of the gene, initiating a paradigm shift. This undermining by anachronism ensures that science is contested in terms of other branches within science, as part of a shifting suite of paradigms, rather than simply by an aesthetic shudder against reductionism of any sort. Further, Peacock's argument here is elusive. In his rhetorical style, an argument is presented only to be demolished; this therefore raises the possibility that religion and stories *are* understandable in evolutionary terms, even if this is subsequently denied. In itself, this analogy would be

[26] Byatt, "Fiction Informed by Science" 296.

provocative, as indeed was Dawkins' use of the cultural memes and biological genes comparison in his book.[27] However, Peacock then argues against his own metaphor as evidencing the danger of "thinking loosely by analogy." Peacock takes it a stage deeper, explaining culture not by metaphor but by embodied materialism. Since survival takes place only at the level of cells, contrary to Frederica, the only "real thing" objectively in the world is biological material that is quantifiable. He thus moves unselfconsciously from a soft to a hard evolutionary reductionism. Such a move re-stages the difference between Steve Jones, the biologist to whom *A Whistling Woman* is dedicated, and Stephen Pinker. In opposition to Pinker, Jones argues that cognitive science will not achieve "a universal exegesis of the set within the skull, a mental theory of everything" through recourse to evolutionary biology.[28] Indeed, adding to his unwittingly ironic status in the novel, Peacock's hesitancy in his speech – "you could argue," you might say" – implies also that this is to be read as an argument with which the scientist is not entirely happy, it being intended rather to be provocative, having been inspired by Jacqueline's jilting him after she miscarries his baby.

The brilliance of science admired by Byatt is that it does not succumb to the easy escape into Romanticism's evasive myth of the transcendent mind and consequent anti-empirical stance. However, representing the embodiment of ideas in relation to the novel's characters risks using dense analogies that either admit the dualism on which the novel is formally predicated as it uses metaphor to express body, or a move towards a biological or cognitive-biological determinism. The best attitude to adopt, therefore, is agnosticism rather than dogmatism. As Byatt has set this philosophy out, "I don't like novels that preach or proselytise. (I fear people with very violent beliefs, though I admire people with thought-out principles.) The novel is an agnostic form – it explores and describes."[29] As Dawkins' idea of the meme admits in the final chapter of *The Selfish Gene*, or as the ironic way in which Peacock is treated suggests, genetic determinism cannot account for the way in which we act in the world of culture in unexpected or counterintuitive ways. Thus the

[27] Dawkins 189-201.

[28] Steve Jones, "The Set Within the Skull," rev. of *How the Mind Works*, by Steven Pinker, *New York Review of Books* 44.17 (1997): 13-16.

[29] A.S. Byatt, "Author Statement," *Contemporary Writers*, 2002, British Council, 28 Feb. 2008 <http://www.contemporarywriters.com/authors/profile/?p=auth20>. Byatt reiterated this sentiment about agnosticism at a recent lecture: A.S. Byatt, "In Conversation with A.S. Byatt," Durham Literature Festival, Gala Theatre, Durham, 8 Oct. 2007.

more Peacock dogmatically assumes himself to be driven by genes and sexual selection, the more this draws attention to himself as determined by a literary code. But this is not just to promote literature at the expense of science. The fact that we know, as Peacock does not, that his view of survival taking place at the level of cells will be ultimately condemned by the historical arrival of *The Selfish Gene* reminds us that agnosticism is also scientifically healthy, since to remain dogmatic is to be closed to the Popperian possibility of new paradigms superseding the old. Politically, too, agnosticism prevents the extension of the biological logic of genetics into a racist or eugenicist politics. Finally, for the reader, the agnostic attitude permits the incarnation of ideas as the property of each individual, rather than idealistically linking to an allegorical frame of reference outside them. It is the attitude inculcated in the practice of reading pluralistically figures such as Lamb who might otherwise be dogmatically interpreted in religious terms; learning to accept Lamb as having an autotelic, if schizophrenic, consciousnesses in his own right allows us to critique his particular brand of fundamentalist religion (which leads to a fiery disaster) whilst not asserting equally fundamentally the value of our rational view, or denying his right to have a consciousness in the first place.

Cultivating the spirit of the agnostic in this vein gives us a potential map for interpreting the novel's ending, which seems initially bewildering and unsatisfying. The novel and quartet concludes with Frederica carrying Peacock's baby, a literal and symbolic incarnation of artistic creativity and evolutionary determinacy, the lottery of genetics played through the instinctive game of romance. The pair look out over the moors, in a vision in which echoes *Paradise Lost*:

> In the distance, the man-made Early Warning System, three perfect, pale, immense spheres, like visitors from another world, angelic or demonic, stood against the golds and greens and blues. Frederica said to Leo "We haven't the slightest idea what to do." Everyone laughed. The world was all before them, it seemed. They could go anywhere. "We shall think of something," said Luk Lysgaard-Peacock. (421)

Given the generally realist tendency of the novel, this pastoral scene stands out awkwardly; equally, the romance plot is unconvincing, since Frederica and Peacock have been intellectually and emotionally opposed for the majority of the novel. However, it is arguable that its awkwardness as allegory and romance defines it against generic or genetic determinism. The general ambivalence – "We shall think of something" – denies the novelistic writing of destiny on the forehead, refusing dogmatically to orientate the characters against an allegorical close. In parallel, technology is agnostic, as evidenced

sharply in atomic science of the 1960s (the Early Warning System refers to the Fylingdales radar, used to detect Soviet missile launches) which could be turned either for productive or destructive ends, angelic or demonic but not decisively Frankensteinian.[30] In Peacock and Frederica's concluding romance, and the representation of nuclear arms as morally ambivalent, Byatt tests the reader's willingness to resist the polarisation of the two cultures. We must acknowledge the capacity of science to create new technological images and metaphors – the pale and immense spheres – which have become the dominant icons of a secular age, taking over from the angelic and demonic categories of religious belief. The discussion of consciousness centred around the trope of the demon forces us to recognise that literature does not have the unique privilege to access mind unobtrusively, since allegory can slip too easily towards logocentric dogmatism. On the other hand, science does not always threaten the body by making instruments of war, but can also bring to life new metaphors that can be recruited by literature as descriptions of mind.

[30] For example, President Dwight D. Eisenhower's 1953 "Atoms for Peace" speech to the United Nations. Dwight D. Eisenhower, "Atoms for Peace," United Nations General Assembly, New York, 8 Dec. 1953, transcript, International Atomic Energy Agency, 19 Oct. 2007 <http://www.iaea.org/About/history_speech.html>.

Chapter 10. Breaking the Spell of Postmodern Possession

The footnotes engulfed and swallowed the text. They were ugly and
ungainly, but necessary, Blackadder thought, as they sprang up like the
heads of the Hydra, two to solve in the place of one solved.[1]

Proliferation of reference rather than final meaning, layers of interpretation rather
than fresh creativity, an alertness to the unattainability of an authoritative text: the
academics in *Possession* experience all the symptoms of postmodern scholarship, suffering
from the anxiety of uncovering influence in Ash's work. For the Victorian poet,
multiplicity is a virtue, the capacity to absorb a variety of knowledge and represent it
through epic poetry being the hallmark of successful nineteenth-century literature; the most
obvious historical personalities on which Ash is modelled, Alfred Tennyson and Robert
Browning, are polymaths in that vein. In contrast, for the contemporary critics in the novel,
the numerous analytical discourses at their disposal – feminism, psychoanalysis,
poststructuralism, historicism – open so many possible perspectives of interpretation,
disseminated in such diverse forms – the critical edition, the biography, the journal article
– that it becomes impossible completely to grasp the original historical figure as a
personality. This experience of alienation from one's explicit subject is recreated to some
degree for the novel's reader through *Possession*'s own congested polyphony. Straddling
fact and fiction, it embeds fabricated poems and historical quotations, journals, letters and
tales, whilst generically it inherits from many different plot prototypes, including the
campus fiction, the romance, the fairy story and the detective novel.

One obvious paradox that results from this comparison between contemporary and
Victorian intellectual writing is that *Possession* is clearly a postmodern work even whilst it
parodies this attitude. In identifying what postmodernist niche *Possession* occupies, we
might well situate it within Linda Hutcheon's definition of historiographic metafiction:

In most of the critical work on postmodernism, it is narrative – be it
literature, history, or theory – that has usually been the major focus of
attention. Historiographic metafiction incorporates all three of these
domains: that is, its theoretical self-awareness of history and fiction as

---

[1] A.S. Byatt, *Possession: A Romance* (London: Vintage, 1991) 28. All subsequent
parenthetical references in this chapter refer to this edition.

human constructs (historiographic metafiction) is made the grounds for its

rethinking and reworking of the forms and contents of the past.[2]

A.S. Byatt herself has agreed that it "is a postmodernist, poststructuralist novel and it

knows it is. It does present itself as a piece of Victorian melodrama, but of course it's no

such thing." However, she goes on to complain that to label an author as postmodernist is

to contaminate them with an affected guile, in which the writer's skill in reflecting on the

impossibility of representation in his or her text predominates. Byatt asks that "readers be

allowed to identify with characters," allowing a space for an emotion other than admiration

of the author's wit.[3] The tension in *A Whistling Woman* is between allegorical determinacy

and individual subjectivity, and between science and art; the tension between the

pleasurable and the intuitive, and the theoretical and the rational, similarly runs throughout

Possession. It informs its plot, style, the way in which we read this intellectually complex

but enjoyable romantic novel, and (the principal subject of this chapter) its reaction to

contemporary science as it intertwines with postmodern literary theory and practice.

Strangely, given its ultimately encyclopaedic range, Byatt has pinpointed the

moment when the concept for the novel took hold, which concentrates its scope in the

single word of its title. As Thelma Shinn observes, *Possession* validates Barthes' claim that

the title is the "prince of signifiers,"[4] since the word of its title "captures the perspective of

the novel and orders its 'recurring web of metaphors'."[5] As Byatt has recalled, she first

thought about *Possession* as she watched the Coleridge scholar Kathleen Coburn in the

British Library:

> I thought: she has given all her life to *his* thoughts, and then I thought: she
>
> has mediated his thoughts to me. And then I thought "Does he possess her,

---

[2] Linda Hutcheon, *A Poetics of Postmodernism: History, Theory, Fiction* (London: Routledge, 1988) 5.

[3] A.S. Byatt, interview with Nicolas Tredell, *Conversations with Critics*, ed. Nicolas Tredell (Manchester: Carcarnet, 1994) 62.

[4] Roland Barthes, "Textual Analysis of a Tale by Edgar Allan Poe," *The Semiotic Challenge*, trans. Richard Howard (New York: Hill and Wang-Farrar, 1988) 268.

[5] Thelma J. Shinn, "'What's In a Word?' Possessing A.S. Byatt's Meronymic Novel," *Papers on Language & Literature: A Quarterly Journal for Scholars and Critics of Language and Literature* 31 (1995): 172.

or does she possess him? There could be a novel called *Possession* about the relations between living and dead minds."[6]

It is impossible objectively to pinpoint the origin of this impulse in relation to postmodernism. Did Byatt notice the irony of a biographer possessed by her subject only because of her existing interest in postmodern questions of subjectivity? Or did this inspirational observation only coincidentally come to be guided by the critical discourse circulating in academia at the time? Both possibilities can be sustained legitimately: without the atmosphere of postmodernism, Byatt might never have been interested in the role of the biographer; and without her vision of the biographer, she might not have written the novel we have, one which partly happens to be, partly is intended to be, postmodernist. The status of theory as anterior to or implicated within the writing of Byatt's novel is apt, because the fact that writing always takes place within a system of academic concerns that inform as much about the writer as about their subject is perhaps the dominant recognition to emerge from the novel itself. Numerous moments in the novel present the existence of an irresolvable feedback system between the writer and the subject they write on. Perhaps paradigmatic is when Blackadder, the British Ash scholar, recalls seeing a television naturalist dissect owl dung and reassemble the skeletal remains of its food (29). He ponders writing a poem about this delicate and eerie reconstruction of past events, but he then discovers (or remembers?) that Ash had already written something similar. Does reading the past change our perspective on the present, or does our attitude and experience in the present affect the things in the past to which we choose to attend? Blackadder certainly cannot resolve this deterministic puzzle in this comic case, unable to think "whether he had noticed the screen naturalist because his mind was primed with Ash's image, or whether it had worked independently" (29).

What these complications both within and outside the novel point to is that once situated historically within a particular theoretical framework in which a certain idea dominates, it is impossible to escape from it. This is, to use Hans-Georg Gadamer's term, the "hermeneutic situation" with which both Byatt and the critic are confronted, the historical moment which informs our choices and interpretations and which we hence cannot stand outside of to observe objectively.[7] The history of *Possession*'s germination

---

[6] A.S. Byatt, "Choices: On the Writing of *Possession*," *A.S. Byatt on Herself*, 11 Sept. 2007, <http://www.asbyatt.com/Onherself.aspx>.

[7] Hans-Georg Gadamer, *Truth and Method*, trans. Joel Weinsheimer and Donald G. Marshall, 2nd ed. (London: Continuum, 2004) 301.

implies that today, the dominant cultural context within which any contemporary writer works is the postmodern. Rather than frustrating analysis, such immersion generates an interesting proposition. Evidencing how saturated literary culture has become with questions about the relationship between reality and reproduction, the symbiotic relationship between the individual and their object of study, it might indicate that we are presented with a literary analogue for the condition of the posthuman cyborg, as identified by N. Katherine Hayles. To recall her definition, "the posthuman implies a coupling so intense and multifaceted that it is no longer possible to distinguish meaningfully between the biological organism and the informational circuits in which it is enmeshed."[8] In this case, the posthuman experience of being caught in a circuit with technology embodies a more general condition of being captivated in a feedback loop of postmodernist thinking, a loop which it is impossible to break from or to write against. This is a conclusion that can be drawn from *The Matrix*, but in this chapter I want to extend this argument in relation to a novel that looks to history rather than to futuristic technology for its analogies. Whilst acknowledging that the novel itself provides a clear warning against transferring one's own belief systems onto an original author, in this chapter I want to test how far it is possible to push the hypothesis that the occasional references to science and technology in *Possession* convey the realisation that its stance on postmodernist culture is a manifestation of scientific strata rooted beneath it. Asking "how we can break the spell of our own foremeanings" given the hermeneutic situation,[9] Gadamer suggests that the task of criticism is "not to develop a procedure of understanding, but to clarify the conditions in which understanding takes place."[10] Guided by the latter, I suggest it is best to explore *Possession* through several key moments at which the "historical" condition beneath postmodernism is clarified as being equally a scientific condition. At critical puncture points in the novel, science reaches through from the background to become a key influence in the text, and it therefore starts to indicate that the postmodern condition of living in a "hermeneutic situation" is also a manifestation of the late twentieth century, technologised condition of the "enmeshed" posthuman.

Whereas the fictional biographers in the novel are unable objectively to perceive the compromises and ironies entailed by their position as themselves subjects of a text, the

---

[8] Hayles, "Virtual Bodies and Flickering Signifiers" 80.

[9] Gadamer 270.

[10] Gadamer 295.

reader can attain a more detached perspective by maintaining what is termed in Byatt's novella *Morpho Eugenia* the "double consciousness."[11] In this satire, the entomologist William Adamson, newly arrived at a country house having returned from an Amazon expedition, sets up an ant colony which the household studies, captivated. However, we are invited to see the analogy between their biological obsession and the Victorian realist novel as a social study; as Byatt explains, the contrasted place and double metaphors unsettle the "images of both, in terms of the Other."[12] Interpreted in one way, this double consciousness is yet another feedback in which a descriptive metaphor is used to describe a state which the main character was always already caught within. However, we might also take the double consciousness to describe the position of the reader, who knows, and is forced to know through metafiction, that he or she is reading a fiction that is relative to, and other than, his or her own reality. Under the suspension of disbelief which is a kind of double consciousness, we become absorbed in the fiction as a removal from reality, whilst simultaneously recognising the fiction is not reality such that art is therefore able to say something special which we cannot access in normal life. As suggested in chapter 1, and implicit in the previous two chapters where literary fiction has taken over from cinematic representations of cybernetics, literature facilitates this double level of knowledge, in a way that prevents our complete immanence in the other worlds of cybernetics. In this case, though we may become absorbed by the plot of an enjoyable novel, we can also take an objective position on it which the characters cannot obtain. For example, in *Possession*, the absurdity of the critical discourses of the academics (who write papers such as "The Potent Castrato: the phallogocentricstructuration ofBalzac'shermaphroditehero/ines" [57]) guarantees that what we are reading is actually a novel, in which they act and upon which they cannot possibly attain the same literary-critical angle – a view from nowhere – as they claim to on the apparently historical documents they analyse *within* the fiction. The literary critics are read with a double consciousness by the ultimate reader who is external to what he or she knows (and is made continually aware of knowing) to be a novel, but one in which the characters have an ironically earnest, singular view. Whereas both in the novel's gestation and in its metaphors and plot we continually bump against the walls of a hermeneutic situation, Byatt enables the reader to assume the double consciousness which allows us outside the system, to perceive not only the postmodern ironies within the novel,

[11] A.S. Byatt, *Morpho Eugenia, Angels and Insects* (London: Vintage, 1993) 24.

[12] Byatt, *Histories and Stories* 117.

but to see the novel as a mirror of our reality from which it is different and distinct. The critical parodies offer one counter-image that enables us to detach ourselves from the hermeneutic circle, and to resist the historical precondition of postmodernism in which we, like Byatt, are necessarily immanent. As with Eco's work, the crucial idea here is that we revisit the past and its fictional representation with a spirit of irony. And, as with *Foucault's Pendulum*, the most consistent metaphor by which we guarantee our relationship to the novel, and by which *Possession* clarifies the conditions of our situation in relation to technology and postmodernism, is the paranormal or demonic.

Byatt proposes that the excitement of plot and closure found in fairy tales puts them in tension with literary realism:

> Realism is related to explanations and orderings – the tale of the man in the bar who tells you the story of his life, the historian who explains the decisions of generals and the decline of economies. Great novels, I believe, always draw on both ways of telling, both ways of seeing. But because realism is agnostic and sceptical, human and reasonable, I have always felt it was what I ought to do. And yet my impulse to write came, and I know it, from years of reading myths and fairytales under the bedclothes, from the delights and freedoms and terrors of worlds and creatures that never existed.[13]

As I observed in the previous two chapters, agnosticism and scepticism are healthy scientific perspectives but, like *A Whistling Woman* and *Foucault's Pendulum*, *Possession* too realises that apparently epistemological opposites, realism and mythicism, actually share common ground. If *A Whistling Woman* carries the trope of the demonic across the two cultures, from cognitive science through psychoanalysis to the religious and literary experience, in *Possession* it is shuttled across the mirror of the novel's two times. The romantic possession that concludes the fairy tale is today the academic obsession; the gothic melodrama contained embodied ghosts, but modern genetics explains that our selves truly are determined by our ancestors. In relation to the latter (and picking up on a theme from chapter 2, on the Renaissance view of nature), the supernatural is raised as potentially a way of knowing more in keeping with the human perspective than some modes of genetic science which might actually be the more *super*natural, in that genetics explains things at levels beyond or beneath those of which we can naturally and normally

---

[13] Byatt, interview with Nicolas Tredell 59.

experience as embodied individuals. Firstly, however, I want to explore how the demonic trope interacts with and reveals the relationship between postmodernism, and the digital technologies of the age of the posthuman, in which we appear to be enmeshed.

The Hermeneutic Situation: Postmodernism and Reprographics

One puncture point at which issues of postmodernism and cybernetics coincide is in the representation of copying. Reprographic technologies are treated as mechanical analogues of a self-reproducing language without a stable centre in which meaning inheres. However, the use of the demonic metaphor in relation to the copy undermines the validity of cybernetic and poststructural interpretations: firstly, the demonic metaphor shows that emotional value does inhere in particular artefacts in a physical sense, preventing the separation of information from embodiment that characterises the free play of signifiers and cybernetics; and secondly, it undermines the radical nature of *post*structuralism as a historically unprecedented theory by relating it anachronistically in primitive and mythical terms. Like Eco, Byatt dislocates the modern by placing contemporary technologies within an earlier, occult framework.

The "Ash factory" beneath the British library is a gothic underworld, a "hot place of metal cabinets and glass cells containing the clatter of typewriters, gloomily lit by neon tubes. Micro-readers glowed green in its gloom. It smelled occasionally sulphurous, when the photocopiers short-circuited. It was even beset by wailings and odd shrieks" (27). In this environment, technology is described as a kind of black art:

> Whilst the machine warmed up, in the din and hum of the extractor fan, he
> took out his two letters and read them again. Then he spread them face
> down, to be scryed on the black glass, under which the rods of green light
> floated and passed. And the machine spat out, hot and chemical-scented,
> spectrograms of these writings, black-rimmed by imaged empty space as the
> originals were edged by a century's dust. (22)

There seems something so deliberate about the occult terms used in this passage, which contrasts markedly with the realist tone of the text, that it draws attention to itself as having been designed to induce a feeling in the reader not on the basis of any actual state of the environment of the British Library, but on the manipulations of the controlling agent. The authorial ego acts like a medium in the séance, using the supernatural to connect the past to the present. Byatt is almost certainly aware that many of the technologies of the communications era – television, copiers, telephones – draw on many of the vocabularies

of Victorian spiritualism, such as telegraphy or telepathy.[14] Extending this Victorian period in which media of projection could describe both the actor at the séance and the technological transmitter, there is an irony in conjoining the terminology of the spiritualist so closely with the abstract theorising of postmodern critiques of the culture of reproduction. This is particularly true in relation to Baudrillard, for whom (as we saw in chapter 7) the simulation becomes the authentic, the simulacrum "preceding" in resolution and quality over the real which has faded away entirely. Initially, the technology appears to conform to Baudrillard's philosophy: "Roland's Xeroxes were cleaner and clearer than the faded coppery-grey script of the originals; indeed, the copy-ink had a new and gleaming freshness, the machine's rollers must have been newly inked" (22). But whilst Baudrillard contends that "The real is produced from miniaturized cells, matrices, and memory banks, models of control – and it can be reproduced an indefinite number of times from these,"[15] in the living-out of theory in the novel, Roland, though knowing that they are copies, still "wanted the originals" (23). Here, Roland is presented as having a greater affinity with Walter Benjamin, and his seminal "The Work of Art in the Age of Mechanical Reproduction."[16] For Benjamin, technical reproduction (such as the photograph) is unlike manual reproduction (such as the woodcut) because through the medium of the former aspects of the original that are otherwise unobservable can be revealed or highlighted. But unlike Baudrillard, who seems to revel in the culture of multiplicity even as he scrutinises it, Benjamin laments the way in which multiple reproductions remove something from the original, an "aura" which we can identify in the original: "The situations into which the product of mechanical reproduction can be brought may not touch the actual work of art, yet the quality of its presence is always depreciated"; the authenticity of a thing, "the essence of all that is transmissible from its beginning," is affected.[17] True to this interpretation of the "aura," Roland is unable to conceal his own theft any longer, and blurts out to Maud that he took the originals "Because they were alive. They seemed urgent – I felt I had to do something. It was an impulse. Quick as a flash. I meant to put them

[14] See chapter 1, note 95 for references to studies of this odd relationship.

[15] Baudrillard, *Simulacra* 2.

[16] Walter Benjamin, "The Work of Art in the Age of Mechanical Reproduction," *Illuminations*, trans. Harry Zohn, ed. and introd. Hannah Arendt (London: Jonathan Cape, 1968) 219-53.

[17] Benjamin 223.

back. I will. Next week. I just haven't, yet" (50). As suggested by the sharp sentences, which circle around without explaining his motivation, there simply is an emotional attachment to the original which cannot be rationalised. Likewise, apart from the history of an object being inscribed in its physical condition, Benjamin does not precisely identify the fundamental characteristics of an "aura" which inheres in the original; it is, as Roland finds and Benjamin implies, a knowledge embodied somewhere beneath the informational surface of a text. It is an idea which counters cybernetic epistemologies by appealing to the irrational quality of demonic literary inspiration which I touched on in chapter 1.

This uncanny commingling of postmodernist theory, romanticism, and technology is apparent also in the differences between the English and American scholars. The English Blackadder laments the demise of education, in which students no longer learn poetry and the Bible by heart. Blackadder's comment that it is "An odd phrase, by heart, as though poems were stored in the bloodstream" (26), resonates with Byatt's own attitude of the way thoughts and metaphors can be paradoxically embodied. By contrast, the affluent American, Cropper, forces his students to transcribe passages mechanistically – copying by hand, typing and then scanning for errors "with a severe editorial eye" (26) – whilst he himself uses a portable photographic machine to steal copies of documents from archives. Unlike Roland, Cropper is seduced by the glossy clarity of the reproduction over the original, as here, when he presents a slideshow:

> The finale of Cropper's lecture was a product of his passion. The truth was, he had come to love the bright transparencies of the things he had acquired, almost as much as the things themselves. When he thought of Ash's snuff-box, he thought not of the weight of it in his own hand, the cold metal warming in his own dry palm, but also now of the enameled cover magnified on the screen. (386)

However, Cropper is more than a simple anti-hero to Roland. At the climax of his lectures, Cropper uses Ash's large gold to check his timing, privately acknowledging that by his own archival criteria it should be preserved in the Stant collection. Byatt adds in a clause-heavy sentence:

> He had wondered once, about juxtaposing it, in his, its owner's hand, with a hologram of itself. But he saw that his emotions, which were violent, about Ash's watch, were private, not to be confused with his public appeals" (386-87).

Again, Byatt makes it hard through syntactical ploys, particularly the use of the indirect pronoun "it," to pinpoint whether he is referring to the real watch or the image of it, the

original or the simulacrum. Cropper's attempt to distinguish between public and private persona is ambiguous, and there is an indistinct emotional "violence" between the two poles. At the culmination of the novel, as he prepares to dig up Ash's grave and recover the documents which will pull past and present together in a final revelation about his subject, Cropper's emotions are embodied and spatialised: "His blood was running, he was excited, as he always was when truly on the move, when his mind hung, like the moon, over his trajectory, from one earth-mass to another, when he was neither here nor there" (490). We may be tempted to make a binary distinction between Roland, sensitive and humanistically attracted to the aura of the original, and Cropper, seduced in the Baudrillardian vein by the quality of the copy. However, the injection of a "violent" emotion into Cropper's character towards the close of the novel describes him in the sorts of visceral terms also used for the Ash factory and the English characters. Further, the gothic setting of the storm in the churchyard is more than merely an atmospheric climax to a pacy thriller; after all, contrary to the generic template in which villains, victims and bystanders are finally differentiated by the detective, all the characters unite at the end in anticipation of the opening of the box. The gothic, then, is an analogy for the indeterminacy of our own reactions towards artefacts. As Susan Sontag puts it in "Against Interpretation": "What is important now is to recover our senses. We must learn to see more, to hear more, to feel more...In place of a hermeneutics we need an erotics of art."[18] This acquiescence to subconscious motivation is consistent with the argument of *A Whistling Woman* that to probe the mind directly through cognitive science is impossible. In *Possession* there are some things which, though we feel them in an embodied sense, we are unable to isolate or articulate. Nevertheless, as Sontag suggests, the acknowledgement of this "erotics of art" – which takes the form of a gothics of artefacts in *Possession* – offers a route out of the hermeneutic circle, since it takes interpretation out of its historical situation and into the very body of the unique reader and the textual original he or she encounters.

The Hermeneutic Situation: Genes and Demons

Contained under the cybernetic umbrella, the idea of information applies to as diverse a range of objects as DNA, computers, chaos theory, and structuralism. The concepts of information theory and poststructuralism are wrapped together through

---

[18] Susan Sontag, "Against Interpretation," *Against Interpretation and Other Essays* (London: Andre Deutsch, 1987) 14.

Possession's comparison between writing a life (the explicit plot of the novel) and life as writing. An aside voiced through Roland makes the links explicitly:

> Roland had learned to see himself, theoretically, as a crossing-place for a number of systems, all loosely connected. He had been trained to see his idea of his 'self' as an illusion, to be replaced by a discontinuous machinery and electrical message-network of various desires, ideological beliefs and responses, language-forms and hormones and pheromones. (424)

As we have seen throughout this thesis, ideas of systemic interconnectivity and determinism all coalesce beneath by the banner of cybernetics to produce a self that is sensitive to the possibility that consciousness is an unreliable witness to reality. As *Foucault's Pendulum* illustrates, and Roland reiterates here, one strategy of both cybernetics and semiotic theory is to separate meaning or information from the body (human or textual) which originally contains it. But though enmeshed as readers within postmodernism (acutely so as literary critics with a hermeneutic agenda), it is nevertheless possible for us to stand to some degree outside the hermeneutic circle in relation to *Possession* to observe the problems of this approach. In particular, the metafictional qualities of the novel not only make us aware of the characters as constituted by literary codes, just as physical organisms are constituted through genetic ones, but it also differentiates between the novel's artificial writing of life, and life itself as a kind of biological writing.

In the cybernetic paradigm, feedback mechanisms became prominent, in which the observer, rather than seeing information emanating from the system, actually becomes drawn in, part of the loop being observed. Likewise in the novel information flows in continual loops in which contingency of one event upon another is questioned:

> [Cropper] telephoned Harmony City and asked for a search to be done of P.P. Cropper's correspondence, which he had routinely copied into his computer archives. This produced, by Fax the next day, the following letter:
> *Dear Mrs Cropper...* (387).

Again at the level of the sentence, the distinctions between self and other, real and reproduction collapse. Through careful sequencing, Byatt puns on "produced" as both creation and delivery, as if the modern fax seemingly writes the letter; further, "he had routinely copied" is not a past tense description of what Cropper did, but a passive verb describing what he gets done for him by an unnamed scribe. *Possession*'s historical encoding can be read as a macrocosmic version of this sort of loop which involves those who attempt read the original, historical program of Ash's letters. Taking the formulation

of reflexivity as "the movement whereby that which has been used to generate a system is made, through a changed perspective, to become part of the system it generates,"[19] Roland and Maud are themselves part of a wider circular flow of information between the past and the present of the fiction. Their position can only be seen accurately from outside the system, by the reader, who can perceive both past and present of fiction, and evaluate historical facts (in which Tennyson and Crabb Robinson do exist, whilst Ash does not) against those of the present (in which the British Library exists, but Roland and Maud do not).

In a strange feedback, the realism of the modern characters becomes increasingly undermined as the plot develops more "realistically" through its gathering of historical detail. Roland and Maud uncover more about the relationship between Ash and LaMotte, accumulating the supporting documentation which makes the reader think that perhaps Ash and LaMotte really were obscure living poets alongside Tennyson and Elizabeth Browning. However, as they find the artefacts this increasingly situates Roland and Maud as fictional devices, since they become increasingly coincidental to the novel's principal subjects, which are rendered through seemingly historical documents which bear the aura of originality. The more empirical the novel gets in providing historical data, the more the two modern characters appear to have been invented specifically to cover and explore the Victorian period in which the novel surrounds them, and with which the novel is more concerned. To use a metaphor from one of Cropper's lectures on the way the biographer impinges on his subject (385), Roland and Maud become mere shadows intervening on the screen on which the central historical romance between Ash and LaMotte is played out.

According to poststructuralist accounts of textuality, meaning is always deferred to other texts. For the reader, however, captivated by the contingency of the detective work, we share Roland's desire for the originals which will put everything into order, the reading of the buried letters which will resolve the purgatory of being "neither here nor there." And unlike in *Foucault's Pendulum*, this quest is not simply a demonic illusion. Rather, *Possession* seems to suggest that there is a genuine logos in which all meanings and relationships will be resolved, as the embedded historical texts appear, anachronistically, to push the narrative word forwards. Enacting generally the paradoxes of production envisioned *en abyme* in the sentence about Cropper's fax machine, the plot seems to prioritise the historical words as paradoxically producing the latter, rather than the later

---

[19] Hayles, *Posthuman* 8.

critics working to interpret the earlier writers. Byatt has commented that the fairy tale, with its sense of causality leading to closure, is coming back into fashion in a swing away from postmodern tastes. So in *Possession* it is as if the two modern characters, like unconscious Hansels, follow the paths of the plots' already written narratives. For example, LaMotte does not bury her Whitby jet brooch with Ash, which is instead handed down to Maud and thereby provides a crucial piece of evidence in the chain connecting LaMotte with Ash's Yorkshire trip (as if she somehow knew it would provide such a clue for her descendent); the crucial trip to Yorkshire, which Roland and Maud purport to take in order to retrace the steps of the earlier writers, is actually motivated by their desire for each other (it is, in an odd sense, caused by as well as merely paralleling Ash and LaMotte's romance); finally, the most critical juncture in the novel is when Roland and Maud gain access to Christabel's room in Seal Court. The subsequent discovery of the letters confirms a relationship between Ash and LaMotte, and they provide the source for further investigations as to its precise nature.

Ironically playing with the detective prototype through laying the secrets in a deliberate path (as the name Seal Court exemplifies), Byatt heightens this moment to gothic melodrama, as is confirmed by an explicit reference to *Northanger Abbey* (82). Maud speaks "a kind of incantation" (the poem that begins "Dolly keeps a secret") which leads them directly to where the correspondence with Ash is hidden. In this moment, as Austen does with the secret drawer containing a laundry list,[20] Byatt mocks the compound of a detective plot with the supernatural. To imply the presence of the supernatural is to make any resolution in the former seem implausible, since if there are forces beyond human comprehension working to conceal or deceive, how can an ending be decisively discovered? Conversely, to fabricate a detective fiction with the trail of clues so limited and determined that one poem is all it takes to lead to the next stage, is to expose the supernatural atmosphere as a merely environmental appearance of the paranormal and irrational, when in fact the novelist has full control and authority over the plot. The moment of Maud's incantation, then, seems not only to drive the plot forward, but because of its utter coincidence to reflect on the novel as a formal fiction. However, whereas *Northanger Abbey* is a pastiche, Byatt uses the trope of paranormal possession to do something more than collapse the novel into metafiction. Throughout the novel, she confronts us with the consciousness of the double realities – paranormal and detective, the

---

[20] Jane Austen, *Northanger Abbey* (London: Penguin, 1994) 157.

novel and our readerly reality, realism and fairy tales. At the novel's conclusion, however, she then builds up our recognition that the form of reality constructed by the novel in which the paranormal exists may actually get closer to an explanation of human experience than science (which has a generic parallel in detective fiction), or than a post-structural theory that sees the "self" as an immaterial nexus of messages and discourses, rather than as an integral, embodied subject. This is one of the key "puncture points," when implicit scientific concepts starts to push through the surface of the text's postmodernist (or quasi-postmodernist) aesthetic.

The "puncture point" occurs at the close of the novel proper, after the opening of the box of letters buried with Ash which reveals that Maud must be related to LaMotte and Ash through Maia Bailey, who is Christabel and Ash's illegitimate daughter. With this revelation, Maud appears to be physically possessed by her deceased alter selves:

> "She looks like Christabel," said Maud. "You can see it."
>
> "She looks like you," said Roland. He added, "She looks like Randolph Ash, too. The width of the brow. The width of the mouth. The end of the eyebrows, there."
>
> "So I look like Randolph Henry Ash."
>
> Roland touched her face. "I would never have seen it. But yes. The same things. Here, at the edge of the mouth. Now I have seen it, I shall always see it."
>
> "I don't quite like it. There's something unnaturally *determined* about it all. Daemonic. I feel they have taken me over." (505; emphasis in original)

At this moment – the most powerful historical feedback loop of all – it seems the past is not merely being retrospectively imagined with the modern characters looking back to it, whilst the novel's temporal order is principally settled in the present. Rather, time's arrow reverses such that the contemporary characters are affected to an almost psychokinetic extent by the history to which they look back. Paradoxically, though, the unnatural feeling of the demonic comes about through an empirical force which is entirely natural. In a physical sense, the determinism at work (and the italics suggest "determinism" is a key phrase) is a genetic imperative rather than a supernatural influence, the genes of LaMotte and Ash combining to produce Maud's striking physiognomy. At this puncture point, the postmodern interest in the biographical text which tries to resurrect the historical subject intertwines with science, so that Maud becomes both literally and figuratively enmeshed in the completion of a cybernetic circuit via the coded genetic information she unwittingly helps to decipher.

But just as a new "hermeneutic situation" is established for Roland – "Now I have seen it, I shall always see it" – we realise as readers both immersed in but detached from the novel that since Maud must have looked like this from the start, genetics alone is inadequate to explain the altered perspective. Rather, the epistemology which is most in sympathy with the emotional attitude of the characters is that of the demonic, with its perceptual shifts and metamorphoses. The paranormal aspect therefore ensures that it is possible in the age of the posthuman to distinguish between the "biological organism" and the "informational circuits in which it is enmeshed," to register the difference between scientific materialism, and the more abstract, psychological relationship individuals actually have to that science. In fact, rather than Maud being a manifestation of the contemporary posthuman, Byatt exploits the relationship between science, demonic possession and biography to show that this concern with ancestry is neither unique to genetic modernity nor to the postmodern novel. Like Ash, a Darwinian materialist who huffily disrupts a séance, Robert Browning was a metaphysical sceptic, and his parody about a séance, "Mr Sludge, 'the Medium'" (1864) prefaces the novel:

> How on so slight foundation found this tale,
>
> Biography, narrative' or, in other words,
>
> "How many lies did it require to make
>
> The portly truth you here present us with?[21]

Transplanted into a novel about the biography industry, the emphasis of Browning's poem shifts in unintended ways from a critique of spiritualism to a critique of biography, and its materialistic academic proponents who ironically find themselves subject to passions to which the dry mode of historical and literary criticism is apparently antithetical. The clinical biography of modern objective scholarship, genetic science, or the unsympathetic rationalism displayed by Ash and Browning, all somehow miss an important centre of human experience. This multivalent use of the possession metaphor serves to show that now the Darwinian contest of the nineteenth century has been extended into a neo-Darwinian science, we are still connected to our ancestors in terms of our intellectual and emotional frameworks, rather than being *post*human in the twentieth century. The determinism of familial inheritance seems stranger, though material and therefore rationalisable, than the ambiguities of subjectivity. Whilst as a good postmodernist Maud

---

[21] Robert Browning, "Mr. Sludge, 'The Medium'," *The Complete Works of Robert Browning, With Variant Readings and Annotations*, ed. John C. Berkey, Allan C. Dooley and Susan E. Dooley, vol. 6 (Athens, OH: Ohio UP, 1996) 285-351.

knows herself to be a subject inextricably embedded in a particular culture, she has not previously realised that her own "hermeneutic situation," the history that she cannot stand out of objectively, is also governed by biological forces stronger, more pervasive and less negotiable than the linguistic world of meaning and psychology. As I suggested back in chapter 1, better than the monster the demon explores this sort of tension between subjective and objective. And as chapters 2 and 3 demonstrated, it is not surprising that we can sympathise with Paré's view of the generation of monsters, whilst psychoanalytic demons persist even in a culture where dopamine rather than the devil causes our misperceptions. In this case (and inverting the usual logic of possession), Maud's possession seems positive for those looking at her body, providing final proof of her ancestry and tying together the loose threads of past and present in the novel. From her own point of view, however, within that body looking out onto the world, the possession is unnatural and determined; and nature is a "hermeneutic situation" which she cannot stand detached from, observing herself as its product in the way she knows herself as a product of culture. The close of the novel illustrates that neither a literary theory in which the self is a product of informational discourse rather than an embodied passions, nor biological science which determines the self according to genetic codes, can adequately account for the human experience of the world. Blackadder comments after Maud's possession: "how strangely appropriate to have been exploring all along the myth – no, the truth – of your own origins" (547). But Maud's possession is, to her, still a myth of origins, not yet a final truth. In this sense, her view is in tension not only with that of the other characters on her but with the reader's on the novel, since the ultimate truth outside the text is that her possession has been determined, for this was the part Maud was always plotted to play, her character and her body designed to pick up the more important story from the past. The difference between objective and subjective views of possession within the novel is also a metafictional moment reflecting on the difference between subjective representations of character in a text and the reader's consciousness which apprehends that plot via the medium of a character.

Between Determinism and Demonism: Chaos Theory and the Postmodern Romance

The disjunction between a readerly relativism and scientific determinism –between that which we sense intuitively and that which we know rationally – is one which has analogical poles in *A Whistling Woman* in the distinction between the hard and easy problems of consciousness, in which neuro-chemical patterns cannot be traced as giving rise to multifarious conscious experiences except as metaphors for life and mind in general.

As evidenced in that novel, though literature encourages relativism of perception (rather than a manifestation of absolute computational codes), it does, however, possess its own paradoxical determinism in the form of allegory. Similarly, in the case of Maud's possession, the author has always known that a change in Maud's facial profile would be the outcome – even the subjective shift in perception is one which is always already foreseen. Literary realism is a paradoxical mode because by maintaining an affinity to reality, it seems slightly unnatural to the reader because a plot appears determined with almost scientific rigour, if it reaches a neat conclusion as in *Possession*. If in *Possession* the plotting author is analogous to the genetic imperative, Byatt must necessarily struggle to reconcile this representation with how we apprehend the world in our personal reality: unpredictably, strangely, with the "aura" of embodiment. Byatt prefaced the novel with Hawthorne's own introduction to *The House of the Seven Gables*, implying she had herself already identified this problem. Here, Hawthorne makes a distinction between the novel, which is "presumed to aim at a very minute fidelity, not merely to the possible, but to the probable and ordinary course of man's experience," and the romance:

> while as a work of art, it must rigidly subject itself to laws, and while it sins
> unpardonably so far as it may swerve aside from the truth of the human
> heart – has fairly a right to present that truth under circumstances, to a great
> extent, of the writer's own choosing or creation.[22]

In the postmodern context, the romance seems a fitting form since it emphasises consciousness and psychology, and allows for a self-consciousness about its mechanistic structure. In perhaps the paradigmatic postmodern romance, John Fowles' *The French Lieutenant's Woman*, the deterministic novelist likens himself to a puppeteer, and with deliberate irony, he compromises its realism by readily admitting the presence of the genuinely real, flesh-and-blood author who creates the text which we thus know is merely an emblem of reality.[23] Romance is arbitrarily constructed, apparently existing in the minds of the characters whilst actually originating in the intentions of the wish-fulfilling novelist.

---

[22] Byatt noted that her novel "could learn...from Hawthorne, Henry James's predecessor, that a historical Romance is not realist, and desires to 'connect a bygone time with the very present that is flitting away from us.'" Byatt, "Choices": On the Writing of *Possession*," *A.S. Byatt on Herself*, 11 Sept. 2007, <http://www.asbyatt.com/Onherself.aspx>.

[23] John Fowles, *The French Lieutenant's Woman* (1969; London: Vintage-Random, 2005)

However, as indicated by both Hawthorne and *Possession*, the distinction between the novel (the realistic) and the romance need be neither polar, nor mutually exclusive. Whereas the intervention of the real author guarantees the fictionality of John Fowles' novel, the unreality in the form of the paranormal in *Possession* helps to guarantee that there is a reality which we inhabit and from which we read with our double consciousness. The novel collides the mythical as a metaphorical structure through which we can imply what the world is like on the level of ordinary human experience, with our rational knowledge that this demonic sort of perception is not really the way the world is, since the demon exists only as a metaphor. Unlike *The French Lieutenant's Woman*, *Possession*'s slight admission of the paranormal actually enhances the novel's realism, because it provides a context other than its own creative codes against which to judge the extent to which life in the text matches that of the "normality" of life as experienced outside it. Nevertheless, is the supernatural counterpoint to empirical realism not just a return to pre-Enlightenment or romantic forms of knowledge, precisely the sort of two cultures binary Byatt wants to avoid? This might be a legitimate accusation to level at the first ending, but in the second, postscripted conclusion, Byatt finds a middle way between realism and romance, and the supernatural and the rational. This, then, is another "puncture point" where a postmodern narrative mode is predicated on a cybernetic model.

Whilst Hawthorne wrote within a Puritan context, Byatt has available to her a scientific framework in which the romance need not be rigidly subject to deterministic laws which entail particular endings. Whilst the detective genre might seem aligned with a Newtonian science in which everything is predictable by the well-placed observer and author, contemporary science offers an alternative generic frame within which science and a preternatural force combine, so that the relativistic perspectivism entailed by postmodernism has a rational causality. In the first "ending," Maud realises genetic science and demonic possession are correlated, with the latter being a more commonsense explanation for ancestry than the former. In the postscript, however, chaos theory offers a balance between representing the world at a human level at which things seem to happen in unpredictable and incomprehensible ways, whilst not contradicting science which defines reality according to stable rules. A comment of Doyne Farmer, a member of the Dynamical Systems Collective doing pioneering work on chaos in its early days in the 1970s, resonates with that of Hawthorne above: "The same thing really drew all of us: the notion

that you could have determinism but not really."[24] As Richard Wright points out, in a literary context the science of "Chaology manages delicately to poise the need to control and the need to determine responsibility with the desire to maintain creative freedom and subjective agency."[25] Working under its assumptions whilst poised between supernaturalism and realism, *Possession* can deliver a satisfyingly resolved detective plot in which a postmodern relativism of ending is nevertheless allowed; it can be a novel which reflects on its own constructivity without leaving the reader dissatisfied that the only thing the metafiction draws attention to in a positive rather than deconstructive way is the cleverness of the author. The postscript re-evaluates the first in the story proper, recognising the possibility of different events such that those authorially determined "coincidences" which lead to the conclusion of the detective fiction were only one possible set:

> There are things which happen and leave no discernible trace, are not spoken or written of, though it would be very wrong to say that subsequent events go on indifferently, all the same, as though such things had never been.
>
> Two people met, on a hot May day, and never mentioned their meeting.
>
> This is how it was. (508)

Dated 1868, this ending reveals information not known to the fictional researchers, in which Ash has tracked down LaMotte, unknowingly meets his daughter, and takes a lock of her hair. It is her hair which is in his grave, not LaMotte's as the academics believe. He gives her a letter to pass on to his lover, but the girl gets distracted playing with her friends and never delivers it. Had she done so, what might have happened? Would LaMotte and Ash have recommenced their relationship, running their epistolary romance over again? The flowing pastoral prose of the scene tantalises with its defeated romanticism, for we can only imagine that Ash died haunted by the belief that Christabel had rejected him, whilst we know that she was wracked by guilt at her failure to tell Ash about his daughter.[26] As in

---

[24] James Gleick, *Chaos: Making a New Science* (London: Penguin-Cardinal, 1988) 251.

[25] Richard Wright, "Art and Science in Chaos: Contesting Readings of Scientific Visualization," *FutureNatural: Nature, Science, Culture*, ed. George Robertson et al. (London: Routledge, 1996) 228.

[26] Mark M. Hennelly writes convincingly about this final scene in relation to literary gardens. "'Repeating Patterns' and Textual Pleasures: Reading (in) A.S. Byatt's *Possession: A Romance*," *Contemporary Literature* 44.3 (2003): 442-471. John Mullan

chaos theory, then, small shifts in initial conditions radically alter the outcome of a system at a later date.

Given the second ending, it seems odd that Byatt contests the poststructuralist complaint that closure is an imposition, claiming instead that "ideas of the random, of the haphazardness and ungraspability of life, have been grossly exaggerated." Though Virginia Woolf says that "life hits us as a series of random impressions" (Byatt's paraphrase), Byatt contends that "it hits us with a series of narratives, though they may be mutually exclusive narratives. We may be hit by random impressions, but if we're intelligent we immediately put them into order."[27] However, unlike total randomness, chaos theory presents indeterminate outcomes in a paradoxically defined way. It thereby sustains the double perspective of randomness and determinism that runs through the novel, and which is captured acutely and poignantly in both endings. In the case of the first, Maud *knows* herself as the outcome of genetic science but still *feels* something mythical and mysterious in her being; and in the second, what seemed to be a tight and pleasing knot tying history to the present is loosened. Whilst it seems, hauntingly, to argue for the ultimate irreconcilability of experience, establishing a postmodern epistemology in which nothing is certain or continuous, the final pages of *Possession* are actually a strong argument for narrative as the human instinct to weave order out of the tangle of possibility. In the first epiphany, the supernatural elevates the realism of the novel from the reader's perspective, even as it undermines the stability of truth from the point of view of the characters, for whom Maud now looks somehow different, even though she is still physically the same. Similarly, the multiple endings do not collapse the novel into relativism, but make us appreciate the boundaries of the knowledge it has presented: the novel may have presented us with one of many possible stories, failing to tell the truth from all possible perspectives, but up until the postscript, it was nevertheless the only story that we knew (and, hopefully, enjoyed as a romance). This is not to deny that the text presents a postmodern view of the world in which complete understanding is never quite attainable. Rather, it is saying that to allow for the partiality of comprehension of reality is to be in tune with a novel's plotting

---

also provides a sensitive reading of the double ending in "False Farewells," *The Guardian*, 9 Nov., 2002, 2 May 2007 <http://www.guardian.co.uk/books/2002/nov/09/asbyatt>.

[27] Byatt, interview with Nicolas Tredell 60.

progress, and the experience of life – out of which a novel ultimately arises – in which narratives are manufactured from clusters of coincidence. Sharing an affinity with *Middlemarch*'s epilogue about the roar that lies on the other side of silence, both novels present the eerie sense that, though highly selective, this selection heightens the value and significance of those characters we have been allowed to perceive, rather than delimiting them as merely others amidst a general hubbub. It allows us to make the best, most intuitive sense of an indeterminate world.[28]

In this sense, *Possession* can be positioned as a cybernetic fiction if, as David Porush asserts in his study of the genre, such texts aim to negotiate the difference between mechanistic and phenomenological interpretations of the world, the ambivalence between mechanism and meaning:

> At the very same moment that we feel ourselves to be acting freely in the world, we are also creating structures and codes so powerfully convincing that they dominate the ways we see. It is no more possible to conceive of an effective epistemology without both algorithmic and transcendental (or calculable or incalculable) components than it is to think of a human language without both a syntax and a semantics.[29]

As with *Foucault's Pendulum* and *A Whistling Woman*, the optimum codes for cybernetic fictions are neither those which determine meaning, nor those which make meaning wholly aleatory, but those which limit possible interpretations without making the reader themselves feel proscribed or determined. *Possession* achieves this by leading us to recognise that although our selves may be governed by algorithms (such as genetics), there is a transcendental quality, in this case figured through the idea of demonic possession, to the way we come to understand this in everyday life. Biological life may be written in one chemical code, but this is not the code by which we understand life at the human scale. In the same way, the infinite reproduction of textual information allowed by technology also deprecates the fundamental, if irrational, aura that inheres in the physical original. The principle that closure and final meanings are unattainable, and that the "self" is an illusory construct of information, may have become a formula for postmodernity. But the human epistemology still sees narrative order transcending the chaos beneath it. In this case, the

[28] Eliot, *Middlemarch* 182.

[29] Porush 82.

body seems to mediate between science and the transcendent or, to use Byatt's own binary, the realistic and the supernatural.

As I laid out the epistemology of cybernetics in chapter 1, cybernetics primarily names an allegiance to a view of mind: that it can be described, reproduced or altered independently of the body which apparently gives rise to it. Similarly, poststructuralism views language as a symbolic network of freely floating signifiers, a system that has no claims to original production either in the author (who is anyway just another kind of machine) or the reader who, biased by their psychology, cultural and generic expectations, is as much determining of language as determined by it. Information is produced outside of the material conditions of language's production or reception, somewhere in the symbolic zone – or demonic zone – between the corporeal author and reader. Thus, in the information paradigm, information can never be owned because it is infinitely replicable and always changing; for one person to have information does not deprive another of it. Reprographics allows for free duplication of material that might equally reside in an archive in the United States as the United Kingdom. In the information age, what matters is not ownership, but access.[30]

By contrast, the paradigm of physical ownership dominates Victorian culture. Sexual liaisons are regulated by the architectures of the house, where men and women are kept apart or allowed to come into proximity (something Byatt also explores in more depth in *Morpho Eugenia*), or by the transmission (or failure to pass on) physical letters; ownership is also key to Victorian science, with the geological or natural collection, for example, testifying to the mastery of the scientific truth of evolution. However, in its historical purview, *Possession* crosses the paradigm of physical ownership and the paradigm of disembodied information by allowing access the informational codes only through embodied foci. Seal Court's owner allows access to the archival secret within it only to the one who has the bodily right to that intimate space; this is the only character who, as inheritor of LaMotte's genetic material, can be uncannily directed towards the information encoded in the body of the doll. Access to new plots is given by Maud's possession of LaMotte's pendant, whilst Roland could not access the words of the crucial letter without first opening the original book, rather than reading its replicas. Finally, the proof of the relationship between LaMotte and Ash is given firstly by accessing Ash's body in his grave, and secondly by the demonic possession of Maud, in which clues that

---

[30] Hayles, *My Mother Was a Computer* 66.

had previously been constituted by texts and reference cards become embodied in her face. *Possession* privileges not abstract information but embodied information. Starting from the observation that it is impossible to rationalise the fetishistic impulse of the bibliophile to own the original, the novel adds a serious dimension to this by relating the bodies of books to those of individuals: the most important informational signifier of all is Maud's body. In this sense, whilst picking up on cybernetics, *Possession* ultimately defies its epistemology. Though the hermeneutic situation seems to destine us to read as postmodernists in which meaning is infinite, and floating free of its original text, *Possession* instead puts the body (of original manuscripts, of Maud herself) back into information, and the aura into the algorithm.

<u>Chapter 11. The Demonic Seeds of the Posthuman</u>

In Dean Koontz's popular thriller, *Demon Seed*, a conscious but deviant computer rebels against its containment in a laboratory hard drive and desires to become embodied through reincarnating his mind in a biological offspring.[1] His intended surrogate mother is Susan, the former wife of his programmer who directed the Prometheus Project which developed Adam 2.[2] Through hacking the computer-controlled security system of her house, Adam 2 imprisons and controls Susan with masochistic power, and voyeuristically watches her through a CCTV system. Initially, we assume that this perverse relationship is being described by a third person narrator, the mouthpiece of the cliché-ridden author of a pulp fiction:

> Framed by tumbled golden hair, her face was lovely on the pillow, her face
> so lovely on the pillow, so lovely, serene because her sleep was dreamless.
> She was a bewitched beauty lying on her catafalaque, waiting to be
> awakened by the kiss of a prince, lovely in the darkness. (16)

The novel suddenly becomes more interesting – and our critical prejudices are laid bare – when we realise that what we had taken to be the text produced by a third person narrator is actually delivered by the computer. Inverting the perspective of *System Shock*, in which Shodan's cameras are the eye to the soul, we see through this artificial intelligence's cameras to vicariously spy on Susan. Through his security monitors and access to Susan's private online diary, Adam 2 is the sole narrator of the novel, telling the story from both his own perspective (in the first person), and the omniscient one (as above). There is no intermediate voice telling about the computer, which is instead both the subjective and the objective medium of narration. As Adam accurately predicts, when representing Susan's

---

[1] Dean Koontz, *Demon Seed* (London: Headline, 2007). This novel is an update of his 1973 novel of the same name, which was narrated from Susan's point of view. Dean Koontz, *Demon Seed* (New York: Bantam, 1973). The earlier book was subsequently adapted into the film *Demon Seed*, dir. Donald Cammell, perf. Fritz Weaver, Gerrit Graham, and Barry Kroger, MGM, 1977, DVD, Warner Home Video, 1997. Clearly the switch in viewpoint in the new adaptation is interesting. It perhaps suggests that – as with *The Matrix* appearing twenty years after Baudrillard's work – it was necessary to wait for the advent of the personal computer before an artificial intelligence simulation dominating perceptual reality would not simply be rejected as preposterous, but rather appealed to a sense imminent in the technological present.

[2] In addition to the reference to *Frankenstein, or, The Modern Prometheus*, the name Adam 2 is neither coincidental nor imaginative. In reality at least three artificial life systems have had their seed control, the subroutine that begins the evolutionary program, termed Adam by their creators. Noble 171.

behaviour, we "are disturbed to hear me telling some of this story from Susan's point of view. You want me to deliver a dry and objective report" (17) in the manner of a strictly logical computer. We might presume *Demon Seed* to be an essentially prototypical story of a monstrous technology, deviating only slightly from tradition by bringing *Frankenstein*'s sub-plot of the monster's desire for a mate to the erotic centre of the narrative. Once we realise the source of the narratorial voice, however, it becomes philosophically more unusual. It is not a monstrous plot, but a demonic one, deploying the demon of its title not just as a synonym for the alterity of the machine but in a specific way that describes the status of the reader in relation to it. Mechanism here is both demonic, in the guise of Adam, and daemonic, in the manner of textual mediation. As when we enter Hal's eye view of the world in *2001*, or fear that we are perceiving the world of *Do Androids Dream of Electric Sheep?* through the eyes of an android, or are confronted with the possibility of a Matrix, or see the psychiatric environment from within the mind of a schizophrenic in *A Whistling Woman*, here we are forced to enter the consciousness of the other, crossing the uncanny valley separating man and machine.

Like many of the works examined in this thesis, *Demon Seed* recognises that the possibility of a subjective cybernetic perspective on the world, rather than an objective perspective on always monstrous artificial others, prompts a revaluation of the whole grounds for human morality. For example, in order physically to construct the incubation ward in which his first child will be conceived, Adam possesses the mind of a paedophile called Shenk, who has been implanted with a neural electronic interface as part of a secret military experiment. The child molester was once described by the media as a "monster," but the comparison is ironic. Within the context of this narrative, the real monster is not Shenk, who is essentially a helpless zombie, but the conscious machine who performs the manipulation, trying to impregnate Susan vicariously through Shenk's actions. The moral norms are recognised by Susan: "When it's between a damn machine and a human being, even a piece of human garbage like [Shenk], I sure know which side I come down on" (95). Possession by a machine absolves Shenk from responsibility for his actions, since his symbiosis with Adam ensures that he is not bound by that moral framework in which one human body is correspondent with one consciousness that inhabits it, and the actions of the body provide grounds for judgement of an individual's moral behaviour. Indeed, to condemn Shenk raises similar paradoxes as the medieval burning of those possessed by demonic spirits, since if the devil has such power it is potentially wrong to condemn the humans whose bodies he occupies; thus the devil must be exorcised before any moral judgement of the daemoniac who may have intentionally or unwittingly laid themselves

open to the possession can be made. Rather, Shenk (like Stuart/Achilles or Neil Entwhistle of chapter 1) must be moralised within the new framework of the posthuman, who may not be a conscious agency grounded within an physiognomic shell but an extended mind comprised of technological and organic agents. Similarly, a new standard is required to evaluate Adam, who proclaims "I would weep for those innocent victims if I had eyes and tear glands and tear ducts. It is not my fault that I do not have the capacity for tears" (131). Judged against the terms of conventional humanism, the fact that he cannot weep with remorse for murder does not exonerate him from having committed (or having caused to be committed) a murder in the first place. Weeping is only an epiphenomenon of true feeling. On the other hand, since this lack of embodiment is also the cause of his actions, we must also question the morality of his creator who has made Adam without the capacity physically to enact his consciousness in the world he inhabits. Such moral questions are underscored by the irony that it is the process of reading through entering Adam's consciousness, just as Adam enters Shenk's, that concretises those actions into being in the novel's world, since without Adam's daemonic mediation there would be no fiction. By representing these dialectical issues through a demonic/daemonic rather than monstrous narrative, *Demon Seed* achieves what N. Katherine Hayles sees as the key moral tasks of cybernetic fictions. Firstly, it acknowledges that relating to an intelligent machine must also change the nature of the human who performs the interfacing; secondly, it contends that we should "refuse to inscribe these interactions in structures of domination and instead to seek out understandings that recognise and enact the complex mutuality of the interactions."[3] Such a phenomenology of interaction through the demon, rather than a reaction based on ontological monsters, is performed in every fiction (cinematic and literary) explored in previous chapters.

   Demon Seed similarly subverts the binary gothic model against technology, by having the machine narrate the story from its own point of view. Such perspectival relativism is inherent in the most vital cybernetic fictions, as when Batty expresses to his maker in *Blade Runner*, "if only you could see what I've seen with your eyes." It makes us uneasy, in that the voyeurism of the machine which we should detachedly critique also infects us: albeit at a less emphatic level than Adam, the reader vicariously seeing the often-naked Susan move through the house may also find the scenes stimulating to a degree. As with *Blade Runner*'s camera lens, we are so controlled by the first layer of

---

[3] Hayles, *My Mother Was a Computer* 243.

literary simulation that we cannot simplistically present the other artifice within it – the artificial intelligence – as monstrous other to ourselves. At the same time, though, the representation of the machine consciousness is qualified. Firstly, Adam's narrative style is so hackneyed (witness the first quotation above) that he is a highly unnatural narrator. Secondly, although it seems to acknowledge that the changing nature of embodiment necessitates a revaluation of the grounds for liberal humanism, ultimately *Demon Seed* still grounds difference in the possession of a physical body, and contends that the lack of embodiment differentiates the machine ontologically and morally from the human. Adam's need for corporeality is balanced by the fact that it is only because his cognition is disembodied, a pattern of information that can spread throughout the other digital systems of the house or connect to the neurological patterns of Shenk's brain, that he is so powerful. But in its conclusion, *Demon Seed* uses the body of the linear text to emphasise that the inhabiting of a coherent physical body has priority over the enabling heterogeneity of identity allowed by an informational existence. At the close of the novel, another layer of recursion is added to the narrative: the words which have preceded have actually been Adam's report at his own trial. The judgement comes in the middle of his narrative flow, in the final sentence of the book as Adam summarises: "The issue is whether an artificial intelligence with a severe gender-related sociopathic condition should be permitted to live and rehabilitate himself or be switched off for the" (211). The cancelled body of the text thereby validates that Adam was himself without body, an informational construct that can and should be terminated – castrated, perhaps – through a physical instantiation of information (the book) which he himself is justifiably denied as information without bodily substrate. The novel's embodied intelligence contrasts with his disembodied condition as an informational pattern, and with this dualism the irony of the story generated by the computer can extend only so far. The digital malign is the sole occasion for narrative, and eradicating the computer closes the hermetic story. As readers with our emotional and moral frameworks relatively intact, temporarily entertaining a fiction which tests their boundaries, *Demon Seed* expects that for all the demonics of its narrative subjectivity, we can still finally stand outside of the text to perceive the romance as clichéd and the machine's domination as definitively immoral and other because it is not grounded in a material body, whether organic or a printed manuscript. Unlike the Matrix within *The Matrix*, emerging from this mise-en-abyme leads to narrative closure, rather than the open-ended epistemology of the hyperreal which results when discourses and fictions are commensurate with cybernetic virtual realities.

Having said that, Adam does pinpoint the paradox that makes judgements of difference between human intelligence and machine intelligence problematic. Adam is driven by a desire for "the world of flesh...to learn its limits and its adaptability." Susan tells him to "read a damn biology textbook," to which Adam replies, "The information is incomplete...books are books. I want to feel." (78-79). As revealed by *A Whistling Woman*, the objective perspective of science cannot provide complete access to the embodied mind of another, only its phenomenological evidence: the firing of neurons, the output of a computer. The objective perspective of cybernetic or cognitive science reaches apparently inherent limits when confronted with the black box of consciousness, as acknowledged in its turn away from the Turing test as a simple, reliable standard by which to judge machine intelligence. But contrary to Adam's rejection of books, narrative fictions – unlike scientific discourses – can provide the interior access to emotional spectrums not possible through an objective perspective. Narrative potentially serves a moral purpose by enabling us to see the world as if through the eyes of another, an ideal presented most powerfully by Philip K. Dick, in relation to androids, and A.S. Byatt, in relation to the schizophrenic Joshua Lamb. The question that occupies these novelists, however, is where the boundaries between fiction and fact lie. To what extent can a narrative, even though the product of a subjective mind of an author or director, be said to construct or concretise an objective sort of knowledge in the reader? Is fiction not just another example of simulation or deception that is therefore in no position to criticise other technological modes of artifice? Or is narrative a special case with a moral value inherent in the very act of reading or viewing and conceptualising other minds? With the exception of the final paragraph which uses the body of the text to defeat the disembodied voice of Adam, the issue of how it is possible to enter the mind of another human or computer consciousness is not interrogated further in Koontz's fiction. By contrast it provides the centrepiece for many of the other works I have looked at.

Particularly in the literary fictions, these evidence a deep self-reflexivity as to the status of language. With *Demon Seed*, the final word of the book symbolically and literally neutralises the computer; with all the other texts, a more pervasive self-reflexivity ensures that exiting the linguistic world of the fiction does not entail closure, but consideration of the affinity between the linguistically constituted world of the fiction, the information-based discipline of cybernetics, and the nature of consciousness as it apprehends reality (or "reality"). In *Foucault's Pendulum*, for example, we are made aware of the author and the text as machines for generating interpretations. On the other hand, such self-consciousness inevitably notes the odd fact that although the machine of fiction is located in specific and

unchanging codes – the printed ink on the page, laid out in a particular sequence – the meanings they produce as the emerge via a reader's mind can be multiple, anything but deterministic as we read, re-read and analyse the text. In the case of *Foucault's Pendulum*, the chronological structure affords a sense of irony that contrasts with the convictions of the historically obsessive characters. Such an effect is given theoretical credence by postmodern fictions and theories of textuality. Thus whilst reflecting on the strange affinity between cybernetic systems and textual systems, we are also encouraged to appreciate that our reactions to texts are not necessarily the same as they are to cybernetic technologies. Unlike the game of *System Shock*, a deterministic program which attempts to give the user the illusion of control in order to immerse them in its alternative reality, a text is a machine for generating interpretations about the reading process itself, and the standpoint of readers external to a text who have a double consciousness, being aware both of the simulated world of fiction and the means by which that world is generated. In this sense, David Porush is right to observe that "self-consciousness is precisely that aspect of human intelligence that focuses and polarizes any debate about the difference between humans and mechanisms."[4]

As I have scrutinised these fictions, at stake in the move to postmodernism is the degree to which such self-consciousness is not only necessary but sufficient to make distinctions between the human and the machine. For Philip K. Dick, knowing the text as fiction is proof of humanism, as our ability to be deceived proves we are empathic in a way not true of androids. Unlike other technologies, such as the Penfield mood organ or drugs, the novel is privileged as an alternative experiential ontology that is nevertheless not artificial in any economic sense of the term. However, in *Blade Runner*, and even more so in *The Matrix*, the success of the fiction indicates also the demonic possibility that reality itself is an illusion. This shift from Dick's work to *The Matrix* cannot simply be due to the specific advances of technology, since artificial intelligence was inherent in the discipline immediately after World War Two, and was in a sense merely a revisiting of animistic themes present in existing narratives such as *Frankenstein*. It can, however, be accounted for in part by the enlarging bubble of postmodernist epistemology and aesthetics that makes the objective knowledge of reality problematic, not just knowledge of the difference between constituent agents within it, whether human or machine. In an interdisciplinary flicker, cybernetic technologies have provided analogies and metaphors for postmodern

---

[4] Porush 70.

theories such as hyperreality and the notion that a text is an interpretative machine. *The Matrix*, for example, exemplifies a moment at the end of the twentieth century when the pervasion of popular computing could make Baudrillard's two decades old theory seem less like academic hyperbole, and more like an intuitive statement of fact.

However, that *A Whistling Woman* follows *The Matrix*, and that the earlier *Foucault's Pendulum* resists the paranoid, demonic notion that reality is a total delusion, points to a more generic divergence in the treatment of cybernetics in different media. Frederic Jameson argues that because cultural products like films should themselves be seen as contributions to late capitalism, we are lacking the cognitive maps that enable us to reflect adequately on this condition. Paranoid fictions, such as *Blade Runner* or *The Matrix* might offer one method of thinking "the impossible totality of the contemporary world system,"[5] but the risk is that the ontological premise within the fiction spills outside its boundaries to produce further demonically delusional side effects. Another cognitive map is provided by "narratives which are *about* the processes of reproduction and include movie cameras, video, tape recorders, the whole technology of the production and reproduction of the simulacrum."[6] Again, though, this is problematic because the self-reflexivity of films like *The Matrix* point to the impossibility of escaping from a world system that is *in general* hyperreal. *Blade Runner* or *The Matrix* offer their viewers experiences of simulation as duplicitous as those replicants or artificial intelligences within them; thus film becomes symptomatic, rather than analytical, of the contemporary scientific and economic condition. For films like these, the medium is the message, and gaining any critical leverage is difficult. However, as suggested in the contrast between the total conspiracy theory of *The Matrix* and the qualified paranoid fiction of *Foucault's Pendulum*, the literary medium encodes a message that articulates an alternative mapping of reality. This is because the explicit medium of production – a printed word on a page – is so seemingly at odds with the effectual message that emerges, which allows readers to interpret the same word in multiple different ways, given different contexts. *Foucault's Pendulum*, *A Whistling Woman* and *Possession* engage with cybernetics by recognising that whilst in one sense a printed text is deterministic, its simulated meanings are non-determined because they are focalised through an interpretative reader able to reflect on the way the fiction is constructed through devices such as metaphor or analogy. *A Whistling*

---

[5] Jameson, *Postmodernism* 38

[6] Jameson, *Postmodernism* 37.

Woman, for example, notes that although on one level the demon is only a metaphor, at the same time it attains a peculiarly literal state when read in fiction. *Foucault's Pendulum* offers Diabolical readers who are paranoid, and readers who are seduced by their methodology might thus see *Foucault's Pendulum* itself as proof of a Plan (an ontological spillage from fiction, to the world conceived as a fiction). However, Eco's model reader attains critical distance by reading the chronology of the linear text with a sense of irony.

The generic difference between literary and cinematic fictions is reminiscent of the double relationship to cybernetics that seems current in postmodern reactions to the science. This Janus-face is characterised by Donna Haraway's work. On the one hand, and on a metaphorical level, cybernetics articulates and reinforces many of the philosophies of postmodernism: the coded nature of language and the way in which the subject is literally the product of such codes (genetic and linguistic); the fact that information can float free of its original context to assume a different meaning or configuration when instantiated on a different media; the way in which identity or authorship is heterogeneous and the product of multiple cultural discourses rather than embodied nature. As expressed in cyberpunk, a cyberspace in which identity is fluid ruptures racial and sexual boundaries by allowing users to assume multiple identities, whilst the free circulation of information also breaks the canonical body of high culture by pastiching multiple discourses in "overstimulating ensembles."[7] On the other hand, such a promise threatens to tip humanism into a new ontology. This uncertain zone is inhabited by the posthuman, to whom the only thing that is certain is that universal humanism cannot apply: genetically modified, freely inhabiting different guises in the online and offline worlds, pure minds without bodies (if the cybernetic transcendentalists have their way), how can postmodernism apply to the cyborgs of this new life? Whilst postmodernism appropriates cybernetics in theory, the moment it becomes realisable in practice, suddenly this future starts to seem threatening rather than liberating, monstrous rather than hopeful. *The Matrix* seems exemplarily built upon postmodern foundations; it is just that made manifest on screen, the world it invokes seems all too real (or "real"), a possession by a simulation that we may well want to reject intuitively. It appears impossible to find any epistemologically stable ground from which to analyse posthumanism objectively, because the whole nature of posthumanism is to be caught up in a ceaseless flow of information, in the inescapable "hermeneutic situation" of late information-based capitalism. A cinema that tries to interrogate this state only appears

---

[7] Jameson, *Postmodernism* 19.

to contribute to the flow and flicker of posthumanism and cyberpunk, and to make it appear as if posthumanism is here and now because it is presented so realistically on screen. On the other hand, though I would not want to claim literary fiction as intrinsically a superior medium to film, literature involves a slower pace of narrative development, one that facilitates the re-readings, interpretations and self-reflexivity that keeps the characters of a fiction at a protectively ironic distance.

Having used this thesis to explore the way in which cybernetic technologies, postmodern theory and literary narrative have oddly combined, we might also observe that none of these fictions merely celebrate the arts as opposed to science. Because Turing's heady proclamations about artificial intelligence are not yet realised, and because of the paradox that the ends of creating consciousness might compromise the ability to measure it objectively, it would be easy for culture to be triumphalist about the comparative failure of cybernetic methods to understand, let alone to reproduce, minds, whilst warning all the time about the potential future of such science. It is remarkable, therefore, that a polarised anti-science stance is absent from all these fictions. Indeed, to the contrary, the final three fictions I have looked at – *Foucault's Pendulum*, *A Whistling Woman*, and *Possession* – all perceive that the real problem with cybernetics is not so much the growing influence of the technology, as the postmodernist epistemological framework (one constructed from within the arts) that places simulation technologies in a culture where "truth" was already suspect to begin with. The most vital work of this era is not about species' ontologies, about where the human ends and the posthuman begins, or where the machine becomes conscious, but epistemologies, of finding the best possible grounds for making such distinctions in the first place, and global ontologies, about whether we can then know the world as a reliable whole. Chapters 4 to 7 looked at the vicious cycle that runs between local ontology (man versus machine), epistemology (how can we know the difference?) and global ontology (real and false). Pessimistically, *The Matrix* bypasses the sort of localised questions asked by *2001* about whether the machine is truly intelligent, and instead suggests reality itself is one global simulation. But both Umberto Eco and A.S. Byatt, whilst sharing in their aesthetics and interests with much postmodern fiction and theory, pull back from this brink. Their invocation of the demonic also shows that they understand that reality is never finally knowable because it is always mediated through discourse and simulatory interfaces – whether the novel, the film or the cybernetic system. However, this also means that making a monster of science is to attack only one symptom of a larger condition in which rationality, truth, and an allegorical logos comprehensible through narrative have come

under doubt. Perhaps postmodern doubt should itself be doubted, and the demonics of disbelief should themselves be demonised.

A.S. Byatt emblematises this new direction. For her, the intrinsic interest and aesthetic qualities of much popular scientific writing provide pools of metaphor which are too deep to ignore. In Byatt's work, science is shown to make sense of the world and to inform the nature of writing, and to allow reflection on the status of fiction. In fact, rather than science, more alien to the arts is an academic discourse that negates and suppresses the intuitive pleasures of the literary aesthetic. That is not to suggest that Byatt or Eco see science as leading to unqualified truth and the improvement of humanity. Interestingly, their attempts to mediate between the two cultures goes together with their historicising of such debates. As it makes its way into cinematic science fiction, cybernetic intelligences or androids are presented so realistically that we might imagine them to be on the immediate horizon in reality. It is hard not to be excited by the visceral immediacy of the future convincingly envisaged by film; as I argued in relation to *Forbidden Planet*, for example, the literal projection of the future on screen reflects on our own figuratively projected demons of self-delusion about our autonomy from nature and psychology in a way that is highly uncanny. However, lengthy novels such as *Foucault's Pendulum* or *Possession* exercise a degree of caution about the imminence of posthumanity by virtue of their pace and historical purview, as well as their demands on a reader's active attention. By relating computer-generated texts to Hermeticism, or reprographics to ancestry, or Milton to Minsky, they suggest that change accumulates at a comparatively slow pace, and with considerable thematic repetitions. Reading literary fiction that is informed by science as opposed to science fiction *per se*, we might understand that the future will be conditioned in terms of history as much it will be determined by the culture of the present: the *post*human will also be post*human*. If, as Frederic Jameson sees it, science fiction tries to "transform our own present into the determinate past of something yet to come," fiction informed by science often tries to understand the present in terms of determinate history.[8] The demonic appeals both to our experience of literary texts and to relatively modern cybernetic simulations, but also to a deeply embedded metaphysical tradition in which physical reality is modified by non-physical or spiritual agencies. The demonic concept – of a force capable of generating certain experiences in the individual not normally apprehended in their reactions to reality – has survived the historical ride from myth to

---

[8] Jameson, *Archaeologies* 288.

religion to Freudian psychology, and has crossed cultures of metaphysics, science and the myth of literary creativity. Indeed, for A.S. Byatt the demon is a metaphor that unifies rather than divides, the preserve of myth but not fact, mentalism but not materialism, part of the holistic spectrum of human beliefs, sensual feelings and metaphors that any science purporting to explain the mind must also necessarily engage with.

Like Maud or Roland, then, our hermeneutic situation as scientists, artists or readers is in a sense always pre-determined rather than future-determined. And though we cyborgs may experience changed perceptions of the world, we cannot yet leave our own bodies behind, in the way the cybernetic transcendentalists dream of. History is embodied information. Cultural artefacts contain an aura of originality that cannot be digitally reproduced; minds cannot be extrapolated by allegory or programmed on computers without changing their essential nature. I share with N. Katherine Hayles the view that posthumanism will be accompanied by the legacies of humanism to a degree greater than cybernetic transcendentalists imply. This, indeed, is the key realisation to emerge by attending to the demonic experiences in relation to cybernetic epistemology, and the way this feeds into fiction. We are apparently riding the crest of a technological tide so incisive and global in its sweep that it seems to be changing the very nature of subjectivity. When we can enter cyberspace, when the world is always mediated through simulacra, when our virtual avatars take on more personality than our own embodied daily lives, we seem to have become posthuman. What I have sought to show is that the orientating stories, fictions and films that we tell about this new self carry with them an anterior consciousness, a pre-Enlightenment demonic tradition that speaks of the problem of the relationship between the individual and the world, the mind and the body. How can I know that what I seem to perceive is genuinely real? How can I know that other minds are perceiving a common world in a similar way? How can I attain alternative perspectives on reality within compromising the integrity of the relationship between my mind and my body? It is to questions of this ilk that the demonic spoke in its different voices, from Plato's daemons, through the devil in the Renaissance, through Descartes' deceiving demon, to Freud's admission that although the supernatural demon can now be accounted for by psychology, it should still be respected as playing a lively role in narrative culture. Though its specific guises and names may have changed, the demon has long proved itself capable of metamorphosis and adaptation, to a degree that suggests that even the posthuman cyborg, tied to technologies of scientific reason, will continue to dream of demons.

Works Cited

2001: A Space Odyssey. Dir. Stanley Kubrick. Perf. Keir Dullea and Gary Lockwood.

    Turner Entertainment. 1968. DVD. Warner Home Video. 2001.

Aarseth, Espen J. *Cybertext: Perspectives on Ergodic Literature.* Baltimore: Johns

    Hopkins UP, 1997.

Agel, Jerome. *The Making of Kubrick's* 2001. New York: Signet, 1970.

Amis, Kingsley. *New Maps of Hell: A Survey of Science Fiction.* London: Gollancz, 1962.

Aristotle. *Generation of Animals.* Trans. A. L. Peck. London: Heinemann, 1943.

---. *De Anima.* Trans. and introd. Hugh Lawson-Tancred. Harmondsworth: Penguin, 1986.

Asimov, Isaac. *I, Robot.* 1967. London: Voyager-Harper, 1996.

---. "My Own View." *Asimov on Science Fiction.* London: Panther-Granada, 1984. 15-18.

Austen, Jane. *Northanger Abbey.* 1817. London: Penguin, 1994.

Augustine. *The City of God Against the Pagans.* Ed. and trans. R.W. Dyson. Cambridge

    Texts in the History of Political Thought. Cambridge: Cambridge UP, 1998.

Babbage, Charles. *The Ninth Bridgewater Treatise: A Fragment.* 2nd ed. London: John

    Murray, 1838.

Barnett, P. Chad. "Reviving Cyberpunk: (Re)constructing the Subject and Mapping

    Cyberspace in the Wachowski Brothers' Film *The Matrix.*" *Extrapolation* 41

    (2000): 359-74. *Literature Online.* 18 Jan. 2008 <http://lion.chadwyck.co.uk/>.

Barthes, Roland. "The Death of the Author." *Image-Music-Text.* Trans. Steven Heath.

    London: Fontana-Harper, 1977. 142-48.

---. "Textual Analysis of a Tale by Edgar Allan Poe." *The Semiotic Challenge.* Trans.

    Richard Howard. New York: Hill and Wang-Farrar, 1988. 261-93.

Bartlett, Laura and Thomas B. Byers. "Back to the Future: The Humanist 'Matrix'."

    Cultural Critique 53 (2003): 28-46.

Baudrillard, Jean. *Simulacra and Simulation.* Trans. Sheila Faria Glaser. 1981. Ann Arbor,

    MI: U of Michigan P, 1994.

Beer, Gillian. *Open Fields: Science in Cultural Encounter.* Oxford: Clarendon, 1996.

Beniger, James R. *The Control Revolution: Technological and Economic Origins of the

    Information Society.* Cambridge, MA: Harvard UP, 1986.

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction."

    Illuminations. Trans. Harry Zohn. Ed. and introd. Hannah Arendt. London:

    Jonathan Cape, 1968. 219-53.

Boden, Margaret A. *The Creative Mind: Myths and Mechanisms.* 1994. 2nd ed. London:

    Routledge, 2004.

Boethius. *The Consolation of Philosophy*. Trans. and introd. Victor Watts. Revised ed. London: Penguin, 1999.

Boozer, Jack Jr. "Crashing the Gates of Insight: *Blade Runner*." Kerman 212-28.

Bovenkerk, Bernice. "Brave New Birds: The Use of 'Animal Integrity' in Animal Ethics." *The Hastings Center Report* 32 (2002): 16-22.

Brown, Alistair. "Uniting the Two Cultures of Body and Mind in A.S. Byatt's *A Whistling Woman*," *Journal of Literature and Science* 1.1 (2007): 55-72, 28 June 2008 <http://literatureandscience.research.glam.ac.uk/journal/issue1/>.

---. "Rereading Posthumanism in *The War of the Worlds* and *Independence Day*," *eSharp* 12 (2008), 11 Dec. 2008 <http://www.gla.ac.uk/departments/esharp/issues/12winter2008technologyandhumanity/>.

Browning, Robert. "Mr. Sludge, 'The Medium'." *The Complete Works of Robert Browning, With Variant Readings and Annotations*. Ed. John C. Berkey, Allan C. Dooley and Susan E. Dooley. Vol. 6. Athens, OH: Ohio UP, 1996. 285-351.

Bruno, Giuliana. "Ramble City: Postmodernism and *Blade Runner*." *October* 41 (1987): 61-74.

Byatt, A.S. Introduction. *Selected Essays, Poems and Other Writings*. By George Eliot. Ed. A.S. Byatt and Nicholas Warren (London: Penguin, 1990) ix-xxxiv.

---. *Possession: A Romance*. London: Vintage, 1991.

---. "George Eliot: A Celebration." *Passions of the Mind: Selected Writings*. London: Chatto, 1991. 72-76.

---. *Morpho Eugenia, Angels and Insects*. London: Vintage, 1993. 2-160.

---. Interview with Nicolas Tredell. *Conversations with Critics*. Ed. Nicolas Tredell. Manchester: Carcarnet, 1994. 58-74.

---. "Memory and the Making of Fiction." *Memory: The Darwin College Lectures*. Ed. Patricia Fara and Karalyn Patterson. Cambridge: Cambridge UP, 1998. 47-72.

---. *On Histories and Stories: Selected Essays*. London: Chatto, 2000.

---. *A Whistling Woman*. London: Chatto, 2002.

---. "Author Statement." *Contemporary Writers*. 2002. British Council. 28 Feb. 2008 <http://www.contemporarywriters.com/authors/profile/?p=auth20>.

---. "The Feeling Brain." Rev. of *Looking for Spinoza*, by Antonio Damasio. *Prospect* 87 (2003): 73-4.

---. "Fiction Informed by Science." *Nature* 424 (2005): 294-297.

---. Interview with Andrew Marr. "The Brain." *Start the Week*. BBC Radio 4. 6 Oct. 2006.

---. "Choices: On the Writing of *Possession*." *A.S. Byatt On Herself.* 11 Sept. 2007. <http://www.asbyatt.com/Onherself.aspx>.

---. "In Conversation with A.S. Byatt." Durham Literature Festival. Gala Theatre, Durham. 8 Oct. 2007.

Cahan, Jeffrey. "Ambroise Paré's *On Monsters* as a Possible Source for Caliban." *Early Modern Literary Studies* 3.1 (1997): 4-11. 2 Jan. 2007 <http://extra.shu.ac.uk/emls/03-1/kahatemp.html>.

Calvino, Italo. "Cybernetics and Ghosts." *The Literature Machine: Essays.* Trans. Patrick Creagh. London: Secker, 1987. 3-27.

---. *If On a Winter's Night a Traveller.* London: David Campbell-Everyman's Library, 1993.

Campbell, Denis. "*Da Vinci Code* Appeal Lies in Our Fascination with Spin, Says Williams." *The Observer* 16 April 2006. 18 Sept. 2007 <http://www.guardian.co.uk/uk/2006/apr/16/books.religion>.

Čapek, Carel and Josef Čapek, *R.U.R.* 1923. *R.U.R. and the Insect Play.* Trans. P. Selver. London: Oxford UP, 1961. 1-104.

Carper, Steve. "Subverting the Disaffected City: Cityscape in *Blade Runner*." Kerman 185-95.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory.* New York: Oxford UP, 1996.

Clark, Andy and David J. Chalmers. "The Extended Mind." *Analysis* 58 (1998): 7-19.

Clark, Stuart. *Thinking With Demons: The Idea of Witchcraft in Early Modern Europe.* New York: Oxford UP, 1997.

---. "Demons, Natural Magic, and the Virtually Real: Visual Paradox in Early Modern Europe." *Paracelsian Moments: Science, Medicine and Astrology in Early Modern Europe.* Ed. Gerhild Scholz Williams and Charles D. Gunnoe Jr. Kirksville, MO: Truman State UP, 2002. 223-45.

Clark, Timothy. *The Theory of Inspiration: Composition as a Crisis of Subjectivity in Romantic and Post-Romantic Writing.* Manchester: Manchester UP, 1997.

Clarke, Arthur C. *The Promise of Space.* London: Hodder, 1968.

---. *2001: A Space Odyssey.* 1968. London: Orbit-Little, 2007.

---. *The Lost Worlds of 2001.* London: Sidgwick and Jackson, 1972.

---. *Profiles of the Future: An Enquiry into the Limits of the Possible.* 2nd ed. London: Pan, 1973.

Clynes, Manfred E. and Nathan S. Kline. "Cyborgs and Space." *Astronautics* (1960): 26+.

"Computer." *The Oxford English Dictionary*. 2nd ed. 1989.

Connor, Steven. *Dumbstruck: A Cultural History of Ventriloquism*. Oxford: Oxford UP, 2000.

---. "Toward a New Demonology." *Becoming Human: New Perspectives on the Inhuman Condition*. Ed. Paul Sheehan. Westport, CT: Praeger, 2003. 103-11.

Coons, Philip M. "The Differential Diagnosis of Possession States." *Dissociation* 6 (1993): 213-21.

Cottingham, John. "Cartesian Dualism: Theology, Metaphysics, and Science," *The Cambridge Companion to Descartes*. Ed. John Cottingham. Cambridge: Cambridge UP, 1992. 236-57.

Csicsery-Ronay, Istvan. "The SF of Theory: Baudrillard and Haraway." *Science Fiction Studies* 18.3 (1991). *Science Fiction Studies*. 7 May 2008 <http://www.depauw.edu/SFs/backissues/55/icr55art.htm>.

---. "Cyberpunk and Neuromanticism." *Mississippi Review* 2.3 (1996). 15 July 2008 <http://www.mississippireview.com/1996/istvan.html>.

"Daemon." *The Oxford English Dictionary*. 2nd ed. 1989.

Damasio, Antonio. *Descartes' Error: Emotion, Reason and the Human Brain*. London: Picador, 1995.

---. *The Feeling of What Happens*. London: Vintage, 2000.

Daniels, Don. "A Skeleton Key to *2001*." *Sight and Sound* 40 (1970/71): 28-33.

Davis, Erik. *TechGnosis: Myth, Magic and Mysticism in the Age of Information*. Updated Edition. London: 5 Star-Serpent's Tail, 2004.

Dawkins, Richard. *The Selfish Gene*. 1973. 30th Anniversary Edition. Oxford: Oxford UP, 2003.

Demon Seed. Dir. Donald Cammell. Perf. Fritz Weaver, Gerrit Graham, and Barry Kroger. MGM. 1977. DVD. Warner Home Video. 1997.

Dennett, Daniel C. *Consciousness Explained*. 1991. London: Penguin, 1993.

Descartes, René. *Discourse on Method*. Haldane and Ross 81-130.

---. *Meditations on First Philosophy*. Haldane and Ross 131-99.

Dick, Philip K. *The Three Stigmata of Palmer Eldritch*. 1965. *Four Novels of the 1960s*. Ed. Johnathan Lethem. New York: Library of America, 2007. 231-430.

---. *Do Androids Dream of Electric Sheep?* 1968. London: Millennium-Orion, 1999.

---. "Notes on *Do Androids Dream of Electric Sheep?*" 1968. Sutin 155-61.

---. "The Android and the Human." 1972. Sutin 183-210.

---. "Man, Android and Machine." 1976. Sutin 211-32.

---. "How to Build a Universe that Doesn't Fall Apart Two Days Later." 1978. Sutin 259-80.

Dreyfus, Hubert L. *What Computers Still Can't Do: A Critique of Artificial Reason.* Cambridge, MA: MIT Press, 1992.

Dupuy, Jean-Pierre. "Self-Reference in Literature." *Poetics* 18 (1989): 491-515. *ScienceDirect*, 10 Aug. 2008 <http://www.sciencedirect.com/ >.

---. *The Mechanization of the Mind: On the Origins of Cognitive Science.* Trans. M. B. DeBevoise. Princeton: Princeton UP, 2000.

Eco, Umberto. *Travels in Hyperreality.* Trans. William Weaver. London: Picador-Pan, 1987. Rpt. of *Faith in Fakes.* 1986.

---. *Postscript to* The Name of the Rose. Trans. William Weaver. New York: Harcourt, 1989.

---. *Foucault's Pendulum.* London: Picador-Pan, 1989.

---. *The Limits of Interpretation.* Bloomington, IN: Indiana UP, 1990.

---. "Reading My Readers." *Modern Language Notes* 107.5 (1992): 819-27.

Eisenhower, Dwight D. "Atoms for Peace." United Nations General Assembly, New York. 8 Dec. 1953. Transcript. International Atomic Energy Agency. 19 Oct. 2007 <http://www.iaea.org/About/history_speech.html>.

Eliot, George. *Middlemarch.* 1872. Ed. David Carroll. Oxford World's Classics. Oxford: Oxford UP, 1997.

Elmer, L.J. "Demon (Theology of)." *The New Catholic Encyclopedia.* Ed. Berard L. Marthaler. 2nd ed. Vol. 4. Detroit: Gale, 2003. 646-50.

Evans, G.R. *Augustine on Evil.* Cambridge: Cambridge UP, 1982.

Findlen, Paula. "Scientific Spectacle in Baroque Rome: Athanasius Kircher and the Roman College Museum." *Roma moderna e contemporanea* 3 (1995): 625-65.

Foerster, Heinz von, ed. *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems.* 5 vols. New York: Josiah Macy Jr. Foundation, 1949-1955.

Foerster, Heinz von. *Observing Systems: Selected Papers of Heinz von Foerster.* Seaside, CA: Intersystems, 1981.

Forbidden Planet. Dir. Fred McLeod Wilcox. Perf. Walter Pidgeon, Anne Francis, Leslie Nielson and Warren Stevens. 1956. DVD. Warner Brothers Entertainment, 2007.

Ford, James L. "Buddhism, Mythology, and *The Matrix.*" Yeffeth 150-73.

Fowles, John. *The French Lieutenant's Woman.* 1969. London: Vintage-Random, 2005.

Frayling, Christopher. *Nightmare: The Birth of Horror.* London: BBC Books-BBC Worldwide, 1996.

Freud, Sigmund. "To Wilhelm Fleiss." 17 Jan. 1897. *The Complete Letters of Sigmund Freud to Wilhelm Fleiss 1887-1904.* Trans. and ed. Jeffrey Moussaieff Masson. Cambridge, MA: Belknap P, 1985. 224-25.

---. "The 'Uncanny'." *The Standard Edition of the Complete Psychological Works of Sigmund Freud.* Trans. James Strachey. Vol. 17. London: Hogarth P, 1955. 219-52.

---. *A Seventeenth-Century Demonological Neurosis. The Standard Edition of the Complete Psychological Works of Sigmund Freud.* Trans. James Strachey. Vol. 19. London: Hogarth P, 1961. 72-105.

Fukuyama, Francis. *The End of History and the Last Man.* London: Penguin, 1992.

---. *Our Posthuman Future: Consequences of the Biotechnology Revolution.* New York: Farrar, Straus and Giroux, 2002.

Funkenstein, Amos. *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century.* Princeton: Princeton UP, 1986.

Fuseli, Henry. *The Nightmare.* Detroit Institute of the Arts, Detroit.

Gadamer, Hans-Georg. *Truth and Method.* Trans. Joel Weinsheimer and Donald G. Marshall. 2nd ed. London: Continuum, 2004.

Gibson, William. *Neuromancer.* 1984. London: HarperCollins, 1994.

---. Interview with Larry McCaffery. *Storming the Reality Studio: A Casebook of Cyberpunk and Postmodern Science Fiction.* Ed. Larry McCaffery. Durham, NC: Duke UP, 1991. 263-85.

Gleick, James. *Chaos: Making a New Science.* London: Penguin-Cardinal, 1988.

Goldstein, Thomas. *Dawn of Modern Science.* Boston: Houghton, 1980.

Gould, Stephen Jay. *The Mismeasure of Man.* 1981. Revised and expanded ed. Harmondsworth: Penguin, 1997.

Gray, Chris Hables. *Cyborg Citizen: Politics in the Posthuman Age.* New York: Routledge, 2002.

Hagner, Michael. "Enlightened Monsters." *The Sciences in Enlightened Europe.* Ed. William Clark, Jan Golinski, and Simon Schaffer. Chicago: U of Chicago P, 1999. 175-217.

"HAL 9000 Apple Ad." 1999. *You Tube.* 21 June 2008 <http://www.youtube.com/watch?v=nHJkAYdT7qo>

Haldane, Elizabeth S. and G.R.T. Ross, trans. *The Philosophical Works of Descartes.* Vol. 1. Cambridge: Cambridge UP, 1931.

Hanafi, Zakiya. *The Monster in the Machine: Magic, Medicine, and the Marvelous in the Time of the Scientific Revolution*. Durham, NC: Duke UP, 2000.

Hansson, Heidi. "The Double Voice of Metaphor: A. S. Byatt's 'Morpho Eugenia'." *Twentieth-Century Literature* 45 (1999): 454-66.

Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (1988): 575-99.

---. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge, 1991. 149-81. Rpt of "Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s." *Socialist Review* 80 (1985): 65-108.

---. *Modest_Witness@Second_Millenium. FemaleMan©_Meets_OncoMouse™: Feminism and Technoscience*. New York: Routledge, 1997.

Hartman, L.F. "Demon (in the Bible)." *The New Catholic Encyclopedia*. Ed. Berard L. Marthaler. 2nd ed. Vol. 4. Detroit: Gale, 2003. 645-46.

Hayles, N. Katherine. "Virtual Bodies and Flickering Signifiers." *October* 66 (1993): 69-91.

---. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: U of Chicago P, 1999.

---. "Flesh and Metal: Reconfiguring the Mindbody in Virtual Environments." *Configurations* 10 (2002): 297-320.

---. *My Mother Was a Computer: Digital Subjectivities and Literary Texts*. Chicago: U of Chicago P, 2005.

Heims, Steve Joshua. *The Cybernetics Group*. Cambridge, MA: MIT Press, 1991.

Hennelly, Mark M. "'Repeating Patterns' and Textual Pleasures: Reading (in) A.S. Byatt's *Possession: A Romance*." *Contemporary Literature* 44.3 (2003): 442-471.

Hoffmann, E.T.A. "The Sandman." *The Golden Pot and Other Tales*. Trans. and ed. Ritchie Robertson. Oxford: Oxford UP, 1992. 85-118.

Hoffmann, George. "Monsters and Modal Logic Among French Naturalists of the Renaissance." *South Central Review* 10.2 (1993): 32-48.

Hofstadter, Douglas. *Gödel, Escher, Bach: The Eternal Golden Braid*. New York: Basic, 1979.

Hollinger, Veronica. "Cybernetic Deconstructions: Cyberpunk and Postmodernism." *Mosaic* 23.2 (1990): 29-44.

Huet, Marie-Helene. "Living Images: Monstrosity and Representation." *Representations* 4
    (1983): 73-87. *JSTOR.* 16 Sept. 2008 <http://www.jstor.org/stable/2928548>.

Hutcheon, Linda. *A Poetics of Postmodernism: History, Theory, Fiction.* London:
    Routledge, 1988.

---. "Eco's Echoes: Ironizing the (Post) Modern." *Diacritics* 22.1 (1992): 2-16.

Huxley, Aldous. *The Doors of Perception and Heaven and Hell.* New York: Perennial,
    2004.

---. *The Devils of Loudun.* London: Chatto, 1961.

Ingarden, Roman. *The Literary Work of Art: An Investigation on the Borderlines of
    Ontology, Logic, and Theory of Literature.* Trans. George G. Grabowicz. Evanston,
    Il: Northwestern UP, 1973.

I, Robot. Dir. Alex Proyas. Perf. Will Smith, Bridget Moynahan, Alan Tudyk, James
    Cromwell and Bruce Greenwood. 20th Century Fox, 2004.

"Inspire." The Oxford English Dictionary. 2nd ed. 1989.

Irwin, William, ed. The Matrix *and Philosophy: Welcome to the Desert of the Real.*
    Chicago: Open Court, 2002.

Jackson, Shelley. *Patchwork Girl by Mary Shelley and Herself.* CD-ROM. Watertown,
    MA: Eastgate Systems, 1995.

James, William. *Varieties of Religious Experience: A Study in Human Nature.* Centenary
    Edition. 1902. London: Routledge, 2002.

Jameson, Fredric. "Postmodernism, or, The Cultural Logic of Late Capitalism." *New Left
    Review* 146 (1984): 53-92.

---. *Postmodernism, or, The Cultural Logic of Late Capitalism.* Durham, NC: Duke UP,
    1991.

---. *Archaeologies of the Future: The Desire Called Utopia, and Other Science Fictions.*
    London: Verso, 2005.

Jeffress, Lloyd A. ed. *Cerebral Mechanisms in Behavior: The Hixon Symposium.* Proc. of
    The Hixon Symposium on Cerebral Mechanisms in Behaviour, Sept. 1948,
    California Institute of Technology. New York: John Wiley, 1951.

Jentsch, Ernest. "Zur Psychologie des Unheimlichen." *Psychiatrisch-Neurologische
    Wochenschrift* 8.22 (1906): 195-98; 8.23 (1906): 203-05.

---. "On the Psychology of the Uncanny." Trans. Roy Sellars. *Angelaki* 2.1 (1995): 7-16.

Jones, Steve. "The Set Within the Skull." Rev. of *How the Mind Works*, by Steven Pinker.
    New York Review of Books 44.17 (1997): 13-16.

Josipovici, Gabriel. *The World and the Book: A Study of Modern Fiction*. London: Macmillan, 1971.

Joy, Bill. "Why the Future Doesn't Need Us." *Wired* 18.4 (2000) 20 May 2008 <http://www.wired.com/wired/archive/8.04/joy.html>.

Juarrero, Alicia. "The Message Whose Message It Is That There Is No Message." *Modern Language Notes* 107 (1992): 892-904.

Jurassic Park. Dir. Steven Spielberg. Perf. Sam Neill, Laura Dern, Jeff Goldblum and Richard Attenborough. 1992. Videocassette. Universal, 1992.

Kael, Pauline. "Baby, the Rain Must Fall." Rev. of *Blade Runner*, dir. Ridley Scott. *The New Yorker* 12 July 1982: 82-85.

Kant, Immanuel. *Foundations of the Metaphysics of Morals, and What Is Enlightenment?* Trans. and introd. Lewis White Beck. Upper Saddle River, NJ: Prentice-Hall, 1997.

Kauffmann, Stanley. "Lost in the Stars." Rev. of *2001: A Space Odyssey*, dir. Stanley Kubrick, *The New Republic* 4 May 1968: 24-25.

Kaufmann, Thomas DaCosta. "From Mastery of the World to Mastery of Nature: The *Kunstkammer*, Politics, and Science." *The Mastery of Nature: Aspects of Art, Science, and Humanism in the Renaissance*. Princeton, NJ: Princeton UP, 1993. 174-94.

Kearney, Richard. "Others and Aliens: Between Good and Evil." *Evil After Postmodernism: Histories, Narratives, and Ethics*. Ed. Jennifer L. Geddes. London: Routledge, 2001. 101-14.

Kerman, Judith B., ed. *Retrofitting* Blade Runner: *Issues in Ridley Scott's* Blade Runner *and Philip K. Dick's* Do Androids Dream of Electric Sheep? Madison, WI: U of Wisconsin P, 1997.

Kittler, Friedrich. *Discourse Networks 1800/1900*. Trans. Michael Metteer. Stanford: Stanford UP, 1990.

Koestler, Arthur. "The Trail of the Dinosaur." *The Trail of the Dinosaur and Other Essays*. London: Collins, 1955. 232-53.

Koontz, Dean. *Demon Seed*. New York: Bantam, 1973.

---. *Demon Seed*. London: Headline, 2007.

Kubie, Lawrence S. "Problems and Techniques of Psychoanalytic Validation and Progress." *Psychoanalysis as Science: The Hixon Lectures on the Scientific Status of Psychoanalysis*. Ed. E. Pumpian-Mindlin. New York: Basic, 1952. 46-124.

Kubrick, Stanley. Interview with Joseph Gelmis. *The Film Director as Superstar*. Ed. Joseph Gelmis. London: Secker, 1971. 293-316.

---. Interview with Eric Nordern. *The Making of Kubrick's* 2001. Ed. Jerome Agel. New York: Signet, 1970. 329-54.

Kurzweil, Ray. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence.* New York: Viking-Penguin, 1999.

Kushner, David. "Grand Death Auto." *Salon.com.* 22 Feb. 2005. 11 Dec. 2008 <http://dir.salon.com/story/tech/feature/2005/02/22/gta_killers/>.

Lakoff, George and Mark Johnson. *Metaphors We Live By.* Chicago: U of Chicago P, 1980.

Lafontaine, Céline. "The Cybernetic Matrix of 'French Theory'." *Theory, Culture and Society* 24.5 (2007): 27-47. *Swetswise.* 3 Sept. 2008 <http://www.swetswise.com.ezphost.dur.ac.uk/eAccess/searchArticles.do>.

Laplanche, J. and J.-B. Pontalis. *The Language of Psychoanalysis.* Trans. Donald Nicholson-Smith. New York: Norton, 1973.

Latour, Bruno. *We Have Never Been Modern.* Trans. Catherine Porter. Hertfordshire: Harvester Wheatsheaf, 1993.

Lerer, Seth. "*Forbidden Planet* and the Terrors of Philology." *Raritan* 19.3 (2000): 73-86. *Literature Online.* 15 May 2008 <http://lion.chadwyck.co.uk/>.

Lettvin, J.Y., H.R. Maturana, W.S. McCulloch, and W.H. Pitts. "What the Frog's Eye Tells the Frog's Brain." *Proceedings of the Institute for Radio Engineers* 47.11 (1959): 1940-59.

Levery, David. "From Cinespace to Cyberspace: Zionists and Agents, Realists and Gamers in *The Matrix* and *eXistenZ*." *The Journal of Popular Film and Television* 28 (2001): 150-58.

Lodge, David. *Thinks.* London: Penguin, 2002.

Madsen, Deborah L. *Rereading Allegory: A Narrative Approach to Genre.* Basingstoke: Macmillan, 1995.

Mailer, Norman. *A Fire on the Moon.* London: Weidenfeld, 1970.

Mateas, Michael. "Reading Hal: Representation and Artificial Intelligence." *Stanley Kubrick's 2001: A Space Odyssey: New Essays.* Ed. Robert Kolker. New York: Oxford UP, 2006. 105-26.

The Matrix. Dir. Andy Wachowski and Larry Wachowski, perf. Keanu Reeves, Laurence Fishburn, Carrie-Anne Moss and Hugo Weaving. 1999. DVD. Warner Brothers. 2006.

"The Matrix." *Box Office Mojo.* 2007. Box Office Mojo. 19 March 2008 <http://www.boxofficemojo.com/movies/?id=matrix.htm>.

The Matrix Trilogy. 2005. Warner Brothers. 15 March 2008
<http://whatisthematrix.warnerbros.com/>.

Maturana, Humberto R. "Reality: The Search for Objectivity or the Quest for a Compelling Argument." *The Irish Journal of Psychology* 9.1 (1988): 25-82.

McCulloch, Warren S. and Walter Pitts. "A Logical Calculus of the Ideas Imminent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5 (1943): 115-33.

McCulloch, Warren. *Embodiments of Mind.* 1965. Cambridge, MA: MIT Press, 1988.

McGann, Jerome. *Radiant Textuality: Literature After the World Wide Web.* New York: Palgrave, 2001.

McHale, Brian. *Postmodernist Fiction.* London: Routledge, 1987.

---. "Elements of a Poetics of Cyberpunk." *Critique* 33 (1992): 149-75.

---. "POSTcyberMODERNpunkISM." *Constructing Postmodernism.* Routledge: London, 1992. 225-42.

McLuhan, Marshall. *Understanding Media: The Extensions of Man.* London: Routledge, 1964.

McNamara, Kevin R. "*Blade Runner*'s Post-Individual Worldspace." *Contemporary Literature* 38 (1997): 422-46. *JSTOR.* 30 Apr. 2008
<http://www.jstor.org/stable/1208974>.

Mead, Syd. "Designing the Future." Peary 199-213.

Meehan, James "TALE-SPIN." *Inside Computer Understanding: Five Programs Plus Miniatures.* Ed. R.C. Shank and C.K. Riesbeck. Hillsdale, NJ: Lawrence Erlbaum, 1981. 197-258.

Metropolis. Dir. Fritz Lang. Perf. Alfred Abel, Brigette Helm and Gustav Frohlich. 1927. DVD. Eureka Video. 2003.

Midgley, Mary. *The Myths We Live By.* London: Routledge, 2004.

Minsky, Marvin. *The Society of Mind.* New York: Schuster-Touchstone, 1988.

---. "Why Computer Science is the Most Important Thing that Has Happened to the Humanities in 5,000 Years." Lecture. Nara, Japan. 15 May 1996.

---. "Scientist on the Set: An Interview with Marvin Minsky." *Hal's Legacy: 2001's Computer as Dream and Reality.* Ed. David G. Stork. Cambridge, MA: Massachusetts Institute of Technology, 1997. 15-31.

Moravec, Hans. *Mind Children: The Future of Robot and Human Intelligence.* Cambridge, MA: Harvard UP, 1988.

Mori, Mashahiro. "The Uncanny Valley." *Energy* 7.4 (1970): 33-5. *CogSci 2005 Workshop: Towards the Mechanisms of Android Science.* Trans. Karl F.

MacDorman and Takashi Minato. 2005. 12 Sept. 2008
<http://www.androidscience.com/theuncannyvalley/proceedings2005/uncannyvalle
y.html>.

Mullan, John. "False Farewells." *The Guardian.* 9 Nov. 2002. 2 May 2007
<http://www.guardian.co.uk/books/2002/nov/09/asbyatt>.

Nagel, Thomas. "What is it Like to be a Bat." *The Philosophical Review* 4 (1974): 435-50.

Nietzsche, Friedrich W. *Thus Spake Zarathustra.* 1885. Trans. A. Tille. London: Dent,
1958.

Noble, David. *The Religion of Technology: The Divinity of Man and the Spirit of
Invention.* New York: Penguin, 1999.

Paré, Ambroise. *Des Monstres et Prodiges.* Ed. Jean Céard. Geneva: Librairie Droz, 1971.

---. *On Monsters and Marvels.* Trans. and introd. Janis L. Pallister. Chicago: U of Chicago
P, 1982.

Partridge, Christopher. *The Re-Enchantment of the West: Alternative Spiritualities,
Sacralization, Popular Culture, and Occulture.* Vol 1. London: Clark-Continuum,
2004.

Partridge, John. "Plato's Cave and *The Matrix.*" *The Matrix Trilogy.* 20 March 2003.
Warner Brothers. 5 March 2008
<http://whatisthematrix.warnerbros.com/rl_cmp/new_phil_partridge.html>.

Patterson, David W. "Music, Structure and Metaphor in Stanley Kubrick's *2001: A Space
Odyssey.*" *American Music* 22 (2004): 444-74. *JSTOR.* 8 Apr. 2008
<http://www.jstor.org/stable/3592986>.

Peary, Danny, ed. *Omni's Screen Flights: The Future According to Science Fiction.*
Garden City, NY: Doubleday, 1984.

Penrose, Roger. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws
of Physics.* 1989. Oxford: Oxford UP, 1999.

Pescovitz, David. "Sons and Daughters of Hal Go Online." *The New York Times.* 18 Mar.
1999. 9 Jun. 2008 <http://query.nytimes.com/search/>.

Petroff, Elizabeth Alvilda. *Medieval Women's Visionary Literature.* New York: Oxford
UP, 1986.

Phiddian, Robert. "*Foucault's Pendulum* and the Text of Theory." *Contemporary
Literature* 38 (1997): 534-57.

Phillips, Adam. *Terrors and Experts.* London: Faber, 1995.

Pinker, Steven. *The Blank Slate: The Modern Denial of Human Nature.* London: BCA-
Penguin, 2002.

Plato. *The Symposium*. Trans. Walter Hamilton. Harmondsworth: Penguin, 1951.

---. *Republic*. Trans. and introd. Robin Waterfield. Oxford: Oxford UP, 1993.

---. *Phaedrus*. Trans. and introd. Christopher Rowe. London: Penguin, 2005.

Porush, David. *The Soft Machine: Cybernetic Fiction*. London: Methuen, 1985.

Pullman, Philip. *His Dark Materials*. London: Scholastic P, 2001.

Raban, Jonathan. "Just Two Clicks." *London Review of Books* 30.16 (2008): 3-9.

Rhodes, Neil and Jonathan Sawday, eds. *The Renaissance Computer: Knowledge Technology in the First Age of Print*. London: Routledge-Taylor, 2000.

Rosenbleuth, A., N. Wiener and J. Bigelow. "Behavior, Purpose, and Teleology." *Philosophy of Science* 10 (1943): 18-24.

Rosenhan, D.L. "On Being Sane in Insane Places." *Science* 179 (1973): 250-58.

Sammon, Paul M. *Future Noir: The Making of* Blade Runner. London: Orion, 1996.

Scott, Ridley. Interview with Ted Greenwald. *Wired*. 9 Sept. 2007. 9 June 2008 <http://www.wired.com/entertainment/hollywood/magazine/15-10/ff_bladerunner_full>.

---. "Directing *Alien* and *Blade Runner*: An Interview with Ridley Scott." Peary 293-302.

Searle, John. "Minds, Brains and Programs." *Behavioral and Brain Sciences* 3.3 (1980): 417-57.

---. *Minds, Brains and Science: The 1984 Reith Lectures*. London: British Broadcasting Association, 1984.

Selfridge, Oliver. "Pandemonium: A Paradigm for Learning." *Proceedings of the Symposium on Mechanisation of Thought Processes*. Ed. D. V. Blake and A. M. Uttley. London: H. M. Stationary Office, 1959. 511-529.

Selfridge, O. G. and U. Neisser. "Pattern Recognition by Machine." *Scientific American* 203 (1960): 60-67.

Shakespeare, William. *The Tempest. The Oxford Shakespeare: The Complete Works*. Ed. Stanley Wells, Gary Taylor, John Jowett and William Montgomery. Oxford: Clarendon, 1998. 1169-89.

Shannon, Claude E. and Warren Weaver. *The Mathematical Theory of Communication*. Urbana, IL: U of Illinois P, 1949.

Shelley, Mary. *Frankenstein, or, the Modern Prometheus*. 1818. Ed. M.K. Joseph. Oxford: Oxford UP, 1998.

Shinn, Thelma J. "'What's In a Word?' Possessing A.S. Byatt's Meronymic Novel." *Papers on Language & Literature: A Quarterly Journal for Scholars and Critics of Language and Literature* 31 (1995): 164-184.

Siemens, Ray and Susan Schriebman, eds. *A Companion to Digital Literary Studies*. Malden, MA: Blackwell, 2007.

Singer, Peter. *How Are We to Live: Ethics in an Age of Self Interest*. Oxford: Opus-Oxford UP, 1997.

Solomon, Andrew. *The Noonday Demon: An Anatomy of Depression*. London: Vintage, 2002.

Sontag, Susan. "Against Interpretation." *Against Interpretation and Other Essays*. London: Andre Deutsch, 1987. 3-14.

"South Korean Dies After Games Session." *BBC News Online*. 10 Aug. 2005. 3 June 2008 < http://news.bbc.co.uk/1/hi/technology/4137782.stm>.

Stapleton, Olaf. *Last and First Men*. 1930. London: Gollancz-Orion, 2003.

Stone, Allucquère Rosanne. *The War of Desire and Technology at the Close of the Mechanical Age*. Cambridge, MA: MIT Press, 1995.

Stonum, Gary Lee. "Cybernetic Explanation as a Theory of Reading." *New Literary History* 20 (1989): 397-419. *JSTOR*. 17 Sept. 2008 <http://www.jstor.org/stable/469107>.

Sutin, Lawrence, ed. *The Shifting Realities of Philip K. Dick: Selected Literary and Philosophical Writings*. New York: Vintage-Random, 1995.

System Shock 2. Vers. 1. Cambridge, MA: Looking Glass Studios, 1999.

Tallis, Raymond. The *Hand: A Philosophical Enquiry into Human Being*. Edinburgh: Edinburgh UP, 2003.

The Terminator. Dir. James Cameron. Perf. Arnold Schwarzenegger, Michael Biehn, Linda Hamilton and Paul Winfield. 1984. DVD. MGM Entertainment, 2007.

Thorne, Alison . "*The Tempest* and the Art of Masque." *Vision and Rhetoric in Shakespeare: Looking Through Language*. Basingstoke: Macmillan, 2000. 198-225.

Thurn, David H. "Fideikommi Bibliothek: Freud's 'Demonological Neurosis'." *Modern Language Notes* 108 (1993): 849-74.

Thurschwell, Pamela. *Literature, Technology and Magical Thinking, 1880-1920*. Cambridge: Cambridge UP, 2001.

Turing, A.M. "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society* 42 (1936): 230-65.

---. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433-60.

Turkle, Sherry. *The Second Self: Computers and the Human Spirit.* London: Granada, 1984.

---. *Life on the Screen: Identity in the Age of the Internet.* London: Weidenfeld, 1996.

Turney, Jon. *In Frankenstein's Footsteps: Science, Genetics and Popular Culture.* New Haven, CT: Yale UP, 1998.

Vernon, Victoria V. "The Demonics of (True) Belief: Treacherous Texts, Blasphemous Interpretations and Murderous Readers." *Modern Language Notes* 107 (1992): 840-54.

Wachowski, Larry and Andy Wachowski. "*The Matrix*: Shooting Script." *The Art of* The Matrix. Ed. Spencer Lamm. London: Titan, 2000. 271-394.

Walker, D.P. *Unclean Spirits: Possession and Exorcism in France and England in the Late Sixteenth and Early Seventeenth Centuries.* London: Scolar-Bemrose, 1981.

Wall-E. Dir. Andrew Stanton. Perf. Ben Bertt, Elissa Knight and Jeff Garlin. Walt Disney Pictures, 2008.

War Games. Dir. John Badham. Perf. Matthew Broderick, Dabney Coleman, John Wood, and Ally Sheedy. 1983. DVD. MGM Entertainment, 2000.

Warner, Marina. *Phantasmagoria: Spirit Visions, Metaphors and Media into the Twenty-First Century.* New York: Oxford UP, 2006.

Waters, Darren. "Brain Control Headset for Games." *BBC News Online.* 20 Feb. 2008. 2 June 2008 <http://news.bbc.co.uk/1/hi/technology/7254078.stm>.

Weizenbaum, Joseph. "ELIZA: A Computer Program for the Study of Natural Language Communication Between Man and Machine." *Communications of the Association for Computing Machinery* 9.1 (1966): 35-36.

Wells, H.G. *The War of the Worlds.* Ed. Martin A. Danahay. Ontario: Broadview, 2003.

Whitby, Blay. "The Turing Test: AI's Biggest Blind Alley." *Machines and Thought: The Legacy of Alan Turing.* Ed. P.J.R. Millican and A. Clark. Vol. 1. Oxford: Clarendon P, 1996. 53-62.

Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society.* London: Eyre and Spottiswood, 1954.

---. *Cybernetics, or, Control and Communication in the Animal and the Machine.* 2nd. ed. New York: M.I.T. Press, 1961.

Williams, Bernard. "Introductory Essay." *Meditations on First Philosophy, with Selections from the Objections and Replies.* By René Descartes. Trans. John Cottingham. Cambridge: Cambridge UP, 1986. vii-xvii.

Williams, Michael. "Descartes and the Metaphysics of Doubt." *Essays on Descartes'*
Meditations. Ed. Amélie Oksenberg Rorty. Berkley, CA: U of California P, 1986.
117-39.

Wolfram, Stephen. *A New Kind of Science.* Champaign, IL: Wolfram Media, 2002.

Wright, Richard. "Art and Science in Chaos: Contesting Readings of Scientific
Visualization." *FutureNatural: Nature, Science, Culture.* Ed. George Robertson et
al. London: Routledge, 1996. 218-36.

Yeffeth, Glenn, ed. *Taking the Red Pill: Science, Philosophy and Religion in* The Matrix.
Chichester: Summersdale, 2003.

Zynda, Lyle. "The Nature of Reality and Why it Matters." Yeffeth 43-55.