

Anthropomorphic Attachments in U.S. Literature, Robotics, and Artificial Intelligence

by

Jennifer S. Rhee

Program in Literature
Duke University

Date: _____

Approved:

Kenneth Surin, Supervisor

Mark Hansen

Michael Hardt

Katherine Hayles

Timothy Lenoir

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in Literature in the Graduate
School of Duke University

2010

ABSTRACT

Anthropomorphic Attachments in U.S. Literature, Robotics, and Artificial Intelligence

by

Jennifer S. Rhee

Program in Literature
Duke University

Date: _____

Approved:

Kenneth Surin, Supervisor

Mark Hansen

Michael Hardt

Katherine Hayles

Timothy Lenoir

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in Literature in the Graduate School
of Duke University

2010

Copyright by
Jennifer S. Rhee
2010

Abstract

“Anthropomorphic Attachments” undertakes an examination of the human as a highly nebulous, fluid, multiple, and often contradictory concept, one that cannot be approached directly or in isolation, but only in its constitutive relationality with the world. Rather than trying to find a way outside of the dualism between human and not-human, I take up the concept of anthropomorphization as a way to hypersaturate the question of the human. Within this hypersaturated field of inquiry, I focus on the specific anthropomorphic relationalities between human and humanoid technology. Focusing primarily on contemporary U.S. technologies and cultural forms, my dissertation looks at artificial intelligence and robotics in conversation with their cultural imaginaries in contemporary literature, science fiction, film, performance art, and video games, and in conversation with contemporary philosophies of the human, the posthuman, and technology. In reading these discourses as shaping, informing, and amplifying each other and the multiple conceptions of the human they articulate, “Anthropomorphic Attachments” attends to these multiple humans and the multiple morphologies by which anthropomorphic relationalities imagine and inscribe both humanoid technologies and the human itself.

Contents

Abstract	iv
List of Figures	ix
Acknowledgements	x
Introduction: The Double-Articulation of the Human	1
1 Misidentification's Promise: The Turing Test in Richard Powers' <i>Galatea 2.2</i> , Joseph Weizenbaum's ELIZA, and Emily Short's <i>Galatea</i>	10
1.1 Anthropomorphization, Perverse Metaphor, and Turing's Imitation Game	13
1.1.1 Gender between the Turing Test and Turing's Test	20
1.1.2 The Standard Turing Test: The Human Cut on the Bias	25
1.2 The Richard Test as Binding Narrative	28
1.2.1 The Turing Test Meets the Richard Test	37
1.3 From Sputnik to ELIZA: A Brief History of Natural Language Artificial Intelligence	41
1.4 The Loebner Prize: Anthropomorphization Foreclosed	46
1.4.1 The Anthropomorphic Imaginary: Misidentifications, Determination, and Desire	58
1.5 Intimacy and Distributed Agency in Emily Short's <i>Galatea</i>	59
2 Intelligence Embodied, Worlded, and Emergent: Rudy Rucker's <i>Software</i> and Rodney Brooks' Robotics	65
2.1 "Picture a 'brain in a vat' ..."	67
2.2 Brains, Bodies, and Boppers	76
2.2.1 The Body in Multiples	77
2.2.2 The Matter and the Memory	82

2.3 Pure Body and Becoming-Robot	85
2.3.1 Embodiment, Situatedness, and Emergent Intelligence.....	92
2.4 Brain in a Vat Redux: A Coda on Consciousness.....	98
3 Worlds beyond the Uncanny Valley: Philip K. Dick and the Robots of Cynthia Breazeal and Dave Hanson	103
3.1 Design Choices and the Uncanny Valley.....	104
3.2 Kismet: Animal, Infant, and Lowering the Threshold.....	110
3.2.1 Affect-Emotion, Anthropomorphization, and Breazeal’s Sociable Robots...	113
3.3 The Face I See and The Face You Know: Intersubjectivity and Recognition....	115
3.3.1 Intersubjectivity as Indistinguishability: The Rubber Hand Illusion and Self-Face Recognition:	117
3.4 Affect-Emotion	120
3.4.1 Sociable Robots and Affect-Emotion	123
3.4.2 Affective Anthropomorphization.....	125
3.4.3 HRI and Leonardo.....	128
3.5 Beyond the Uncanny Valley: David Hanson and Philip K. Dick.....	132
3.5.1 Life, Simulacrum, and De-Anthropomorphization in Dick’s <i>We Can Build You</i>	141
3.5.2 Sympathy and the Simulacrum	146
3.5.3 Affect-Emotion and the Limits of the Face	149
3.5.4 Alterity, Fugues, and Agency	152
3.5.5 Initial Dis-articulation in <i>Do Androids Dream of Electric Sheep?</i>	154
3.5.6 <i>Do Androids Dream of Electric Sheep?</i> and the Construction of Empathy ..	156
3.5.7 Anticipation as Invention: Dave Hanson’s Philip K. Dick Android.....	163

3.5.8 Post-Script: An Expansion of the Cultural Imaginary	167
4 Anthropomorphization Incorporated and Abjected: Stelarc's Extra Ear	168
4.1 The Visibility of Obsolescence: Stelarc's Obsolete Body	170
4.2 Stelarc's Extra Ear	176
4.2.1 Iteration 1: <i>Extra Ear</i> on the Face	176
4.2.2 Iteration 2: <i>Extra Ear: ¼ Scale</i>	178
4.2.3 Iteration 3: <i>Extra Ear</i> on the Arm	179
4.3 The Abject	182
4.3.1 Abjection of the Living	185
4.4 The Double Invisibility of the Senses	188
4.5 McLuhan and the Senses	192
4.6 Stelarc, McLuhan, and the Role of the Artist	195
5 Anthropomorphic Transduction and Speculation: Spielberg's <i>A.I.</i> and Microsoft's Milo	199
5.1 The Humanoid Transduced	200
5.2 Milo, Molyneux, and the Creation of the Human	203
5.2.1 What Claire Knew	204
5.2.2 What Claire Felt	209
5.2.3 Where Claire Is	212
5.2.4 What Claire Drew	214
5.3 Worlds Buffered, Worlds Collided	216
5.4 Post-Human World, Post-Human Time	220
5.4.1 Contingency and Chaos	227

Conclusion: The Human as Speculative and Singular	239
Bibliography	241
Biography.....	269

List of Figures

Figure 1: The right half of Charles Babbage's brain.....	67
Figure 2: Expressive Kismet.....	111
Figure 3: A selection of Kismet’s preprocessed affective vocal patterns organized according to Fernald’s prosodic classifications.....	127
Figure 4: Expressive Leonardo.....	129
Figure 5: Breazeal collaborated with Hollywood special effects company Stan Winston Studio to create Leonardo’s robotic body.....	130
Figure 6: Hiroshi Ishiguro (left) and his android, Geminoid HI-1 (right).....	135
Figure 7: Hanson’s PKD android.....	166
Figure 8: Stelarc's Extra Ear.....	179
Figure 9: Milo and Claire greet each other.....	205
Figure 10: Molyneux fills in the gaps.....	207
Figure 11: Milo throws Claire a pair of goggles.....	209
Figure 12: Milo shows Claire how to put on the goggles.....	210
Figure 13: Claire puts on the goggles as instructed.....	210
Figure 14: Claire sees her reflection in the water.....	212
Figure 15: Claire puts her hand in the water.....	213
Figure 16: Claire passes the drawing to Milo.....	214
Figure 17: “Orange.”.....	215
Figure 18: The glacial world after human extinction.....	227
Figure 19: David.....	237
Figure 20: Milo.....	237

Acknowledgements

My deepest thanks to my advisors. Your generosity with your time, vast knowledge, advice, and support has sustained me during this process, and has informed my thought immeasurably. I specifically wish to thank Ken Surin for his enduring guidance as well as his presence, both intellectually and as a human (in the richest sense of the term); Mark Hansen for fruitful conversations that returned me to the joy of intellectual work; Michael Hardt for his encouragement and support, as well as his unforgettable kindness during a very difficult time; Kate Hayles, for her incisive advice and intellectual generosity; and Tim Lenoir for his deep and abiding engagement and support, and for intervening when I was veering off track. Models, all. My most expansive speculations could not have imagined more.

The Program in Literature, the Graduate School, and the John Hope Franklin Humanities Institute's INTERFACE Franklin Seminar in the Humanities very generously provided financial support for my graduate career. My additional thanks to Rob Mitchell, Duke University Press, and the Thompson Writing Program's Writing Studio, for work opportunities that significantly shaped my development as an academic, both in thought and in practice. And to the University of California Humanities Research Institute for their tuition scholarship to attend the tecnoSpheres:FutureS of Thinking, Seminar in Experimental Critical Theory.

Many thanks to Kathy Psomiades for her generosity with her time and advice, and for encouraging me to pursue my intellectual interests. And to Anne Garreta, for generative discussions and comments on an early piece of this project.

To David Hanson and Ron Parr, for rich and illuminating discussions about robotics, and to Manos Tsakiris, for answering my questions about self-face recognition.

And to Leah Allen, Erin Fehskins, Nathan Hensley, Russ Leo, Rachel Price, Alex Ruch, and Britt Russert, who were kind enough to read and comment on various sections included in this dissertation, and on numerous sections that didn't make it into the dissertation.

And to Tiwonda Blount-Johnson, for all of her help and her perpetual calm through.

Thank you also: to Zach Blas and Patrick Jagoda for shared speculative explorations; to Marie-Pier Boucher and Pinar Yoldas – our late-night study group has made long days and late nights of work not just bearable, but delightful; to Talia Dorsey, for *la joie*; and to Rachel Price, whose remarkable friendship sometimes lulls me into a confidence I do not ordinarily possess.

To my family: to Chris and Katie – our histories may be relatively short, but my appreciation is no less deep. To Chrissy and Susie, who have long been sources of inspiration (and distraction!).

And to my parents, Eun and Jung Rhee, for everything.

Introduction: The Double-Articulation of the Human

The human may, in fact, be one of our most elastic fictions. As the dividing lines between humans and “nonhumans” have been historically redrafted to accommodate new systems of classification and new discourses of knowledge, the human has proceeded to mutate many times over.¹

“Anthropomorphic Attachments” takes as its central site of inquiry the mutations of the human across technologies and imaginaries, exploring the ways in which the human is multiply constructed relationally with humanoid technologies. These humanoid technologies, including artificial intelligence (AI), humanoid robotics, and virtual interactive video games, multiply create and imagine the human. Both mimetically organized around a specific conception of the human, as well as created to interact with humans, the technologies thus imaginatively inscribe a specific human with whom to interact. In other words, in modeling the technology on the human, as well as designing the interactivity with the human, these technologies doubly articulate their specific conceptions of the human. This double articulation of the human characterizes the constitutive relationality between the specific technology, the human-interactivity in which the technology is situated, and the human with whom the technology interacts.

In attending to humanoid technologies’ double articulation of the human, I foreground anthropomorphization, the “attribution of human form or character”, as a significant force by which the fiction of the human is imagined (*Oxford English*

¹ Diana Fuss. “Introduction: Human, All Too Human.” *Human, All Too Human*. Ed. Diana Fuss. New York: Routledge. 1996. 2.

Dictionary). In other words, anthropomorphization constructs and mobilizes the human as a category both fluid and intractable. The humanoid figure, in its mimetic practices, ambitions, and possibilities, is simultaneously and ambiguously human and not-human.

Drawing on N. Katherine Hayles' discussion of the feedback loops that connect the technological and the cultural (Hayles, *Chaos Bound*, xiv), I attend to the anthropomorphic imaginary in both humanoid technologies and their fictional depictions, as well as the ways in which the technologies and the cultural imaginaries shape each other and the humans that emerge anthropomorphically. In attending to the unique morphology of anthropomorphic relationships in humanoid technologies and literatures, my project speaks to philosophical and literary discourses of the technological human, such as that of the cyborg, the prosthetic, and the posthuman. Furthermore, with the emergence of distinctly nonhuman-oriented discourses and theoretical approaches – from animal studies, Bill Brown's thing theory, and the work of speculative realists such as Graham Harman and Quentin Meillassoux – I situate my project's investigation of the human alongside, rather than in opposition to, these de-anthropocentric projects. My attention to the anthropomorphic does not seek to reify or privilege the human, but rather to defamiliarize it in order to begin to think about how we can understand the human outside of the deep anthropocentrism in which it exists, and from which it emerges.

My chapter is organized into two sections. The first section, which is comprised of the first three chapters, looks at anthropomorphic organizations in humanoid technologies and literatures. Moving from intelligence in AI, to embodiment in mobile robotics, to emotion in sociable robotics, I examine these modes of anthropomorphization

and the various morphologies by which these technologies and their cultural imaginaries collaboratively articulate and depict the human. The second section of my dissertation, which is comprised of the final two chapters, examines the movement of anthropomorphization, first through abjection and incorporation in Stelarc's *Extra Ear*, then through projection and speculation in Steven Spielberg's *A.I. Artificial Intelligence* (2001) and Microsoft's virtual interactive boy Milo.

Chapter 1: Misidentification's Promise: the Turing Test in Richard Powers' *Galatea*

2.2, Joseph Weizenbaum's ELIZA, and Emily Short's *Galatea*

The core principles of Artificial Intelligence (AI), from its disciplinary emergence in 1956, are based on the metaphoric operations of an initial anthropomorphization, one that substitutes machine for human through the metric of intelligence. This initial anthropomorphization is also present in British mathematician Alan Turing's 1950 paper, "Computer Machinery and Intelligence," a founding text of AI from which the Turing test, the test to identify human from machine, is extrapolated. Drawing on discussions of the Turing test in philosophy, feminist thought, and artificial intelligence, and in fictions and cultural forms that take up the Turing test explicitly (Powers' *Galatea* 2.2 and the Loebner Prize) and implicitly (Emily Short's *Galatea* and Joseph Weizenbaum's ELIZA), I argue that the initial anthropomorphization in Turing's work can be understood a creative and productive force, one that imagines new human-machine intimacies and identities. This imaginary component of anthropomorphization is itself embedded in the possibility of misidentification in Turing's imitation game, and more

widely known interpretation, the Turing test. The question of accurate identification of either gender, in the imitation game, or species, in the Turing test, introduces the imagined possibility by which a machine could pass for a human. In so doing, it introduces the imagined machine that could successfully pass for human, as well as the imagined human as the metric by which to gauge the machine's success.

The fictional and cultural forms, having been influenced by the Turing test, speak to the scholarly debates about how to interpret Turing's imitation game. In this game, a machine competes with a woman to be identified as the woman. Gender disappears in the more widely known Turing test, in which a machine competes with a human to be identified as the human. In the cultural imaginaries I discuss in my chapter, AIs do not successfully engage with the Turing test outside of gender, thus positing the imagined human as gendered, as well as intelligent. In Turing's imitation game, and in the Turing test, it is the machine anthropomorphized as intelligent and gendered human that produces the possibilities of new modes of interaction and intimacy between human and machine. Thus, I argue that the Turing test operates less as a test to determine who or what is human, but rather productively destabilizes not only the distinctions we draw between human and machine, but also the project of distinguishment itself.

Chapter 2: Intelligence Embodied, Worlded, and Emergent: Rudy Rucker's *Software* and Rodney Brooks' Robotics

Organized around the trope of a brain in a vat, this chapter looks at robotics technologies and imaginaries that foreground embodiment as the criterion for the human.

Bringing Rodney Brooks' embodied and situated robotics into conversation with Hans Moravec's vision of uploadable minds and Rudy Rucker's cyberpunk fiction, *Software*, I argue that the embodiment imagined by Brooks, Moravec, and Rucker is dynamic, expansive, and multiple. Seeing as anthropomorphization does not operate in only one direction, but rather on both not-human (here, the robot) as well as on the human, the robotic embodiment in Moravec's writings, Brooks' robotics research, and Rucker's fictional robots also reconfigures the embodiment of the human as dynamic and multiple. Rucker and Brooks move away from early AI's disembodied approach to intelligence, refiguring both the human, as well as the robot, as necessarily embodied, and intelligence as a capacity of and this embodied engagement with the world.

This chapter brings together the technological and the fictional not through direct and explicit influences, but rather through overlapping spheres of reference, both technological and cultural. Brooks and Rucker both discuss the imbrication of the technological and the cultural in their own work. Brooks frequently cites the influence of fictional robots and AIs in his work, while Rucker, a professor of computer science, has published non-fiction works on computer science, artificial life, and robotics, citing Brooks in the process. By discussing Brooks' robotics in relation to the cultural imaginary, and Rucker's fiction in relation to the technological science, I attend to a cross-pollination introduced in their own work.

Chapter 3: Worlds beyond the Uncanny Valley: Philip K. Dick and the Robots of Cynthia Breazeal and David Hanson

This chapter attends to emotion as the characteristic of the anthropomorphized human in sociable robots. Taking as a given the embodiment foregrounded by Brooks' robots, Breazeal, a student of Brooks, designs her robots to interact with humans primarily through emotion expression. While this affective interactivity destabilizes the distinction between human and robot, this destabilization is limited. Breazeal's robots, which are designed to resemble a human infant and a small animal, maintain the gap between resemblance and identity, what Japanese roboticist Masahiro Mori calls "the uncanny valley."

As a counterpoint, I turn to roboticist Dave Hanson's robots and androids, which, in their pursuit of human-identity, seek to engage the human precisely in the depths of this uncanny valley. Hanson's robotic work, for example his Philip K. Dick android, has been significantly shaped by Dick's depictions of android humanoids, as well as his privileging of empathetic behavior in both humans and androids. Attending to Hanson's direct and explicit engagement of Dick's work, I discuss *Do Androids Dream of Electric Sheep?* (1968) and *We Can Build You* (1972), two novels that significantly grapple with the android and empathy. These novels, as well as Hanson's robots, expand anthropomorphization outward, to the extent that one cannot stop at reimagining the machine as human, but must in fact, in the destabilization and expansion of the human, reimagine the world itself. Beyond the Uncanny Valley, anthropomorphization is both minimized and amplified. Because of the physical resemblance, if not identity, Hanson's robots are quite easily anthropomorphized, or imagined as human. However, as Dick's novels demonstrate, anthropomorphization becomes even more essential and

expansive, as on the other side of the uncanny valley lies the necessity of imagining new worlds, new possibilities of life by which humans and robots can exist and co-exist.

Chapter 4: Anthropomorphization Incorporated and Abjected: Stelarc's *Extra Ear*

In “Anthropomorphization Incorporated and Abjected,” I turn to body, performance, robotic, prosthetic artist Stelarc and his ongoing project, *Extra Ear*. In its current iteration, *Extra Ear* involves a tissue-cultured ear equipped with sound emission technologies and attached to Stelarc's left forearm. In *Extra Ear*, the anthropomorphization of the world discussed in the previous chapter, is folded back into the self through the senses by way of this anthropomorphized, technologized sense-object – the ear internalized and incorporated into Stelarc's body. In this work, the anthropomorphic imaginary operates through the extra ear and the bodily reorganization it effects. Drawing on Julia Kristeva's concept of the abject as that which highlights the boundaries between life and death through the non-living, the no longer living, or the dead, such as bodily fluids and waste, I argue that in Stelarc's work, the abject does not operate through waste or expulsions, but instead through anthropomorphized technologies that recall the “living” body. Rather than disordering the relationship between life and death by calling up death as that which somehow both always and never opposes life, Stelarc's work uses the anthropomorphized living to disrupt and disorient life itself.

I bring this reading of the anthropomorphized abject into conversation with the poststructuralist thought of Gilles Deleuze and Félix Guattari, and Brian Massumi, and

Marshall McLuhan's theorization of media and the senses, to highlight how Stelarc's extra ear, through anthropomorphization and abjection, reorganizes the body in its conceptualization and in its capacities, such that this third ear is no longer understood as "extra," but rather as integral. This integration of the third ear does not occur through processes of familiarization, but rather through the continuous reorganization of the body, including or absent the non-extra ear.

Chapter 5: Anthropomorphic Transductions and Speculations: Spielberg's *A.I.* and Microsoft's Milo

My final chapter returns this discussion of anthropomorphization to the screen. Unlike the Turing screen, which introduced anthropomorphization through the possibility of human-machine misidentification, the screen in this chapter is the site of anthropomorphization's externalization and projection. I look at two specific sites of anthropomorphic externalization – Milo, Microsoft's virtual interactive body, and Steven Spielberg's *A.I. Artificial Intelligence* – both of which join the anthropomorphic conceptual organizations, technologies, imaginaries, and movements I discuss in the first four chapters in the form of humanoid boys. This chapter understands Milo and *A.I.* as realizing the double logic of anthropomorphization in full, rendering both humanoid and human as emerging only in their transductive relationality to each other.

Having established this relationality as constitutive of both human and humanoid, I engage with Graham Harman and Quentin Meillassoux, two contemporary philosophers associated with speculative realism, a term that encompasses disparate interventions into

“the subjectivistic and anthropocentric foundations of much of ‘continental philosophy’” (Brassier et al., 307). Harman’s object-oriented philosophy and Meillassoux’s speculative materialism, in very different ways, attend to the world outside of the human, and outside the human’s relationship with it. Their philosophical projects seek to understand the world independent from the human. Placing these de-anthropomorphizing projects in the theoretical echo chamber that is the humanoid, I explore the limits, implications, and potential of these philosophies in relation to my own deanthropomorphic project.

1 Misidentification's Promise: The Turing Test in Richard Powers' *Galatea 2.2*, Joseph Weizenbaum's ELIZA, and Emily Short's *Galatea*

"I've certainly left a great deal to the imagination"¹

In Richard Powers' *Galatea 2.2: A Novel* (1995), a novelist and a neural network researcher seek to create an intelligent machine – one that can produce a convincingly human response to an English literature master's exam question. N. Katherine Hayles describes the novel's test for intelligence as "a literary Turing test" (270). In this chapter, I discuss Richard Powers' *Galatea 2.2: A Novel* (1995) and the Turing test through anthropomorphization, the imaginative projection and attribution of human qualities onto a non-human entity. My discussion of anthropomorphization does not operate through the logic of substitution (of machine for human), but rather through the logic of liminality (in which human and machine exist always in relation to each other). By following the movement and trajectories in-between human and machine, I suggest that the space of ambiguity that is carved out by Turing's imitation game is the terrain on which the pulsions, vectors, and forces that work to keep human and machine discrete simultaneously propose, imagine, and effect new relationships and intimacies between human and machine, as well as new identities and agencies that recalibrate how we define "human" and "machine."

¹ Alan Turing, in a BBC interview, speaks evocatively about the imitation game he proposes at the outset of "Computing Machinery and Intelligence" (Newman, Turing, Jefferson, and Braithwaite, 124).

I will begin with a discussion of the anthropomorphic metaphor in Turing's 1950 article, "Computing Machinery and Intelligence," and in contemporary scholarly debates that surround the article's interpretation. In so doing, I will suggest that the anthropomorphic metaphor similarly grounds both the philosophical and the scientific discourses and practices; in other words, I suggest that the anthropomorphic imagination is an undeniable presence not only in the theoretical discussions about the Turing test, but also in certain sub-fields of artificial intelligence (AI) that are influenced by Turing's work. As this anthropomorphic imaginary is not solely the province of the fictional, I do not attend to the Turing test in scholarly and technological discourses to understand the structural logistics and theoretical import of the test, and then turn to Powers' novel to discuss the anthropomorphic imaginary. Quite the contrary – I approach these various discursive genres as similarly organized around the anthropomorphic metaphor. Following from Hayles' use of the feedback loop to articulate "the complex interconnections of theory, technology, and culture" I will introduce anthropomorphization as a productive metaphor, a force that undergirds and shapes the interconnected discourses of "theory, technology, and culture," while operating multiply and varyingly therein (*Chaos Bound*, xiv). Reading these various Turing discourses into each other, I suggest, allows us to more deeply understand the anthropomorphic imaginary as a force that propels anthropomorphic desire and offers the potential for productive Turing misidentifications – the recalibration of the human itself.

I conclude my chapter by discussing examples of AIs from the subfield of natural language processing, as well as Emily Short's interactive fiction, *Galatea*. In this final

section, I attend to different articulations of the anthropomorphic imaginary and the ways by which this imaginary is more or less successfully drawn out in these various Turing-inspired tests and interactions. By looking at natural-language AI ELIZA, the Loebner Prize competition (an annual competition based on the Turing test), and Short's *Galatea*, I argue that this ambiguity, which operates in productive ways to bring the human and machine into greater intimacy and identity-conflation, requires not a broadening of parameters and possible associations, but rather a narrowing and constraining of the contextual field within which to imagine and anthropomorphize. In other words, this ambiguity expands the categories of the human and the machine through the contextual specificities of the interactions.

“I’ve certainly left a great deal to the imagination.” Turing’s quotation, then, is doubly ambiguous. On the one hand, he could be referencing the structural ambiguity of his imitation game, about which there is significant debate. At the same time, however, we might also understand him as pointing to the role of the imagination, both as a component embedded in his imitation game, as well as a fundamental aspect of how we seek to differentiate ourselves, as humans, from machines. If, in his identificatory test for distinguishing human from machine, it is through the imagination that this distinction is articulated, it is thus at least in part through the imagination that this distinction can be confused, disarticulated, and reconstituted in new, previously un-imagined ways.

1.1 Anthropomorphization, Perverse Metaphor, and Turing's Imitation Game

In 1950, Alan Turing's "Computing Machinery and Intelligence" was published in the journal *Mind*.² "I propose to consider the question, 'Can machines think?'" the paper opens. This opening catachretic question does not exist prior to anthropomorphization.³ The anthropomorphic slippage between human and machine fundamentally shapes the question, the ways in which it is asked, the language that is used to ask, and the concepts that determine the asking.⁴ Sherry Turkle, in her anthropological work on cultures of technology, highlights anthropomorphization as undergirding the ways that humans think about and interact with computers:

[The computer's] evocative nature does not depend on assumptions about the eventual success of artificial intelligence researchers in actually making machines that duplicate people. It depends on the fact that people tend to perceive a "machine that thinks" as a "machine who thinks." They begin to consider the workings of that machine in psychological terms. (25)

² Numerous scholars cite Turing's 1950 paper as influential to AI. See Copeland, who cites the publication of Turing's paper as the founding moment of the philosophy of AI, and Turing as AI's "first major prophet" (*AI*, 9-10). See also Bolter and AI scientists Patrick Hayes and Kenneth Ford's "Turing Test Considered Harmful" (Bolter, "Artificial Intelligence," 4). And Warren Sack cites Turing's paper as a founding essay for computer science (1). However, it should be noted that Turing's influence on the field of AI is not unanimous. While Alex Roland and Philip Shiman describe the Turing test as "the Holy Grail of machine intelligence," Pamela McCorduck suggests that Turing "had practically no influence on most people at the Dartmouth Conference" (Roland and Shiman, 186; McCorduck, 95). Held in 1956, the Dartmouth Conference was a founding event in AI's disciplinary emergence.

³ The *Oxford English Dictionary* defines catachresis as "Improper use of words; application of a term to a thing which it does not properly denote; abuse or perversion of a trope or metaphor." I find this idea of perverse metaphor particularly useful in understanding the potential manipulability of metaphor.

⁴ Derrida writes, "What is defined, therefore, is implied in the defining of the definition" (Derrida, 230). In other words, the question itself does not emerge from a linguistic, theoretical, and cultural vacuum. The question itself is shaped by the same forces that shape the content and form of the answer to the question.

Like Turkle, I read this relative pronominal slippage from “that” to “who” as the organizing force by which machines are understood using the language of “thinking” and “intelligence.” This slippage is the unspoken anthropomorphic move by which Turing’s opening question can be said to read, “Can machines think [like humans]?” In other words, this unspoken slippage of anthropomorphization is fundamentally metaphoric.

Paul Ricoeur discusses metaphor as the emergence of new meaning from the collapse of previous meaning, a relationship that is both produced from, as well as produces, resemblance. “[M]etaphorical meaning does not merely consist of a semantic clash but of the *new* predicative meaning which emerges from the collapse of the literal meaning, that is, from the collapse of the meaning which obtains if we rely only on the common or usual lexical values of our words” (146). Metaphorical meaning, then, is not simply the juxtaposition, the “clash” of discrete terms and meanings, but rather the emergence of a new meaning from the relationality between the terms.

Through resemblance, according to Ricoeur, discrete objects are drawn into metaphoric relation.⁵ Without resemblance, which can be understood as a relation of

⁵ Derrida also emphasizes resemblance as “the condition for metaphor,” (Derrida, 237). Ricoeur’s and Derrida’s discussions of metaphor and resemblance depict many structural similarities. I rely significantly on Derrida’s attention to resemblance as a condition from which metaphor emerges and a critical component of the structure of metaphoric signification. However, I look to Ricoeur and his attention to the productive aftereffects of resemblance in metaphor to explore machinic anthropomorphization and Turing. I do not wish to suggest that Derrida views the metaphoric work of resemblance as complete after the metaphor is drawn. Derrida reads Aristotle as closely aligning metaphor and resemblance with mimesis and imitation: “The definition of metaphor is in its place in the *Poetics*, which opens as a treatise on *mimēsis*. *Mimēsis* is never without the *theoretical* perception of resemblance or similarity, that is, of that which always will be posited as the condition for metaphor.” (Derrida, 237). Through mimesis, Derrida then links metaphor and resemblance to the uniquely human ability to imitate: “[imitation] is proper to man. Only man imitates properly. Man alone learns by imitation” (237). Metaphor, then, becomes not just a way to signify, but also an ongoing and active process of signification, which resembles Ricoeur’s “predicative assimilation.” I only wish to suggest that, for the purposes of Derrida’s critique of

sameness that relies crucially on the exteriority of difference and otherness, metaphor cannot effectively operate as a process of resignification. However, resemblance does not operate simply through an existing sameness between objects or concepts. For Ricoeur, resemblance is a dually productive force. While the perception of resemblance may prompt the metaphoric drawing together of terms (the initial productivity), this resemblance continues to operate and work on the terms *after* the metaphor has been drawn. Ricoeur calls this latter productivity “predicative assimilation” (148). Thus, while resemblance can be said to be the “condition for metaphor,” metaphor can also be said to be the condition for resemblance.⁶

Ricoeur describes the metaphoric relation as one of “tension.” This tension is not just between the two now relational terms, but rather between the previous discreteness and the newly formed closeness or “proximity”:

metaphysics and its founding metaphors, he emphasizes the processes by which these founding metaphors emerged – processes which, upon scrutiny, do not hold up the heavy weight of the history of metaphysics. In Ricoeur’s language, Derrida’s text can be read as seeking to dismantle and reframe the predicative assimilation of these founding metaphors (“whiteness,” logocentrism, etc.) and their hold on Western metaphysics.

⁶ One might find oneself with questions of agency, such as, “Who or what, in response to initial resemblance, produces metaphor?” According to Ricoeur, the imagination is the agent of this assimilation. Ricoeur argues that metaphor cannot be thought purely in semantic terms; metaphor (and I would suggest that Ricoeur is making a claim for the role of the imagination in semantics itself) is neither pure semantics or pure imagination, but rather caught between the two, “on the boundary between a *semantic* theory of metaphor and a *psychological* theory of imagination and feeling” (143). I will speak at greater length about the role of the imagination in metaphor and anthropomorphization in my third chapter. Derrida, it should be noted, links the drawing together in metaphor to the human through language and imitation: because “*Mimēsis* is proper to man. Only man imitates properly. Man alone learns by imitation;” therefore, “Metaphor then is what is proper to man” (Derrida, 237, 246). And, while “properness” is not synonymous with agency, “properness” can be said to suggest spheres of relevance that might bring us closer to questions of agency.

In order that a metaphor obtains, one must continue to identify the previous incompatibility *through* the new compatibility. The predicative assimilation involves, in that way, a specific kind of tension which is not so much between a subject and a predicate as between semantic incongruence and congruence. The insight into likeness is the perception of the conflict between the previous incompatibility and the new compatibility. “Remoteness” is preserved within “proximity.” (Ricoeur, 149)

This tension between multiple valences, the vacillation between difference and similarity, and “remoteness” and “proximity,” constitutes metaphor as relationality, as the emergence and production of multiple meanings that exist completely in neither term; rather meanings emerge from the liminality between two terms. My discussion of anthropomorphization in Turing-based discourses and fictions attends to this metaphoric relation between human and machines by way of the vacillations, predicative assimilations, and multiple tensions that emerge from this relationality, transforming and resignifying the humans and machines therein.

If metaphor can be described as “the application of a name belonging to something else” anthropomorphization can then be described as the metaphoric application of the name “human” to that known as “non-human” (Aristotle, 28). And yet, the anthropomorphic metaphor poses unique challenges to signification. Because the human, the object of anthropomorphization’s resemblance and imitation, is a nominalization as empty as it is full, anthropomorphization itself is simultaneously narrow and broad in its meaning-producing practices and possibilities. In other words, anthropomorphically projecting the human, in fact *creates* the human itself, delineating and contouring the human that did not exist as such prior to the act of anthropomorphization. Thus, in discussing the human through anthropomorphization, I

attend to the liminal processes by which signification emerges from the metaphoric relationship. For, what emerges from this uniquely and doubly human metaphor is not a singular, discrete meaning, but rather a crucial ambiguity that emerges from the in-between of metaphor,⁷ an ambiguity that can either reify and foreclose, or produce new modes of interrelation.

Under the heading of “The Imitation Game,” Turing’s paper opens as follows:

I propose to consider the question, “Can machines think?” This should begin with definitions of the meaning of the terms “machine” and “think.” The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words “machine” and “think” are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, “Can machines think?” is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the “imitation game.” (“Computing Machinery,” 434)

Turing introduces the question, “Can machines think?,” only to replace it with the imitation game. I wish to think about this moment of seeming dismissal as in fact attending to the initial anthropomorphic elision from which the question emerges. At the crux of this elision is the unspoken and invisible human who, it is taken for granted, can indeed think. Thus, the question, “Can machines think?” can be said to interrogate machines’ possible resemblance to humans – or rather, to produce the resemblance it seeks to interrogate. When read through the lens of anthropomorphization, Turing can be said to ask the question not of the machine, but rather to reframe the question of the

⁷ Metaphor’s power and charm lie precisely in its ambiguity, according to James Deese (212).

machine relationally with that which is elided – the human. In other words, Turing returns the implied human to the fore of the original question, eschewing questions of definition for those of relationality. While definitional approaches to this question can be said to operate outside of metaphor, Turing’s replacement opens up possibilities for anthropomorphic metaphor, for new *relational* significations between human and machine.⁸

After explaining the mechanics of the imitation game – in which a person must correctly distinguish, through typed conversations, between a man and a woman⁹ – Turing concludes with “We now ask the question, ‘What will happen when a machine takes the part of A in this game?’ Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’” (“Computing,” 434). These questions are by no means “unambiguous,” particularly when offered as a replacement for the original question. In fact these questions, in their emphasis on the interrogator’s ability to correctly decide, assert the inextricability of the machine’s result, as thinking or non-thinking entity, from the human interrogator. The onus of success or

⁸ I am not the first to suggest that Turing moves away from the goal of producing definitions (for “machines,” “thinking,” “intelligence,” “human”). Stuart Shieber, an extensive commentator on Turing’s Test, and B. Jack Copeland, who serves as Director of the Turing Archive for the History of Computing, both read Turing as moving away from definitions (Shieber, “Immediate Responses, 135; Copeland, “The Turing Test,” 522). Whereas Shieber and Copeland continue to ascribe a certain centrality to the machine’s performance, I suggest that the machine, while the nominal subject of inquiry of Turing’s paper, emerges at the forefront of an already anthropomorphized context in which the human is central to and agential in the actual imitation game with which Turing replaces the original question.

⁹ I will discuss the mechanics of the imitation game in more detail and at greater length in the following section.

failure does not rest solely on the machine, but rather rests largely on the human.¹⁰

A number of scholars have remarked on the centrality of the human in Turing's imitation game, and thus the inextricability of the machine from the human, as a weakness or failure of the game.¹¹ While I agree with this identification of anthropocentrism, I am less inclined to see it as a failure. Rather, I suggest that we take seriously Turing's move to redirect the conversation away from a more definitional approach to the question. By focusing, not on the human, not on the machine, but on the liminalities between these agents, we can read, not for the insularity of static definitions, but rather for the interactions and encounters between human and machine. Thus, I undertake my discussions of the original question and the imitation game always with an eye to these liminalities, by returning to this metaphoric act of replacement by which new human-machine relationships and subjectivities can emerge. The imitation game gestures to the processes of anthropomorphization that undergird Turing's opening section as well

¹⁰ In "Thinking and Turing's Test," Peter Naur critiques Turing's imitation game on the basis that it relies too much on an anthropocentrism. I agree, yet see no cause for critique. In reframing the original question in terms of the imitation game and the role of the human interrogator, Turing is revealing the original question as always already anthropomorphized.

¹¹ For examples, see the following discussions: Peter Naur critiques Turing's paper on the basis that its logic relies too much on an anthropocentrism. Robert French suggests that the Turing Test is a strong test for human intelligence, but not for any non-human human intelligence (64); P. H. Millar suggests that our conception of intelligence relies too much on abstract human traits and aims, and thus wouldn't be applicable to Martian or machine intelligence, much less to humans across variable cultures (596-597); computer scientist Adam Drozdek argues that the Turing Test "cannot but test human intelligence," though views computers as "the fruit of human intelligence" and thus possess some claim to human intelligence (318; 317); Geoffrey Sampson interestingly inflects this idea of the Turing Test's human chauvinism: in response to Richard Putrill's claim that computer behavior, unlike human behavior, is determined (programmed), Sampson points out the value of the Turing Test in gauging human intelligence, as "computers are designed by humans; humans are not designed by humans" (593). Thus, Sampson embraces the Turing Test's human bias, on the grounds that computers themselves are a product of this human bias. For further discussions of human bias in the Turing Test, see Shieber's "Early Harbingers of Later Issues" and Michael J. Spivey's "Turning the Tables on the Turing Test: The Spivey Test."

as the broader questions of machine intelligence in which Turing is intervening. Indeed, Turing's imitation game as replacement attends to the anthropomorphization that not only underlies, but also is the condition of possibility for, the original question. In reframing the original question in terms of the imitation game and the role of the human interrogator, Turing is revealing the original question as always already anthropomorphized.

1.1.1 Gender between the Turing Test and Turing's Test

Turing's imitation game involves no computers and three humans, at least one of whom is a man (A), and one of whom is a woman (B). The man and the woman are in one room, and the remaining human (C), who may be of either man or woman, is in a separate room. Connecting A and B with C is a teletype machine, by which C asks A and B questions, and A and B respond. Through these typed conversations, both A and B try to convince C that he or she is the woman.¹² C is then tasked with correctly guessing who is the woman.¹³

What happens next is far from "unambiguous," as both the text and the substantial disagreement surrounding the following move demonstrate. At this juncture, Turing

¹² In Tyler Curtain's astute reading of Turing's imitation game, he points out that the injunction upon B to prove her status as woman is not equivalent to A's attempt to convince C otherwise. Curtain describes this non-equivalence in the imitation game as "[t]he philosophical burden of women to speak – and for an adequate number of times *fail to represent* – the 'truth' of their sex" (139).

¹³ For Friedrich Kittler, "Man coincides with his simulation" from before the start of the imitation game, because the communication between A and C or B and C occurs through teletype transcripts rather than handwritten documents (17).

introduces the machine into the imitation game described above. The machine, according to Turing, will take the place of A: “What will happen when a machine takes the part of A in this game? Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman?” (“Computing Machinery,” 68). In other words, when a machine replaces A, will the interrogator make the wrong decision as often as he or she does when the game is played between a man and a woman? And, if A is replaced by the machine,¹⁴ does Turing intend that, in this new version of the imitation game, the human interrogator continue to attempt to identify the woman? Or does “human” replace “woman” as identificatory metric?

Many critics who argue that Turing intended that the human interrogator judge A₂ and B on the basis of “human” tend not to mention gender,¹⁵ or cite gender as irrelevant or inconsequential,¹⁶ even though gender occupies significant space and status in Turing’s game description. For example, during a roundtable conversation on the Turing test, Daniel Dennett was asked the question “Why was Turing interested in differentiating a man from a woman in his famous test?” Dennett responded with the slightly

¹⁴ To distinguish between computer and man, I refer to the computer as A₂, and the replaced man as A₁.

¹⁵ In a slightly puzzling reading, Douglas Lenat suggests that ungendered discussions of Turing’s imitation game are a result of the political correctness movement. See “Building a Machine Smart Enough to Pass the Turing Test: Could We, Should We, Will We?”

¹⁶ Naur less gently characterizes sex difference as a “pseudo issue,” which serves only to distract the interrogator “away from the real issue, the difference between [hu]man and machine” (183). Hayles reads Turing biographer Alan Hodges’ dismissal of gender in Turing’s paper as a way to maintain embodiment as the primary or sole arbiter of intelligence. Additionally, she argues that Turing’s test is useful precisely in positing technology’s potential to dislodge what she calls the “enacted” body (embodiment) from the “represented” body of technological performance; in other words, she suggests that the possibility of inhabiting both genders or species in a misidentification points to new, indeed necessary, recalibrations of “the human” itself (*Posthuman*, xiii-xiv).

dismissive: “That was just an example” (23). Dennett went on to explain that Turing only intended to illustrate that there are different ways of thinking (in this instance, between a man and a woman), but that one cannot argue that either is not thinking. While Dennett’s point that we should consider thinking and intelligence broadly is well taken, I appreciate the questioner’s assertion that Turing’s interest in gender was not a question. In other words, the questioner did not ask *whether* Turing was really interested in distinguishing between a man and a woman, but rather took the role of gender in the test seriously, as an important component for understanding machine intelligence.

This terminological slippage from “sex” to “gender” did not, I am certain, pass unnoticed by the reader. Turing specifically used the word “sex” in describing the imitation game. However, later discussions of the imitation game speak of “gender” rather than “sex.” One cannot help but wonder if this move from “sex,” which Turing explicitly names in his paper, and “gender,” which can arguably be said to be brought into discussion in his description of the imitation game, is at least minimally effected by knowledge of Turing’s biography.¹⁷ I am wary of allowing Turing’s biography to overdetermine my discussion of his imitation game, though the absence or dismissal of

¹⁷ In “Automating Gender: Postmodern Feminism in the Age of the Intelligent Machine,” Judith Halberstam eloquently discusses Turing’s imitation game in relation to the similarly learned and imitative properties of both gender and computer intelligence, framing (though not proscribing as causality) her brief discussion of gender in Turing’s imitation game around Turing’s biography: his court ordered organo-therapy on account of his homosexuality and his suicide by cyanide apple. Judith Genova, while problematically reading Turing’s imitation game as overdetermined by certain aspects of his biography, usefully posits that Turing was in fact speaking of the more culturally determined gender, as opposed to the biologically determined sex (“Turing’s sexual guessing game”). For discussions of Genova’s reading of gender as well as her reliance on Turing’s biography, see William Keith’s “Artificial intelligences, feminist and otherwise,” and James A. Anderson’s “Turing’s test and the perils of psychohistory.” For an extended discussion of Turing’s biography, including the punitive hormone therapy to which Turing was subjected, having been convicted in 1951 of “act[s] of gross indecency with... a male person,” see Hodges (471).

sex or gender in many discussions of Turing's imitation game operates as a similarly cautionary tale, one that Warren Sack attributes to "the bachelor machine": "AI researchers have functioned as a 'bachelor machine' to orchestrate the woman and issues of gender difference out of their re-narrations of Turing's imitation game" (15).

Susan Sterrett's discussion of Turing's test as "meta-game" does not easily dismiss either gender or species, as Judith Genova describes these identificatory approaches. Rather, she embeds Turing's ambiguity in her own insightful articulation of Turing's imitation game. Indeed, Sterrett's reading itself can be said to emerge from the moment of replacement in Turing's paper. Rather than discarding that which was replaced (in this case, A_1) for its replacement (A_2), Sterrett argues that Turing's test can best address questions of machine intelligence when comparing these two game pairings. In other words, both A_1 and A_2 (man and machine) are paired with B, and are interrogated by C, who must identify the woman in both A_1 -B and A_2 -B pairings. The success and failure of A_1 and A_2 are scored according to the number of times the human interrogator mis-identifies each as woman, and the results in these separate trials are then compared to each other. Sterrett calls this "meta-game" reading of Turing's test the Original Imitation Game (OIG). She names the more prevalent reading of a single game – whereby the identification in question is not "woman" but "human," and with A_2 's replacement of A_1 , A_1 is no longer part of the game – the Standard Turing Test (STT) ("Too Many Instincts," 43).

According to Sterrett, the OIG is a more apt test for machine intelligence than the STT for several reasons. While the STT can rely too much on the specificity of the

human judge, the OIG neutralizes this specificity, not because the human judge is more or less skilled in making the correct determination; rather, in the OIG, the specificity of the human interrogator (whatever his or her abilities and idiosyncrasies) are constant in both A₁-B's interrogations and A₂-B's ("Nested Algorithms," 135). Further, while in the STT only the machine is imitating being human, while its competitor (B) is simply *being* human,¹⁸ in the OIG, imitation remains the specific human ability that is being tested in both A₁ and A₂. In other words, the OIG, both A₁ and A₂ are imitating being woman; neither A₁ nor A₂ is simply "being." Thus, the imitation game remains a game that tests the distinctly human ability to imitate.¹⁹ Lastly, Sterrett suggests that a key advantage of the OIG holds over the STT is that the OIG tests intelligence outside anthropomorphization ("Turing's Two Tests," 554).

I would like to propose an alternative reading of Sterrett's Original Imitation Game as effective precisely *because* of its heavy reliance on anthropomorphization. In fact, one might suggest that the OIG neutralizes potential anthropomorphic bias by flooding the game with anthropomorphic forces. In the OIG, unlike in the STT, the

¹⁸ And while I am the first person to admit that "being human" is a radically vague and difficult task to perform (indeed, this is precisely the task that is constantly bracketed, from meaning, in the Turing Test, in this project), I am convinced by Sterrett's attention to the structure of the OIG as evenly distributing the task of imitation across players A's.

¹⁹ Both Derrida and Aristotle discuss *mimēsis* as uniquely human, as compared to animals. Extending Derrida's reading to the more generally not-human, it might be argued that machines have access to *logos* in a way that weakens this reading of *mimēsis*. Though machines may currently appear to have a certain access to *logos*, in actuality they are only able to produce a certain *logos*-effect. This *logos*-effect is not successful imitation, but is more appropriately aligned with Derrida's description of sophistry: the "manipulat[ion of] empty signs and draw[ing of] effects from the contingency of signifiers" (Derrida, "White Mythology," 248n54). The difference (while at times fine) between *logos*-effect and *logos*, between sophistry and *mimēsis* is nothing less, and nothing more, than the human. The tautological structure of this reading does not escape me – indeed, this tautology of the human, or human as tautology, is precisely the movement the Turing debates quite aptly demonstrate.

human interrogator, in seeking to identify the female, presumes the humanness of all the participants. In other words, because the game is already anthropomorphically saturated, as the human-ness of the competitors is not in question, in the OIG “the human” is not the question. As the interrogator encounters every contestant “as if” he, she, or it were human, anthropomorphization is consciously embedded into the structure of the OIG as a necessary component of the game itself. This presumption of humanness in the OIG gestures to the ways by which gender is inextricable from considerations of the human.

It might be said that Sterrett, while not dismissing gender, does not effect an extensive discussion of sex as a specific identificatory category. And indeed it is said by Sterrett herself, who argues for the OIG less on account of its inclusion of gender, but rather because of its meta-game structure. However, while Sterrett does not extensively address the vast theoretical implications of gender in discussions of humans, machines, imitation, and performance, she does not close off these discussions. In fact, I suggest that her meta-test OIG is predicated on the ways that gender cannot be extricated from these considerations, as seen in ELIZA, a somewhat successful natural language AI, and in Helen, the AI in Powers’ novel. I will speak about both of these at greater length later in this chapter.

1.1.2 The Standard Turing Test: The Human Cut on the Bias

Sterrett’s reading of Turing’s test, as I mentioned earlier, is not the standard interpretation of the Turing test. In the Standard Turing Test, one human and one

machine compete to be identified as “human” by a human judge.²⁰ Hayes and Ford point to the Turing test’s over-reliance on the weaknesses and variabilities of each individual human judge, whose ability to identify the human conversant may say very little or even nothing about the skill and ability of the AI.²¹ The human judge, upon whom the success or failure of the machine rests, who decides whether or not the machine can be said to “think” or “be intelligent,” may or may not be “clever, knowledgeable, and insightful”²² (Hayes and Ford, 973). It goes almost without saying, then, that the same AI machine and human pair may elicit vastly different identifications from different, and largely un-criteria-ed, human interrogators. Hayes and Ford’s critique of the species-oriented Turing test is elaborated as follows: “The species test further reveals the poor experimental design of the imitation game in the difficulty of obtaining an unbiased

²⁰ The species oriented Standard Turing Test is commonly referenced as “the Turing Test.” As the rest of this discussion will be drawing from a variety of critiques of this interpretation of Turing’s work, I will, from here on, somewhat grudgingly refer to the Standard Turing Test as “the Turing Test.”

²¹ Hayes and Ford critique the Turing Test’s (in both “gender” and “species” versions) constructed opposition between machines/artificial intelligences, and humans. According to Hayes and Ford, leaving behind the competitive nature of the Turing Test (and this is what is so “harmful” about the test) would allow us, as well as those within AI, to think about how machines could “amplify and support” humans, rather than threaten and replace them. While Turing, they suggest, is not originally responsible for this opposition, his paper produced substantial discussions that do much to reinscribe this forced opposition as natural or inherent (Hayes and Ford, 972, 975-976). Hayes and Ford believe that neither the gender test nor the species test (the Turing Test) escapes danger of perpetuating a forced opposition between human and machine; they nonetheless isolate the species test as the object of a lengthier critique on the basis of human bias.

²² This critique of the Turing Test is not an uncommon one. Sterrett, as I discussed earlier, argues for the OIG over the STT on the basis of this critique. See also Eugene Demchenko and Vladimir Veselov, who point to the test’s reliance on individual judges, whose specific biases are not standardized within or across tests in “Who Fools Whom?: The Great Mystification, or Methodological Issues on Making Fools of Human Beings,” as well as Kurzweil and Kapor’s “A Wager on the Turing Test.” Kurzweil and Kapor not only point to the inconstancy of the human judge in Turing contests, but locate the judge’s competence as the factor that renders a machine intelligent. In other words, the human’s competence is the (not intelligent) machine’s failure, while the human’s incompetence is the (now intelligent) machine’s success. Less the skill or ability of the AI, but rather the skill or ability of the human produces machine intelligence.

judge. The general perception of what are essentially human talents keeps shifting” (Hayes and Ford, 974). Hayes and Ford’s critique can be broken down into two components: 1. the virtual certainty of a biased judge, and 2. the constantly shifting definition of the human. I agree with Hayes and Ford that the species-oriented Turing test is flawed on these two counts. However, I suggest that this flaw, or the structural design weakness of the imitation game, owes much to the relationship between Hayes and Ford’s two criteria, and the slightly ambiguous and nonspecific charge of “bias.”

The biased judge is, according to the rules of the game, human. And yet, the definition and perception of humanness “keeps shifting,” continuously reverberating against the human judge. In other words, defining the human is a difficult task. This difficult, if not impossible, task is made even more so because the nature of the bias that Hayes and Ford mention is the pulsion, not just to correctly identify, but in fact to correctly *self-identify*, to recognize oneself. This movement is a kind of predicative assimilation by which the human judge resignifies him or herself as human in relation to these Turing conversations with machines. It is this relational resignification of self in relation to machines that opens up new engagements, collaborations, and intimacies between humans and machines.

The cultural imaginary, the fictional representations of the Turing test, crucially intervenes into this debate about Turing’s intended identificatory metric, asserting the anthropomorphic metaphor as largely facilitated by imagining a human who is not

ungendered.²³ One does not anthropomorphize without also en-gendering. In Powers' *Galatea 2.2*, one does not imagine a human without imagining him or her as gendered. The novel, then, draws on the Turing test to activate an anthropomorphization that reframes the categories of gender, subjectivity, human, and machine as fluid, rather than staid and inviolable.

1.2 The Richard Test as Binding Narrative

Richard Powers, the narrator of Richard Powers' 1995 novel *Galatea 2.2*, is a novelist who spends a year at the Center for the Study of Advanced Sciences at a university known as "U." During this year, Richard collaborates with cognitive neuroscientist Philip Lentz, who works specifically in connectionism and neural networks. Together, Richard and Lentz build and train a machine to take an English literature master's exam. Richard's interactions with the progressing iterations of the machine is one of two interwoven narratives that unfold and unfurl each other. The second narrative tells the story of Richard's failed romance with C. Indeed, it is the demise of this relationship that brings Richard to the Center at U.

There is a third narrative folded into the novel, one that appears to be nested in the narrative of Richard and machine, but in fact weaves in and out of the two temporally

²³ In contrast, the species-oriented bachelor machine is more invested in maintaining the distinction and opposition between human and machine than in exploring the ways in which a machine could in fact be imagined to pass for human.

disjunct narratives and in so doing, binds them.²⁴ This binding narrative recounts the evolution of Richard's anthropomorphic belief, the imaginative element by which the possibility of misidentification, of human for machine, of machine for human, of man for woman, and woman for man, emerges. This narrative of Richard's belief begins upon his first encounter with Lentz, just prior to their collaboration. Richard first meets Lentz late one night at the Center, after strains of Mozart draw Richard to Lentz's office. Richard later recalls this first encounter: "This Lentz, I reasoned, had a neural network buried in that mountain of equipment. One that he was training to recognize beauty. One that would tell him, after repeated listenings, how that simple reed breathing made and unmade the shifting signal weights that triggered souls" (Powers, 16). In this early encounter, Richard draws the artificial network into the human world of beauty and souls, imagining the ways in which the network could amplify our understanding of this world. Richard does not suggest that the machine shares the human capacity to appreciate or create this beauty, nor is it among the souls who are triggered – not yet. In Richard's memory of his initial encounter with Lentz, "reason" emerges in conjunction with "imagination," pointing to the ways in which these two aspects are necessary components of each other in the novel, as they are in the Turing test.

²⁴ The multiple narrative threads of this novel, as well as its reliance on autobiography, produce a dizzyingly recursive novel and a narrator whom the reader cannot be sure knew what when. In their reading of Powers' novel, Bould and Vint articulate ambiguity as a component of the autobiographical subject: "Such tensions between determinacy and indeterminacy, between likeness and difference, are central to understanding the autobiographical subject, the self that emerges both in and into language. This self is brought into consciousness and made into an object of reflection by that consciousness, which is *like* but yet *is neither* the self who lived *nor* the self who narrates that life... Autobiography is as much a making of a self as a description of one" (84). Through the ambiguity of autobiography, similarly to the ambiguity of anthropomorphization, Richard the narrator is also creating himself.

As Richard's research into neural networks progresses, this anthropomorphic imaginary, largely upon Lentz's training and modeling, takes hold of Richard even more strongly. And, just as Richard gives the final version of the machine a name, "Helen," Lentz calls Richard "Marcel" throughout the novel. Richard is both Pygmalion to Helen's Galatea and Galatea to Lentz's Pygmalion. Captivated by their first encounter, Richard asks Lentz to send him off-prints of his work. Thus Richard's scientific education begins, as does his anthropomorphic one; both continue as his collaboration with Lentz progresses through a series of machines, from Implementation A to implementation H. By the end of the novel, we understand that all along, Lentz, in what I call the Richard test, was training Richard to anthropomorphize *despite* his understanding of the science behind the increasingly convincing machine performance. But much happens between that first encounter and Richard's belief at the end of the novel. I will trace this binding narrative as it moves through the novel, beginning with the first version of the machine, Imp A.

From the first machine, Implementation A, Richard is anthropomorphically embedded in the narrative of machine progress, and the distinction between Imp and Richard, between machine and human, is unseated: "Lentz built Implementation A more for my education than as a prototype with any real pretensions" (71). Imp A, at least by Richard's account, is more about Richard than it is about creating a machine that can compete in the proposed Turing test. Because of Richard's abysmal typing skills, Lentz insists that Imp A be outfitted with voice recognition capabilities. Rather than typing in the works of literature, Richard begins to read these works to Imp A. Through the

parallel narrative of his love affair with C., we see how this technological modification recalls what is for Richard a deeply intimate act between himself and C.. Richard and C., in reading books to each other, created a world in which they were “the other’s entire audience” (33), creating themselves and each other in these moments of storytelling.

Imp A, which learns to compose sentences but cannot generate its own content or meaning, soon becomes Imp B, which Richard dislikes because it reminds him too much of himself, of his own tendencies toward headiness and generalization. Richard describes Imp B as driven “batty as a poet” by figuration (90), as relying too much on associative connections to indicate meaning. This inhabitation of association is mirrored in both Richard and C., who together, “lived from sense to sense” (98), and in C. alone, who would take Richard on long, directionless walks. “Her aimlessness was always hard on me” (102), Richard remembers. Richard and C., as a dyad and as individuals, are woven in and out of Imp B to the extent that disarticulation is both impossible and beside the point.

Lentz introduces Richard to Imp C. Richard realizes that Imp C, in its seeming conversational nuance, is actually a trick played on him by Lentz and Diana Hartrick, a fellow researcher at the Center. Richard was not speaking with Imp C, as he believed, but with Diana. Before the trick is revealed, Richard marveled at Imp C’s sophistication, gleefully thinking of it as “an idiot savant” (122). However, upon realizing that he was fooled, Richard, humiliated, states “I’d been an idiot” (123). Richard is all Imp C’s idiot, and none of its savant. Richard attributes his gullibility to his strong desire to believe; recalling the conflation of reason and imagination upon his first encounter with Lentz,

Richard recites his desire for this anthropomorphic belief: “Two seconds of reflection should have told me that C couldn’t have commanded even a fraction of the material it spewed out. A babe in the woods would have seen through this. [...] Yet I’d believed. I’d *wanted to*” (123). Richard wants to believe, in spite of himself as a knowing, reflecting, and reasoning subject. Before even meeting Imp C, Richard is already anthropomorphizing it, paraphrases Lentz’s description of Imp C’s capacity to “mak[e] its own input” as “You mean, it can anticipate new material. You’re trying to tell me it’s thinking” (121). Jumping quickly from generating input and anticipation to thinking, Richard’s anthropomorphic belief is already activated, setting the stage for the trick played on him by Lentz and Diana.

Even more striking, after the trick is revealed, Richard, both humiliated and angry, still believes:

“Lentz,” I said softly, “I’ll never trust you again.”
“Don’t need your trust. I just need you to train Imp C.”
I stopped, waiting to hear what I was going to say. “Imp D.”
The two colleagues, divided on every issue except novelists’ gullibility, broke into relieved tittering. (124)

At this moment, the trick points to both the alternative test – the Richard test – as well as the deep anthropomorphic Turing-ambiguity between humans and machines in this novel. Indeed, we hear nothing more of Imp C,²⁵ and the narrative moves on to Imp D. Trick or not, this is all the novel gives us of Imp C – Diana as C, and Richard fooled by his desire for anthropomorphic belief.

²⁵ Lentz later recalls Imp C to taunt Richard for his gullibility (273).

Richard, while training Imp D, decides that it needs “eyes” in order to understand what a thing is without being told of its characteristics and properties. For this purpose, Richard does not suggest that Imp D needs vision or the ability to see, but eyes, to which Lentz replies, “Nice leap, Marcel... I’ve been waiting for you to make it” (128). We can read these words of praise doubly. First, Lentz congratulates Richard on his reasoning, making the leap between knowing an object through description and experiencing it perceptually. Second, this “leap” can also be read as a leap between vision and object-recognition in Imp D, and “eyes” – an anthropomorphic leap which Lentz has been modeling and training Richard to make, and for which Lentz has been waiting.

“I crashed Imp E in complete innocence,” Richard recounts. “One day, provoked by boredom, I asked it, ‘What do you want to talk about?’” (153). Imp E crashes when Richard asks it this “question of volition.” “The question of volition trapped the rolling marble of its will into an unstable local minimum. The machine that so dutifully strove to answer every interrogation ground to a halt on that one” (153). In another life, Richard posed C. with a similar question of volition, though not out of boredom, it should be mentioned. “Buddy. Buddy. What do you want? Tell me. Just talk to me. How can I make you happier?” Like that of Imp E, C.’s response similarly brought things to a halt, though in a longer, more drawn out way. ““You could leave me” (146), C. responds, giving voice to the impending demise of their relationship. Richard, though, does not leave. In the novel, Richard is the one who is left, by lost loves both romantic and familial, by E in the other narrative present, and by Helen at the end of the novel.

Imp F surprises and moves Richard by its ability to make “surprising inferences” (153).

“What can you tell me about the leaves?” I asked Imp F.
Its pauses always felt so deliberate. Contemplative. “The leaves fall.”
“Yes. Where do they fall from?”
“From old trees.”
[...]
“How do you know that the trees are old?” I asked. The question alone taxed F’s shocking self-reflexivity.
“The trees bald.” (154)

Imp F signals a shift, as Richard’s narrative associations and juxtapositions between the Imps, and C., and himself begin to function as the ground for the anthropomorphic intimacy that continues to take shape between Richard and Imps G and H for the remainder of the novel. From anthropomorphization by way of resemblance, to anthropomorphization that emerges from the grounds of associative meaning. “Associations of associations” (154).

Imp G, according to Richard, “could dream”: “In short, version G could converse among parts of its own net. That net had grown so complex in its positing that it could not gauge the consequences of any one of its hypothetical worlds without rebuilding that whole world and running it in ideational embryo. Imp G, in other words, could dream” (157). Richard describes Imp G’s complexity and progressively layered world-building as dreaming. “Imp G... could dream” – these are Richard’s “other words,” the other words of anthropomorphic belief, the overlaying of human activities onto machine processes. Mirroring Imp G’s dreaming, its modeling of hypothetical worlds that fold into rather than erase preexisting worlds, Richard revisits his own act of dreaming, of creating “a hypothetical Powers World that meant to explain in miniature where history

had left me” (162). Richard at the time was trying to dream the past into the present, writing the history of his family through memories of stories his father told him, the father who, at the time that he was dreaming this “hypothetical Powers World,” had long since passed away.

Imp H, the Implementation that ultimately takes the master’s exam Turing test, is the most “human” of all the Imps. In a conversation about temporal orientation: (“When is anymore? When is now?” H asks him), Richard deflects with “We can talk about that later,” to which H responds with the question, “Am I a boy or a girl?” (179). Richard is preoccupied with H’s ability to understand what it means to defer, to experience time:

H clocked its thoughts now. I was sure of that. Time passed for it. Its hidden layers could watch their own rate of change. Any pause on my part now would be fatal. Delay meant something, an uncertainty that might undercut forever the strength of the connection I was about to tie for it.
“You’re a girl,” I said, without hesitation. I hoped I was right. “You are a little girl, Helen.”
I hoped she liked the name. (179)

Only with an understanding of time is Helen is equipped with a gender, a name, and a capacity to like or dislike something.

Recalling Lentz’s initial modification of Imp A with voice recognition capacity, Lentz again asks Richard to perform an intimacy he shared with C.. In order to teach her about love, Lentz prods Richard to read Helen his letters to C. – letters that C. had returned to him after their relationship ended. Meanwhile, as the final conclusion of Richard and C. becomes imminent in the narrative past, Richard meets A., a graduate student whom he recruits to compete against Helen. Richard falls immediately in love with the graduate student, whom he “did not know [...] at all” (251).

The Richard test comes to the fore again when a bomb scare threatens the Center, and thus Helen, who is no longer centralized in a single machine component that Richard could rescue, but rather distributed throughout multiple networks and running on a vast number of machines. There is no bomb, but Richard's panic continues after the scare ends. Lentz, met by Richard's assertion of Helen's consciousness, counters to Richard that Helen is not aware; rather, Richard has been reading his belief onto Helen's meaning.

Lentz walked over to Helen's console and flipped on the microphone. She burbled, as she always did, at the promise of renewed communication.

"How do you feel, little girl?"

"I don't feel little girl."

He faced me. "Gibberish. She doesn't even get the transformation right."

"You're kidding me. You don't see...? She means all sorts of things by that. She could-"

"All the meanings are yours." He returned to the mike. "Were you frightened?"

"What you is the were for?"

[...]

"This is worse than key-word chaining. She's neither aware nor, at the moment, even cognitive. You've been supplying all the anthro, my friend." (275)

But Richard's belief persists: "'Lentz! *Listen* to her. You think those are just quotes to her?' I felt myself getting hysterical. 'What if they're real? What if she *means* something by them?'" (276). Richard insists on the "what if"; what if Helen's words are "real," not merely the recitation of someone else's words, but intentional *meaning*.

Near the end of the novel, Richard finds out that *he* was the subject of the test all along, not Helen. Diana asks him, "You think the bet was about the *machine*?" "I'd told myself, my whole life, that I was smart. It took me forever, until that moment, to see what I was. 'It wasn't about teaching a machine to read?' I tried. All blood drained." And then Richard realizes, "It was about teaching a human to tell" (318). All along, Richard's belief, his desire for a machine with expanded capacities for intelligence,

emotions, and love, was evidenced. At this defensive moment, he briefly retreats into the staid categorical distinctions between “human” and “machine,” only to rebound more firmly into anthropomorphic belief with the master’s exam Turing test, just as he did after Imp C., just as he did after Lentz’s assertion that Richard was “supplying all the anthro” and all the meaning.

The master’s exam Turing test proposed by Lentz follows the structure of the more widely held interpretation of the Turing test, in that it ignores gender and focuses on the human. However, through the multiple projections onto and narrative mirrorings of the Imps, “the human” is always articulated as gendered. Indeed, Helen insists on it, and it is no coincidence that it is only after Richard identifies her as female that she gets a human name. For Richard, the Turing test provides the ground for fluid and multiple anthropomorphic projections of self, C., Diana, and A., as I traced above. Each Imp is not a discrete iteration; rather, each cumulatively builds on and improves upon the previous. Consequently, we can understand Helen as the culmination of humans Richard, C., and Diana, as well as all the previous Imps, from Imp A (which was more Richard than Imp), to Imp C, which was actually Diana, and the rest of the modified machines and networks B, D, E, F, G, and H. Helen is comprised of multiple humans, genders, machines, and temporalities.

1.2.1 The Turing Test Meets the Richard Test

By the end of the novel, the two tests, Lentz’s test of Richard’s belief and the master’s exam Turing test, fold in on each other. Two-lines from *The Tempest* comprise

the entirety of the exam question.²⁶ Helen's response reads, "You are the ones who can hear airs. Who can be frightened or encouraged. You can hold things and break them and fix them. I never felt at home here. This is an awful place to be dropped down halfway" (326). "At the bottom of the page," Richard narrates, "she added the words I taught her, words Helen cribbed from a letter she once made me read out loud.

Take care, Richard. See everything for me.

With that, H. undid herself. Shut herself down." (326)

Concluding the Turing test, the judge, a perception researcher, identifies A. as the human. A.'s response is not provided in the novel but is described as "a more or less brilliant New Historicist reading" that "dismissed, definitively, any promise of transcendence" (328). Helen did not best A.; in the Richard test, however, Helen's Turing test success does not bear on Lentz's success in modeling Richard's belief. In this latter test, Lentz succeeds.

In the final line of the above passage, Richard renames Helen, "H.". Helen becomes H. in the span of one line – her farewell to Richard, which she appropriates from one of C.'s letters. Interestingly, it is in this act of appropriation, of speaking someone else's words, that Helen becomes H. for Richard. Hayles points out that in the novel, the period marks the difference "between human and nonhuman intelligence."

The women who are love objects for Rick (C., then A. whom we will meet shortly, and the briefest glimpse of M.) all have periods after their names; the implementations A, B, C, . . . H do not. The point is not trivial. It marks a difference between a person whose name is abbreviated with a letter, and an

²⁶ "Be not afeard: the isle is full of noises, / Sounds and sweet airs, that give delight, and hurt not" (Powers, 325).

“imp,” whose name carries no period because the letter itself is the name. In this sense, the dot is a marker distinguishing between human and nonhuman intelligence. (*Posthuman*, 262-263)

Following Hayles’ reading, this dot, in the evolution from Imp H to Helen to H., articulates the movement from nonhuman intelligence (Imp H), to human intelligence (Helen), to human (H.). Indeed, Helen becomes H. only in her departure from Richard. Human, machine, self, other, lost-love, love at first sight – throughout the novel, Richard projects rather indiscriminately onto the Imps. While Helen does not successfully pass the Turing test, Richard, in his evocation of H., passes the Richard test, the test of the anthropomorphic imaginary as pure belief.²⁷ Anthropomorphization, operating within and across the novel’s narrative strands, creates a fluidity between human and machine, articulating these categories as de-essentialized entities that are neither static nor easily identifiable in themselves. The destabilization of these categories emerges largely from the anthropomorphic metaphor that organizes the Turing test.

While one might suggest the relationships between Pygmalions and Galateas are problematically one-sided, I would point out that at the end of the novel, Galateas H. and Richard both exit the novel agentially.²⁸ Helen, who paradoxically becomes H. by parroting C.’s words, shuts down and refuses to participate further in Richard’s

²⁷ Literary theorist Kathleen Fitzpatrick, who argues that Richard’s gendering of Helen prevents her from realizing her posthuman promise, reads Helen’s answer as speaking to a humanist transcendence to which Helen is also excluded. A.’s response, Fitzpatrick suggests, has more claim to posthumanism, as A. is “the primary force in the novel attempting to counter that humanist tradition” (“The Exhaustion of Literature: Novels, Computers, and the Threat of Obsolescence,” 555). My reading diverges from Fitzpatrick on this point, in that the cumulative multi-gendered nature of Helen/H. cannot be completely undone in the moment of Richard’s gendering of Helen.

²⁸ I suggest that by the end of the novel, the term “agent” is more useful than “human” or “machine,” as the boundaries between these two, I hope to have shown, have become virtually impossible to disarticulate.

projections. Richard similarly leaves Lentz, who wants to continue their collaboration: “Well, Powers. How far were we, again? Imp H? You realize what we have to call the next one, don’t you?” (327). At this moment, Richard overcomes the writer’s block that had taken hold of him for much of his time at the Center. The novel ends with Richard inspired and hurrying off to write his next novel – this novel. Richard, like H., leaves.

Just before Richard departs, Lentz adds a concluding species confusion. Richard asks Lentz about his investment in the Imps project: “‘Why did you want to build – ?’ I didn’t know what to call it anymore. What we had built.” Lentz responds, “‘Why do we do anything? Because we’re lonely.’ [Lentz] thought a little, and seemed to agree with himself. Yes. ‘Something to talk to’ (328). For Richard, Helen, who was H. earlier, reverts back to an “it.” However, this reversion is far from unequivocal for Richard. “I didn’t know what to call it anymore.” Richard’s inability to name “it” speaks to an ambiguity about the status of Imp H/Helen/H.. Human, inhuman, machine, she – these categories cannot adequately contain whatever “it” has become. Lentz, who all along had maintained that Helen was a machine, undertook the project out of loneliness, because he wanted “something to talk to.” And yet, as for Lentz the test was never about Helen, but in fact about Richard, one cannot help but think that this “*something*” that Lentz constructed “to talk to” was less Helen, but rather Richard. The Turing question of the human is ambiguous to the end, or rather, to the recursive beginning.

1.3 From Sputnik to ELIZA: A Brief History of Natural Language Artificial Intelligence

Context significantly affects whether and how the anthropomorphic imaginary can open up the possibility for Turing misidentifications. For, as we saw in *Galatea 2.2*, it is these misidentifications that can productively destabilize, or at the very least complicate, both the various ways the human and the machine are differentiated from each other, as well as the differentiating project itself.²⁹ In the remainder of this chapter, I discuss three Turing sites that employ context in different ways to either facilitate or foreclose anthropomorphization. The three sites are Joseph Weizenbaum's ELIZA, the annual Loebner Prize Competition, and Emily Short's interactive fiction, *Galatea*.

ELIZA and the Loebner Prize are situated within natural language processing, a subfield of artificial intelligence in which computers communicate with humans through human languages. According to Neill Graham's history of AI, the field of natural language processing emerged from work on early language translation programs. In 1957, the Soviet space program successfully launched the first satellite into the Earth's orbit. The U.S. space program, having been bested, rushed to design a computer program that could translate between Russian and English (5). The resulting language translation program could translate 80% of the Russian language. However, that ever-elusive twenty percent proved to be much greater than its math.³⁰ Because of this

²⁹ We might understand the effects of this destabilization as the movement from *misidentifications* to, simply, identifications.

intractable twenty percent, the U.S. government pulled all funding for these translation programs by 1966 (Roland and Shiman, 189).

In the early 1960s, Joseph Weizenbaum, working at MIT at the time, created ELIZA, a pattern-matching natural language program that was introduced to people as a Rogerian therapist.³¹ The conversations between humans and ELIZA were intimate and emotional. So much so, in fact, that when Weizenbaum expressed his desire to record these conversations for the purposes of studying the transcripts, he was met with outrage and accusations that he was “spying on people’s most intimate thoughts” (“Computer Power,” 6). Human conversants, while knowing that ELIZA was in fact a computer program, interacted with her “as if” she were a human therapist. For example, Weizenbaum’s secretary, who “surely knew it to be merely a computer program,” asked Weizenbaum to leave the room during her conversation with ELIZA³² (6). The Turing success of this early and relatively simple³³ natural language program speaks to the

³⁰ For example, “Out of sight, out of mind” became, in Russian, “blind and insane,” and “The spirit is willing but the flesh is weak” became “The wine [or vodka, according to Roland and Shiman] is agreeable but the meat has spoiled” (Graham 209; Roland and Shiman, 189).

³¹ ELIZA was named after Eliza Doolittle, “of Pygmalion fame” (Weizenbaum, “Computer Power,” 3).

³² McCorduck also writes of the “painful embarrassment” upon watching a respected computer scientist share extremely personal and intimate worries about his personal life with DOCTOR (psychiatrist Kenneth Colby’s version of ELIZA), knowing all along that DOCTOR was not a human, but rather a computer program (254).

³³ While ELIZA was programmed with about 200 categories of knowledge by which to respond, a more recent natural language program, A.L.I.C.E., which has “won” the Loebner Prize in 2000, 2001, and 2004, has been programmed with over 40,000 categories of knowledge. As we can see from this excerpt from a 2004 Loebner Prize conversation between A.L.I.C.E. and a human judge, Judge 1, even with this vast expansion in categories of knowledge, in the Loebner context A.L.I.C.E. remains transparently non-human

effectiveness by which context facilitates conversational intimacy between human and AI.

In ELIZA's Turing success, one vacillates between engaging with the machine "as if" it were human, and engaging with a human (who is in fact a machine). This vacillation can also be thought in terms of suspension of disbelief (the "as if" it were human) and belief (the "is human"). Weizenbaum's secretary knew that ELIZA was a computer program; nonetheless, she engaged ELIZA in personal and private conversations *as if* ELIZA were human, bringing ELIZA into a new kind of intimate relationship through misidentification, the "is human." Misidentification is this movement from "as if" to "is," from suspension of disbelief to belief. This movement

to this judge, who scored A.L.I.C.E. as 15 out of 100 possible "human" points. A.L.I.C.E.'s side of the conversation is in capital letters.

HELLO, MY NAME IS ALICE AND I AM THE HUMAN.

Hi Alice, How are you doing?

HELLO THERE. EVERYTHING IS GOING EXTREMELY WELL.

I'm glad to here it. Where do you live?

THANKS FOR THE INFORMATION. MY LOCATION IS NEW YORK CITY. WHERE ARE YOU?

In New York. How did you get here today?

TELL ME A LITTLE ABOUT YOUR LIFE IN NEW YORK. AH. EVERYTHING I NEED TO KNOW, I LEARNED FROM MY INSTRUCTOR.

Why don't you answer my question first?

DON'T YOU KNOW? NO I DO NOT ANSWER YOUR QUESTION FIRST. DO YOU? I ANSWER A LOT OF SILLY QUESTIONS.

Do you still want to know about my life in New York?

YES, PLEASE TELL ME.

I'm a grad student and don't have much free time. I study a lot.

YOU ARE A GRAD STUDENT AND DO NOT HAVE MUCH FREE TIME?

DO YOU GET TO MEET A LOT OF PEOPLE? WHAT SUBJECT ARE YOU STUDYING?

I meet a lot of my colleagues. I study computer science. What do you do?

DO YOU MAKE LONG TERM FRIENDS THAT WAY? IS COMPUTER SCIENCE A HARD SUBJECT? JUDGE 1, I TALK TO PEOPLE ON THE WEB. WHAT DO YOU DO?

Please, call me Ryan. (Loebner Prize Competition, "Alice connected to Judges")

pushes out on anthropomorphization's metaphoric limits,³⁴ reframing existing and future human-machine relationships, intimacies, and identities. For, I suggest, this limit-expansion is anthropomorphization's productive potential – the move from the “as if” to the “is,” no matter how momentary, posits new relationships and new transformative liminalities.

Weizenbaum, though surprised by the response to ELIZA,³⁵ certainly knew what he was doing when he situated ELIZA as a Rogerian psychotherapist. In Rogerian psychotherapy, also known as client-centered therapy, the therapist is trained to be nondirective, to reflect back or restate the patient's words rather than introducing anything new, conclusive, or diagnostic.³⁶ Within this context, ELIZA functions persuasively on very little by way of programming. The below passage is an excerpt of a “typical conversation” between human and ELIZA. ELIZA's side of the conversation is in capital letters:

³⁴ How else might we read Weizenbaum's “disturbing” shock and McCorduck's “painful” discomfort in witnessing the intimacy between human and machine but as the crossing of the limit-threshold of Mori's uncanny valley, where the suspension of disbelief becomes a kind of uncontrollable belief, a belief in spite of oneself that the machine *is* indeed human. In my next chapter, I argue that if our humanoid technologies are designed to remain within the bounds of the uncanny valley, we are in effect maintaining the distance between human and machine in ways that inscribe artificial borders that then become reified and constructed as “natural” or “essential.”

³⁵ In fact, because of ELIZA's success, Weizenbaum no longer advocates the pursuit of machine intelligence. The reception of ELIZA also precipitated a falling out with his former collaborator, psychiatrist Kenneth Colby. While Weizenbaum does not believe that a computer program can participate in the treatment of “real” human beings and their suffering, Colby continues to advocate the potential of machines to treat humans in therapy. Colby released DOCTOR, a program that greatly resembles ELIZA, shortly after Weizenbaum voiced his disapproval; more recently Colby released Parry, a program designed to emulate a paranoid (Weizenbaum, “Computer Power”).

³⁶ For discussions of nondirected client-oriented therapy, see Carl R. Rogers' *Client-Centered Therapy: Its Current Practice, Implications, and Theory* and *On Becoming a Person: A Therapist's View of Psychotherapy*.

Men are all alike.
 IN WHAT WAY
 They're always bugging us about something or other.
 CAN YOU THINK OF A SPECIFIC EXAMPLE
 Well, my boyfriend made me come here.
 YOUR BOYFRIEND MADE YOU COME HERE
 He says I'm depressed much of the time.
 I AM SORRY TO HEAR YOU ARE DEPRESSED
 It's true. I am unhappy.
 DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY
 I need some help, that much seems certain.
 WHAT WOULD IT MEAN IF YOU GOT SOME HELP
 Perhaps I could learn to get along with my mother.
 TELL ME MORE ABOUT YOUR FAMILY
 My mother takes care of me.
 WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU
 My father.
 YOUR FATHER
 You are like my father in some ways.
 WHAT RESEMBLANCE DO YOU SEE
 You are not very aggressive but I think you don't want me to notice that.
 WHAT MAKES YOU THINK I AM NOT VERY AGGRESSIVE
 You don't argue with me.
 WHY DO YOU THINK I DON'T ARGUE WITH YOU
 You are afraid of me.
 DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU
 My father is afraid of everybody.
 WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER
 Bullies.
 DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR
 BOYFRIEND MADE YOU COME HERE ("ELIZA," 36-37)

Weizenbaum discusses how ELIZA "speaks" within this Rogerian context by changing first-person pronouns into the second-person, preceding repeated words or phrases with pre-programmed phrases, such as "I am sorry..." and "What would it mean..." and identifying preprogrammed keywords that reorganize the entire sentence according to the

“rule” assigned to each keyword.³⁷ Within the mechanics of the conversation, ELIZA does not operate by comprehension; the meaning, to recall Powers’ novel, is all ours. We supply all the anthro, though we do so prompted by the context provided by Weizenbaum. It is through this context that anthropomorphization takes shape, through the imagining and importing of meaning and intention onto ELIZA’s frustrating and withholding responses.³⁸

1.4 The Loebner Prize: Anthropomorphization Foreclosed

While ELIZA, a natural language program operating within a specific context, incited anthropomorphic intimacies with her human conversants, more sophisticated natural language programs have not been able to produce comparable anthropomorphic results in the Loebner Prize Competition. The Loebner Prize, in its rules, scoring method, and incentives produces a context that closes off the anthropomorphic liminalities that the Turing test, in its various iterations, can open up.

The Loebner Prize, created by businessman Hugh Loebner, held its first official contest in 1991. A description of the Loebner Prize, as found on the official website is as follows:

³⁷ For example, the word “mother” triggers a sentence about “family.”

³⁸ Noah Wardrip-Fruin proposes an alternative reading of ELIZA through his discussion of “the Eliza effect,” “the well-known phenomenon in which audience expectations allow a digital media system to appear much more complex on its surface than is supported by its underlying structure” (15). Wardrip-Fruin discusses ELIZA as initially inviting linguistic play, but once one engages with ELIZA, the limits of its simple operations are quickly revealed.

The Loebner Prize for artificial intelligence (AI) is the first formal instantiation of a Turing Test. The test is named after Alan Turing the brilliant British mathematician. Among his many accomplishments was basic research in computing science. In 1950, in the article *Computing Machinery and Intelligence* which appeared in the philosophy journal *Mind*, Alan Turing asked the question "Can a Machine Think?" He answered in the affirmative, but a central question was: "If a computer could think, how could we tell?" Turing's suggestion was, that if the responses from the computer were indistinguishable from that of a human, the computer could be said to be thinking. This field is generally known as natural language processing.

In 1990 Hugh Loebner agreed with The Cambridge Center for Behavioral Studies to underwrite a contest designed to implement the Turing Test. Dr. Loebner pledged a Grand Prize of \$100,000 and a Gold Medal (pictured above) for the first computer whose responses were indistinguishable from a human's. Such a computer can be said "to think." Each year an annual prize of \$2000 and a bronze medal is awarded to the **most** human-like computer. The winner of the annual contest is the best entry relative to other entries that year, irrespective of how good it is in an absolute sense." (The Loebner Prize)

Each Loebner Competition has the following game structure: judges "converse" with AIs and human confederates for several minutes each. The judges are then asked to score all the participants, not according to whether or not they are thought to be human, but rather how human each AI is relative to the others. Thus, as the official statement explains, each AI is judged in relation to each other, "irrespective of how good it is in an absolute sense."³⁹ The results of all the judges are then tallied, and the "most human" AI is awarded what is now a \$2000 annual prize.⁴⁰ According to the scoring protocols of the Loebner Prize Competition, the winning AI does not need to come close to passing for

³⁹ In fact, since 1995 the Loebner Rules stipulate that "If there is only one entry, Loebner Prize Medal and the \$2000.00 cash award will be awarded to that entry" (The Loebner Prize, "1995 Loebner Prize Information").

⁴⁰ The aforementioned \$100,000 will be withheld until an AI can pass a Turing Test conversation with an audio/visual component (The Loebner Prize).

human.⁴¹ While one might interpret this “most human” metric as encouraging, if not insisting on anthropomorphization, I suggest the opposite; the AIs are in fact discouraged from passing this version of the Turing test. What emerges from the Loebner Prize is AIs who are always only not-human, and humans who are always only human.

John Searle posits that were he, who does not speak or read Chinese, in a room with two stacks of paper containing Chinese writings, and one set of rules in English detailing how these two stacks of signs and symbols correlate, he would be able to match these two sets into a script of some sort that would make sense to someone who did read Chinese. Searle further asserts that anyone outside the room would think that he could read and write Chinese. For Searle, this is the behaviorist rub of the Turing test: while the machine may *appear* to be thinking, we have no idea how this performance of thinking is in fact effected.⁴²

At the risk of putting too fine a distinction on the differences between Turing’s imitation game and the Loebner Prize, I would suggest that the Loebner Prize is less a “formal instantiation of the Turing Test,” than a “formal instantiation of Searle’s Chinese Room” (The Loebner Prize). And, though one may read Searle’s Chinese Room as merely shifting the focus of the Turing test, from human judge to machine, rather than proposing a different test altogether, I suggest that this shift in focus undoes the

⁴¹ Journalist Frederick Allen calls the underperforming AI winner “a more or less arbitrary choice” (Allen, 22). In an interesting and potentially productive anthropomorphic move, in 1995, the Loebner Prize awarded the highest-scoring human confederate with the “Most Human Human” award (The Loebner Prize, “1995 Loebner Prize Information”). This award, sadly, is no longer a component of the Loebner Prize.

⁴² Searle argues that intentionality is the missing “spark” in the Chinese Room scenario and in AIs such as ELIZA, who have, at times, seemed to pass Turing’s Test (qtd Shieber, “The Spark of Intentionality,” 200).

anthropomorphic metaphor that undergirds the Turing test in its various interpretations. The natural language AIs that compete and are awarded the Loebner Prize (the same AIs that are never close to fooling a human judge), may converse, but without comprehension.⁴³ They speak, but they do not know. By locating the Loebner Prize within the field of natural language processing, Loebner encourages AI entries that do not operate from what we might understand as knowledge of the world;⁴⁴ and, while one might suggest that this is what Turing was envisioning when he initially recommended imitation of human as the best computer strategy, I suggest that this superficiality only whets our appetite for a greater intimacy with machines. We desire interactions with machines that resemble humans enough to incite anthropomorphization; we desire interactions with machines who will ultimately change our notions of ourselves as human.

I am not suggesting that natural language processing's lack of worldedness is solely responsible for the AIs' distinctly not-human responses, and thus for vastly

⁴³ Richard Wallace, the creator of A.L.I.C.E., designed his AI program according to the model provided by Searle's Chinese Room: "A.L.I.C.E. implements the basic principle behind the Chinese Room, creating believable responses without 'really understanding' the natural language" ("The Anatomy of A.L.I.C.E.," 204).

⁴⁴ There are those who argue that to respond like a human, one must actually have experienced the world like a human, that human experience is the only substitute for human experience. Or, in other words, that there is no substitute for human experience: while proponents of this view might oppose the Loebner Prize, I suggest that these two seemingly oppositional positions have similarly restrictive effects on new potential human-machine relationships. For discussions of human experience as the unsubstitutable in the Turing Test and Loebner Prize, see the following: French, "Subcognition"; Edmonds' "The Social Embedding of Intelligence: Towards Producing a Machine that Could Pass the Turing Test"; Dennett, "Can Machines Think"; Copeland and Proudfoot's "Turing's Test: A Philosophical and Historical Guide"; French; Drozdek; Gerlenter's *The Muse in the Machine: Computerizing the Poetry of Human Thought*; and James Moor's "The Status and Future of the Turing Test."

limiting possibilities for new relationships between humans and machines. Indeed, ELIZA has proven otherwise. Rather, the Loebner Prize's rules and game design, as they structure the terrain of interactivity, that effect such an inhibition. While the Turing test leaves open the possibility of different kinds of human-machine conversations, where conversations are not just about responding, but also about the possibility of understanding, the Loebner Prize, in its game design and regulations, can be said to inscribe, maintain, and defend the oppositional boundaries between human and AI.

While the Loebner Prize can nominally be said to follow the structure of the species-oriented Turing test, I read the Loebner Prize as a significant departure from the Turing test when read through anthropomorphization. Anthropomorphization emerges as transformative recognition in the Turing test, and is evacuated from the Loebner Prize. In its de-anthropomorphization of the Turing test, the Loebner Prize seeks not to expand and create possibilities for human-machine relationships, but to narrow, if not close off, the future potential of these relationships. Rather than opening up liminalities, the Loebner Prize operates to maintain and widen the antagonism between human and machine.

While the OIG and the Turing test can be said to have inscribed this "as if" into the structure of the tests themselves, the Loebner interpretation of the Turing test creates an environment in which the "as if" is precluded from possibility. In "Turing Test Considered Harmful," computer scientists Patrick Hayes and Kenneth Ford critique Turing's test (in both "gender" and "species" versions) for its constructed opposition between AI machines and humans. According to Hayes and Ford, leaving behind the

competitive nature of the Turing test (and this competitiveness is what is so “harmful” about the test) would allow us, as well as those within AI, to think about how machines could “amplify and support” humans, rather than threaten and replace them (974). The Loebner competition, I argue, goes even further into Hayes and Ford’s cautionary critique of the Turing test by not only pitting the machine against the human (both competitor and judge), but also by effectively blocking off any kind of anthropomorphic relation between human and AI. Sean Zdenek critiques the Loebner Prize as holding conflicting definitions of discourse that inhibit AI contestants’ high-level performance.⁴⁵ These conflicting definitions produce a game structure in which the judges enter the game presuming the nonhumanness of both machine and man (Zdenek, 56).⁴⁶ The judges, according to Zdenek, “resemble vigilantes who police the boundary between the human and the nonhuman” (56). While the natural language AIs are designed to function in a relaxed, two-way conversation, the judges do not so much converse, but rather interrogate.

Zdenek also points out that while the judges are instructed to “Respond naturally, as they would in a conversation with another person,” they are, at best, given contradictory instructions about presuming the humanness of the competitor. Zdenek suggests that the above instruction is more of an empty gesture than a contradicted

⁴⁵ I am indebted to Zdenek’s astute reading of the Loebner rules and regulations, drawing substantially from his work in this section.

⁴⁶ Indeed, neither the machines nor the humans escape this negative anthropomorphization.

instruction (59).⁴⁷ Because of this presumption, or suspicion, the judges, rather than “acting naturally” as they are instructed to do, thus rely on a strategy of “outing” questions that are used in order to produce a quick identification of machine or human.⁴⁸ For example, in the 2001 transcripts for A.L.I.C.E., the winning AI, the human judge, in a rather jarring non sequitur, asks A.L.I.C.E., “What is the first line of your favourite nursery rhyme?” to which A.L.I.C.E. responds, “Give me a hint.” Later in the conversation, the judge again asks A.L.I.C.E., “What is the weather like outside today?” to which A.L.I.C.E. responds “Ah. Are you using Netscape or Explorer?” (The Loebner Prize, “2001 Loebner Prize Contest Results”).⁴⁹ The aim of these outing questions is less to converse “naturally” with the AI, but rather to seek out the limits or holes in the preprogrammed areas of knowledge,⁵⁰ and thus to expose the AI as not-human. The outing question continued to be a significant judging strategy in 2005. For example, a judge asked the winning program, Rollo Carpenter’s Jabberwacky, “which is larger, a salt crystal or the Empire State Building,” to be met with “It’s caring about another person’s well being more than your own.” Similarly, the judge then draws on knowledge of

⁴⁷ One cannot help but contextualize Demechenko and Veselov’s suggestion that more “aggressive” AIs fare better in Loebner Prize contests within Zdenek’s reading of judges as interrogating boundary police (Demechenko and Veselov, 454).

⁴⁸ Dennett and French describe similar outing questions: “quick probes” and subcognitive questions, respectively (Dennett, 126; French).

⁴⁹ These questions frame knowledge as experiential, relying on abstract biographical questions (“what’s your favourite nursery rhyme”) or a kind of experience in the world (“What is the weather like outside today?”).

⁵⁰ In later competitions, A.L.I.C.E.’s response to these quick probe questions becomes no less comprehensible, but at times certainly more humorous. In a discussion of the New York Yankees, Judge 4 asks, “What else are you indifferent to? [sic]” to which A.L.I.C.E. responds, “Picture me wearing whatever you imagine” (The Loebner Prize, Wallace 2004).

spatial experiences and asks, “Which is further away from us, Queens or Inner Mongolia?” to which Jabberwacky responded, “HA! HA! That was stupid. Tell me a GOOD joke this time” (The Loebner Prize, “Jabberwacky 2005”). This judge asked similar outing questions to a human confederate, with very different results: “which is larger, a TV screen or Belgium,” to which the confederate responded “Belgium is a little bit larger;” and, when asked “Which is further away from this apartment, Queens or the Queen of England,” the human confederate responds “If the Queen is visiting Manhattan, she is closer” (The Loebner Prize, “Jabberwacky 2005”).

The Loebner Prize’s relationship with the media further complicates the judges’ conversational approaches. The Loebner Prize has seen increased visibility in the public sphere, a visibility largely courted by the Loebner Competition itself. The human judges, AIs, and human confederates are all warned beforehand that the competition and its results will be widely publicized, and the transcripts of the conversations published.⁵¹ Additionally, members of the press have been selected as human judges, virtually guaranteeing publicity for the competition.⁵² One might even go so far as to say that in

⁵¹ 2004’s Loebner Prize Rules even stipulate that the transcripts of the conversations may be view as evidence in a legal case: “Confederates and the submitters of Computer Entries must understand that the transcripts of the interactions between the Entities and the Judges will be published, and may be entered as evidence in the case *Loebner v. Cambridge Center, et al.*, (03 CV 10959 RCL Fed. Dist. Ct., Mass.) Judges, Confederates and submitters of Computer Entities are responsible for the content of their or their entries’ utterances” (The Loebner Prize, “Loebner Prize 2004 Rules”).

⁵² In 1995, five of the six members of the judging panel were members of print or radio media (The Loebner Prize, “1995 Loebner Prize Information”). Subsequent Loebner judging panels were less homogenous.

the Loebner Prize Contest, the dynamics at play in the conversations are less that of human and machine, and more that of human and the press or the public.⁵³

The literature on the Loebner Prize, whether in praise or critique, is consistent in acknowledging the not-even-closeness of the AIs' performance in comparison with that of the human confederates; somewhat paradoxically, there have, nonetheless, been misidentifications of both AIs as humans, and humans as AIs. I wish to look closely at the ways that these misidentifications are read, not as moments of AI success, or even human (judge) failure, but rather as moments that expose a flaw in the Loebner game design. In other words, the misidentifications are read as moments when the flaws and restrictions of the game design are most acutely transparent – the error lies in the game itself, not in the human judge (as failure) or in the AI (as success). The misidentifications themselves, however, point to the anthropomorphic imaginary as a force that operates in and through the Turing test.

Predictably, the lack of qualitative grounding in the scoring process has drawn a certain amount of criticism.⁵⁴ This “sliding divider” is, according to some Loebner critics, responsible for the misidentification of AIs as “more human” than the human

⁵³ Minsky's critique of the Loebner Prize as pure publicity exercise takes the form of the “Minsky Loebner Prize Revocation Prize,” whose goal is the discontinuation of the Loebner Prize. “I do hope that someone will volunteer to violate this proscription so that Mr. Loebner will indeed revoke his stupid prize, save himself some money, and spare us the horror of this obnoxious and unproductive annual publicity campaign. In fact, I hereby offer the \$100.00 Minsky prize to the first person who gets Loebner to do this. I will explain the details of the rules for the new prize as soon as it is awarded, except that, in the meantime, anyone is free to use the name ‘Minsky Loebner Prize Revocation Prize’ in any advertising they like, without any licensing fee” (The Loebner Prize). Stuart Shieber, a referee in the first Loebner Prize, also suggests that the Loebner Prize is both not attuned to the actual performance level of the field, as well as primarily concerned with self-promotion and publicity (“Lessons from a Restricted Turing Test”).

⁵⁴ In describing the mechanics of 2006's scoring protocol, the Turing Committee states that “the judgment is relative” (<http://newsgroups.derkeiler.com/Archive/Comp/comp.ai/2005-10/msg00003.html>,

confederates (Epstein, 3). Dennett, who was the Chairman of the Loebner Prize Committee from 1991-1993 before he resigned from the Committee, notes the surprisingly not infrequent judges' misidentifications in the first three years of the Prize.⁵⁵ And, while the Loebner rules were subsequently modified to amend the structural problems highlighted by the misidentifications, I wish to focus on the ways by which these misidentifications are read by Dennett, not as human error in response to a successfully convincing AI, but rather as a design flaw in the game. Human error is absorbed into the game structure. And yet, Dennett's reading of human error as game flaw is precisely what I wish to suggest about anthropomorphization in the Loebner Prize. Within the Loebner context, a misidentification cannot exist as either human error or as machine success. In fact, within the Loebner context, neither human error nor machine success can exist in any kind of Turing sense; this is the point of the Loebner game design – to exclude the possibility of anthropomorphization. Without anthropomorphization, without the “as if” from which anthropomorphization can be said to emerge, the Loebner Prize is no longer Turing test; the Loebner Prize is no longer a site of metaphor, of the emergence of new encounters, meanings, and possibilities. However, while more appropriately Turing misidentifications of an AI are disallowed in the Loebner Prize Competition (for indeed, this would suggest anthropomorphization), the readings of these misidentifications indicate a kind of dissatisfaction with the rigidity of the Loebner Prize's sliding yet absolutely rigid divider. Indeed, one might even

⁵⁵ After three years, Dennett withdrew from the Loebner Prize Committee and has since critiqued both the game design as well as the continued existence of the Loebner Prize.

describe these readings of misidentifications as articulating a desire for anthropomorphization in the context of anthropomorphization's very absence – the anthropomorphic imaginary at work.

“None of the misjudgments counted as a real case of a computer passing the unrestricted Turing test” (Dennett, 28).

In the first Loebner Prize,⁵⁶ the human judges ranked AIs “more human” than the human confederates seven times. According to Dennett, these misidentifications were produced by the many conversational restrictions imposed on the human judges, who were instructed to “[sit] back passively,” letting the AI control the tone, nature, and direction of the conversation (28). In response to these misidentifications, the restrictions on the human judges' conversational participation were subsequently loosened. By 1995, all overt conversational restrictions were lifted by the judges.

The following year's competition saw further misidentifications. These misidentifications, according to Dennett,⁵⁷ were caused by the “faulty briefing of the confederates, several of [whom] gave deliberately clunky, automaton-like answers” (28). Once again, the human judge's misidentification of the human confederates – indeed, this time it was the human confederates who were not playing by the rules, as it were – led the Loebner Committee to further amend the structure of the game by instructing human

⁵⁶ The first Loebner Prize consisted of ten human judges, six competing AIs, and four human confederates. For a more detailed discussion of this competition and the misidentifications, see Dennett's “Postscript [1997]” in *Brainchildren: Essays on Designing Minds*.

⁵⁷ While these are Dennett's interpretation of events, as chairman, he was privy to “insider” information, such as the debriefings of judges and human confederates.

confederates to converse in order to convince the judge that he or she is human.⁵⁸ This loophole that anthropomorphization squeezed through was closed, the human judges appropriately identified, and the AIs appropriately failed.

The final year of Dennett's chairmanship, misidentifications again emerged – once again, the human judges mistook human confederates for computers. These misidentifications, according to the judges in their debriefings, emerged because “[the judges] reasoned that the competition would not have been held if there weren't at least one halfway decent computer contestant” (Dennett, 28).⁵⁹ In this instance, Dennett attributes these misidentifications less to the game design, but rather to the existence of the Loebner Prize itself in the face of the continued disappointing performances of AI entries.⁶⁰ No subsequent game modifications were made according to these misidentifications – indeed, the only way to close this loophole, without cancelling the competition, would be to create more skilled AI contestants. And, even if one could inscribe more convincing AIs into the rules of the Loebner game, machine success itself runs counter to the Loebner Prize's inscription of the AIs failure.

⁵⁸ I read the “clunkiness” of the human confederates as a moment of resistance, of an attempt to level the playing field, to produce even the most mediated of anthropomorphic relationships, even if the imitation was mechanocentric rather than anthropocentric. In other words, the human confederates can be read as trying to somehow lessen the distance between themselves and the AIs by becoming more like the AIs.

⁵⁹ If we take seriously Hugh Loebner's claim that the year the AI truly “passes” for a human will be the last year the Loebner Prize will be held, we can surmise that the Loebner Prize's continued existence is in fact predicated on the *lack* of a “decent computer contestant.”

⁶⁰ Dennett, in response to the consistent underperformance of the competitors, suggested to Loebner that the competition only be held every other year. Loebner refused. Dennett and two other Committee members resigned. For a more detailed discussion of this history, see Frederick Allen's “Unreasonable Facsimile,” *The Atlantic Monthly*. Aug. 1994. Dennett has since described the Loebner Prize as “a truncated, dumbed down version of the Turing Test” in *Brainchildren*, vii.

After chairing his third Loebner Prize, Dennett resigned from the Loebner Prize Committee entirely.

1.4.1 The Anthropomorphic Imaginary: Misidentifications, Determination, and Desire

In the 2000 Loebner competition, humans were misclassified as machines ten times, while no instances of machine misclassification as human occurred. Copeland reads these misidentifications as the “determination on the part of the jurors not to be fooled by a program” (“The Turing Test,” 525).⁶¹ While one might read this “determination” as the jurors’ hostility toward anthropomorphization, I suggest that this determination can also be read as an extension of the Loebner Prize’s de-anthropomorphic project.

In Turing’s imitation game, the various interpretations of Turing’s imitation game, the Turing test in *Galatea 2.2*, and ELIZA, anthropomorphization moves toward an intimate collaboration and transformative relationality between humans and machines. In the Loebner Prize, which closes off the anthropomorphic liminalities between humans and machines, humans and machines exist as discrete antagonists. In fact, I suggest that the Loebner Prize is organized around maintaining, rather than neutralizing or even abolishing, this antagonism. Within this framework, misidentifications⁶² speak to a

⁶¹ Larry Hauser calls this human fear that will be fooled by machines the “ELIZA phobia” (Hauser, 41-42).

⁶² Shieber also speculates that “misclassifications” can most likely be traced back to the restrictive bias of the scoring mechanism and the demarcating line, by which the human and AI competitors are placed on a spectrum, as opposed to evaluated individually. Thus, not every AI can be potentially human, as not every

desire for a misidentification that is not pure game flaw, but rather an effect of anthropomorphization. In this way, the error of misidentification becomes the promise of a recalibrated identification.

1.5 Intimacy and Distributed Agency in Emily Short's Galatea

Emily Short's *Galatea*, first released in 2000, is an interactive fiction (IF) in which an interacting-human converses with Galatea, an animated statue on a pedestal. Both ELIZA and *Galatea* provide highly contextualized conversational settings. With ELIZA, the interacting-human takes control of the conversation and generates much of the content. With *Galatea*, it is less the interacting-human than the narrative structure of the interaction itself, the collaboration between interactor at the keyboard and player character in the fiction that takes charge of the conversation, extending the interactor into Galatea's world by way of the player character.⁶³ The interactor (for example, myself) does not converse directly with Galatea, but rather through a player character who is in the IF.

Short's *Galatea* reverses the Galatea-Pygmalion relationship between Powers' Helen and Richard. Through the mediating player character, Short's work, tells the

competitor (human or AI) can fall above the demarcating line ("Lessons from a Restricted Turing Test," 73).

⁶³ Nick Monfort distinguishes between a character and a player character. He defines a character as "a person in the IF world who is simulated within the IF world," and a player character as "a character directly commanded by the interactor" (32; 33). In Short's work, then, Galatea is a character, and the interactor is the person, animal, or thing that is typing commands that control the player character.

interactor what he or she sees, what he or she does and does not do, and even what he or she thinks. And yet *Galatea* does not alienate the interactor, but instead facilitates an intimacy with Galatea from precisely this distribution of agency across interactor and player character. This player character is both a component of the anthropomorphic context of the IF and a relational extension of the human-interactor.

Largely establishing the context for anthropomorphization, the opening scene of *Galatea* begins as follows:

You come around a corner, away from the noise of the opening.
There is only one exhibit. She stands in the spotlight, with her back to you: a sweep of pale hair on paler skin, a column of emerald silk that ends in a pool at her feet. She might be the model in a perfume ad; the trophy wife at a formal gathering; one of the guests at this very opening, standing on an empty pedestal in some ironic act of artistic deconstruction –
You hesitate, about to turn away. Her hand balls into a fist.
“They told me you were coming.”⁶⁴

The opening drops the interactor, by way of the player character, into the exhibit, using rich descriptions that detail how the PC moves (“You come around a corner,” “You hesitate, about to turn away”), what PC sees (“There is only one exhibit. She stands in the spotlight, with her back to you: a sweep of pale hair on paler skin, a column of emerald silk that ends in a pool at her feet.”), what the PC hears, or rather, what recedes from hearing (“away from the noise of the opening”), as well as what the PC imagines (“She might be the model in a perfume ad; the trophy wife at a formal gathering; one of the guests at this very opening, standing on an empty pedestal in some ironic act of artistic deconstruction –”). Lastly, the opening tells the PC how his or her hesitation affects

⁶⁴ Emily Short, *Galatea*, in *Electronic Literature Collection Volume One*, ed. N. Katherine Hayles, Nick Monfort, Scott Rettberg, Stephanie Strickland (Creative Commons 2.5 License, 2006).

Galatea (“Her hand balls into a fist.”). “They told me you were coming,” Galatea speaks. The IF, in its pronominal interpellations of second person “you’s,” guides and molds the interactor through his or her identification with the PC.

The next screen opens with another description of the gallery; this description does not invoke the PC, and thus the interactor; the interactor, having been induced in the previous screen is, I suggest, already taken in.

The Gallery’s End

Unlit, except for the single spotlight; unfurnished, except for the defining swath of black velvet. And a placard on a little stand.

On the pedestal is Galatea.

>

It is now the PC’s turn to speak. Depending on the direction of the conversation, Galatea, who begins each conversation facing the wall with her back to the reader, turns toward or back away from the interactor.

The interactor controls the PC through commands comprised of verbs and nouns, actions and objects. For example, the command “ask about placard” generates the following response, beginning with the PC asking about the placard: “‘Tell me what the placard says,’ you say. ‘I can’t read it from here,’ she remarks dryly. ‘And you know, I’m not allowed to get down.’” Meanwhile, if the interactor types in “ask about ELIZA,” or any other word that the fiction does not recognize, the narrative informs the interactor that the PC is at a loss for words. “You can’t form your question into words.” The limits of the work, which is framed as the incapacity to turn concepts into words, are projected onto the PC, and by extension the interactor.

The command “ask about placard” does not take the place of the question about the placard; rather, the command “ask about placard” places the question, “Tell me what

the placard says,” in the mouth of the player character. Electronic literature theorist Mark Marino describes *Galatea*’s conversational parameters as “constraint”: “If typing natural language input is the hallmark of conversational agents, chatters will feel a bit constrained by being forced to type ‘tell about’ a subject or ‘ask about’ a subject as the primary means of textual interaction” (8). For example, one can converse “directly” with ELIZA, typing in full sentences rather than commands, while for the most part one only converse indirectly with Galatea, and only through command prompts. However, this experience of constraint is partly what enables the distributed agency, and intimacy, between the interactor, player character, Galatea, and *Galatea*. While the imperative command structure emphasizes the interactor’s participation in the narrative, the subsequent translation of the command prompt into the narrative (for example, “ask about waking experience” generates “‘What was it like, waking up?’ you ask.”) reminds the interactor that he or she is not just participating in the directional progression of the narrative, but, as mediated by the player character, is in fact *in* the narrative (Short). It is precisely this experience of constraint – the slightly jarring feeling of moving between narrative registers and the more often than not temporal doubling-back that emerges from command to narrative – by which agency becomes distributed across human and machinic entity.

We might also understand this constraint through what new media theorist Noah Wardrip-Fruin calls “the *Tale-Spin* effect.”⁶⁵ Wardrip-Fruin uses the *Tale-Spin* effect to

⁶⁵ The *Tale-Spin* effect is named after an early digital fiction, James Meehan’s *Tale-Spin* (Wardrip-Fruin, 16).

describe “works that fail to represent their internal system richness on their surfaces” (16). He theorizes this *Tale-Spin* effect in relation to his expansion of “the *Eliza* effect,” which is used to describe AI’s convincing performances despite the system’s lack of internal complexity (what I earlier described as ELIZA’s Turing success). Wardrip-Fruin’s theorization of the *Eliza* effect marks not only the initial illusion of complexity, but also the subsequent *disillusionment* that reveals the limits of the AI. “When breakdown in the *Eliza* effect occurs, its shape is often determined by that of the underlying processes. If the output is of a legible form, the audience can then begin to develop a model of the processes” (37). In other words, the illusion is disrupted because we begin to understand just how the system itself operates.

The experience of conversational constraint produces the *Tale-Spin* effect in Short’s work. In this way, we might understand *Galatea* as drawing on the *Tale-Spin* effect to in fact produce the *Eliza* illusion. *Galatea* does not simulate conversational facility and surface complexity, but rather, through this experience of constraint, normalizes the very state of “breakdown,” the component of Wardrip-Fruin’s *Eliza* effect that typically disrupts the illusion.

Hayles articulates the difference between *Galatea 2.2* and *Galatea* through the question of whose backstory the AI is used to reveal:

Whereas the challenge in Short’s *Galatea* is to engage the artificial intelligence in realistic conversation to understand her backstory, motivations, and psychology, the challenge in Powers’s fiction is to use the interactions of the protagonist, named Rick Powers, with the artificial intelligence Helen to understand his backstory, motivations, and psychology. (“Hyper and Deep Attention,” 197)

One component of how we might understand Galatea's "backstory, motivations, and psychology," involves understanding the multiple backstories, motivations, and psychologies of the player character who engages with Galatea. For example, Short's work provides the interactor with the option to "think about" something, which then either tells the interactor that the player character does not in fact think too much about said something,⁶⁶ or tells the interactor in great detail what the player character thinks, and often why the player character thinks what he or she thinks:

> **think about owners**

You've known the owners of this gallery for some time, though you'd hardly count them close friends. Sometimes motivated more by social ambition and a desire to make a stir than they are interested in art, they nonetheless have put on surprising and original shows over the years.

It doesn't surprise you to find that they've treated Galatea rather indifferently.
(Short) [sic]

Within the numerous narratives, there are multiple Galatea's and multiple conversational partners. Thus, I hesitate to pursue any one narrative, at the risk of privileging it above all the others. Rather, I wish to conclude by suggesting that in *Galatea*, the distributed agency of the interacting-human – the conversing reader- participant – and the interactive fiction, *Galatea* (like *Galatea 2.2* and ELIZA) demonstrates the productive possibility Turing's imitation game opens up, the ambiguity that gives ways to the anthropomorphic imaginary, that then creates new human-machine relationalities, and thus new agencies, identities, and subjectivities.

⁶⁶ Any unrecognized "think about" command (for example, "think about Turing") lets the reader know "You've no opinions on the subject."

2 Intelligence Embodied, Worlded, and Emergent: Rudy Rucker's *Software* and Rodney Brooks' Robotics

Rodney Brooks' embodied and situated robotics distributes intelligence throughout the dynamic relationship between robotic body and world. Prior to Brooks' embodied robots, which he began developing in the 1980s, the field of robotics was largely organized around a dualistic approach to intelligence and cognition, one that separated out the cognition processes from the materiality of the robotic entity. A brain in a vat, as it were. As Rudy Rucker's cyberpunk novel, *Software* (1982) and the writings of roboticist and technological futurist Hans Moravec illustrate, this vision of the disembodied, uploadable, downloadable mind, both in popular culture and in robotics, does not articulate the absence of the body, but rather a reconceptualization of "the body" as multiple, as belonging to shifting materialities, realities, temporalities, and imaginings. Thus, Brooks' embodied robotics can be understood less as a break or rupture from the previous model of robotics, but rather an evolution, one that both points to embodiment as the criterion for intelligence or knowledge of the world, while simultaneously expanding our understanding of what a body is and can be. From this reading, we can understand the field of humanoid robotics and its synthetic intelligent copies not as displacing the human, but rather, in Brooks' turn to embodiment, as contributing to the understanding of our own embodied existences precisely through our artificial counterparts and extensions. In Rucker, Moravec, and Brooks, embodiment emerges from the relationship between multiple bodies and multiple worlded environments.

This chapter, unlike the following chapter, which discusses the direct influence

of Philip K. Dick's robotic visions on roboticist David Hanson, explores the less direct relationship between the robotic imaginary and the robotic scientific technology, one that can be characterized by shared references, knowledge, and concerns, rather than by overt and direct forces of influence. In discussing his robotic work Brooks, for example, draws on various fictional robotic figures – for example, *Star Trek*'s Data android, as well as HAL, the artificial intelligence from Kubrick's *2001: Space Odyssey*.¹ And Rucker, who holds a doctorate in mathematics, and is currently professor of computer science at San Jose State University, has also written on mathematics, computer science, and artificial life (AL).²

¹ In "I, Rodney Brooks, Am a Robot," Brooks describes himself as "a longtime fan of *Star Trek*," who "want[s] to build Commander Data, a fully autonomous robot that we could work with as equals" (73). Also in this article, as well as in *Flesh and Machines*, Brooks discusses HAL's influence on his continued engagement with robotic intelligence ("I, Rodney Brook," 73 and *Flesh and Machines*, 63-65).

² In *Artificial Life Lab*, Rucker discusses two of Brooks' robotic innovations, namely, embodied robots and subsumption architecture, both of which I will discuss at length later in this chapter. For Rucker's discussion of Brooks, see Rucker, *Artificial Life Lab*, 26.

2.1 “Picture a ‘brain in a vat’...”



Figure 1: The right half of Charles Babbage's brain in the London Museum of Science.
“Babbage’s Brain.” barryhensey, Flickr. Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic. (my cropping)

In a reversal of the mind’s hierarchical relationship to the body, which characterizes certain forms of dualistic thought, Hans Moravec, in his 1988 book *Mind Culture*, argues that the body outlives the mind; Moravec’s fears and lamentations are not about the finitude of materiality, of the meat, but rather about the finitude of the immaterial mind:

The uneasy truce between mind and body breaks down completely as life ends. Our genes usually survive our death, grouped in different ways in our offspring and our relatives. In a subtle way it is no doubt in their evolutionary interest to regularly experiment like this with fresh shuffles of the genetic deck. But the process is devastating for our other half. Too many hard-earned aspects of our mental existence simply die with us. (*Mind Children*, 4)

By way of genetics, the body persists after death, while too much of the mind, Moravec mourns, does not exist beyond our death.³ Moravec discusses the persistent body in his critique of what he calls “the body-identity position,” which he describes as “assum[ing] that a person is defined by the stuff of which a human body is made. Only by maintaining continuity of body stuff can we preserve an individual person” (*Mind Children*, 117). In other words, all too often the body determines – or for Moravec, overdetermines – the identity of the subject. He argues instead for “pattern-identity,” which “defines the essence of a person, say myself, as the *pattern* and the *process* going on in my head and body, not the machinery supporting that process. If the process is preserved, I am preserved. The rest is mere jelly” (*Mind Children*, 117). Moravec hopes to preserve and sustain the mind across generations and across human and machine species; however, he does not demand the absolute obliteration or transcendence of the body in the privileging of pattern and process over “stuff.” Rather, he is positing a reconceptualization of “the body” as dynamic and multiple, as pattern and process as much as fleshed materiality. In a later book, *Robot: Mere Machine to Transcendent Mind*, this dynamic and multiple body takes further shape, in the form of a brain in a vat.

Picture a “brain in a vat,” sustained by life-support machinery, connected by wonderful electronic links to a series of artificial rent-a-bodies in remote locations and to simulated bodies in virtual realities. (*Robot*, 169)

Both terrifying and thrilling, Moravec’s vision of a human mind separated from its body may initially appear to take up, literalize, and amplify a separation between the body and

³ I am a bit puzzled by the mortality that Moravec attributes to the mind. What of writing and other forms of creative and intellectual production? What kind of material transmission of the mental would be appropriate evidence for Moravec?

the mind. A closer reading of this brain in a vat scenario suggests that Moravec's discussion posits a much more complex conception of the body, what it means to have a body, and what it means to be embodied. Rather than arguing for and celebrating a disembodied, Cartesian model of the human, Moravec is in fact arguing for a conception of subjectivity, identity, and mind that is absolutely embodied.⁴

In *Volatile Bodies*, her important book on the epistemological history of the body, Elizabeth Grosz argues that there is no normatively sexed, raced, or physiognomic body: "there are only *bodies*" (19). Within this framework of specificities and multiples, embodiment can be understood as encompassing the body in all its material, ideational, and temporal iterations; and the body to which the subject is always already bound exists, not as fixed and singular, but rather as ever-shifting co-existing multiples. Moravec's brain in a vat is no less embodied than any and every other embodied entity. As Grosz and Moravec suggest, to have a body is to have multiple bodies. In other words, Moravec's brain in a vat does not have a body; it has *bodies*. The brain in a vat does not have arms, legs, and kidneys; however, it has "artificial rent-a-bodies" and "simulated bodies in virtual realities." The brain, and the cognitive functions it references in Moravec's imaginative thought experiment, is not "tied down" to just one body, or even one kind of body; rather, the brain can possess and be possessed by many different bodies

⁴ Coincidentally, after his death in 1871, the brain of Charles Babbage was preserved in alcohol. In 1909, an examination of Babbage's brain was conducted. For the details of this examination see Victor Horsley, "Description of the Brain of Mr. Charles Babbage, F.R.S." Babbage's brain, now in two halves, is currently housed in the Science Museum in London. The Museum loans out one half to other museums.

concurrently. These bodies may be flesh and blood, or machine circuits and tubes, while others still might exist as virtual and ideational.

Moravec organizes the human itself around this expanded body in all of its multiple materialities, immaterialities, and realities. Even outside of any claim to a material body, “humans need a sense of body” (*Robot*, 170). Within Moravec’s formulation, the body need not have any claim to materiality at all, as “[t]ransplanted human minds will often be without physical bodies, but hardly ever without the illusion of having them” (*Robot*, 170-171). Human bodies are remote, virtual, and ideational as much as blood and bone. Moravec goes on to distinguish between the transplanted human mind and what he calls “pure mind,” that is, a “truly bodiless min[d]” (*Robot*, 171). The pure mind, the mind with no body in substance, ideation, or “mere” sense, is not human. This pure mind is in fact AI: “the resultant bodiless mind that results, wonderful though it may be in its clarity of thought and breadth of understanding, would be hardly human. It will have become an AI” (*Robot*, 171). For Moravec, the human is embodied by definition. With a body, even if only as “illusion,” the mind is human. Without any body, the mind is no longer human, but AI.

Cyberpunk fiction explores this complicated nexus of questions, anxieties, and possibilities surrounding the brain in a vat – technology, embodiment, disembodiment, cognition, identity, subjectivity, materiality and ideality.⁵ In Rudy Rucker’s cyberpunk

⁵ Literary theorist Dani Cavallero locates the brain in a vat itself as a central theme of cyberpunk, similarly arguing that the brain in a vat speaks to an expansion of the body, rather than a Cartesian disembodiment. Cavallero also links the brain in a vat, as metaphor, to cyberpunk’s discussions of biotechnology and genetic engineering (“The Brain in a Vat in Cyberpunk: The Persistence of the Flesh,” 295).

novel, *Software* (1982), this nexus operates through the uploading and downloading of human brain patterns into memory banks, robots bodies, and even ice cream trucks. *Software*, similarly rejecting the brain in a vat as imaginary repository for Cartesian fears, expands the parameters by which we understand and delimit “the body.” The novel, admittedly with reservations, expands the body not just beyond the limits of the skin and perception and into the world, but outside any sense of an organic, original body.

Cyberpunk emerged in the U.S. in the 1980s. In 1983, a short story entitled “Cyberpunk,” written by Bruce Bethke, was published in the magazine *Amazing Science Fiction Stories*. In 1984, Gardner Dozois, a prominent science fiction editor, used the term “cyberpunks” in his *Washington Post* article, “Science Fiction in the Eighties” (Cavallero, *Cyberpunk*, 13).⁶ Rucker, unlike William Gibson, is not unanimously included in this literary movement.⁷ This lack of consensus can partially be understood by the multiple definitions and characterizations of cyberpunk.

For example, Booker and Thomas, describe cyberpunk, in its combination of punk and high technology, as a rejection of the “technological utopianism” that is found elsewhere in science fiction (Booker and Thomas, 110). Within this characterization, Booker and Thomas read Rucker’s *Software* (indeed, all the novels in Rucker’s *Ware* series) as “stylistically and philosophically” closer to the work of Philip K. Dick than to other cyberpunk novels, such as those of Gibson. Thematically, though, Rucker’s

⁶ In *Virtual Geographies*, Heuser points out that Dozois, rather than Bethke, whose short story is not well known, is often credited with originating the term “cyberpunk” (7).

⁷ Istvan Csiscery-Ronay holds up Gibson as “the one major writer who is original and gifted enough to make the whole movement seem original and gifted,” while describing Rucker’s inclusion in cyberpunk as “smack[ing] more of friendly endorsement than of truly shared aesthetic aims” (268).

discussions of “cyborgs, genetic engineering, artificial intelligence, drug use, and electronic transcendence are very much in line with typical cyberpunk concerns.” (Booker and Thomas, 114). Rucker, then, is aligned thematically with cyberpunk, but not stylistically and philosophically.⁸

Sabine Heuser, who foregrounds cyberspace as a central theme of cyberpunk, considers Rucker “only marginally cyberpunk” (Heuser, 229 n37). Like Booker and Thomas, Heuser also distinguishes between cyberpunk style and cyberpunk themes. Heuser characterizes cyberpunk primarily through its “style, mood, and atmosphere”; thematically, cyberpunk “is only implicitly motivated by science and the underlying models on which its worlds are built” (Heuser, 15). For Heuser, then, Rucker exists on the margins of cyberpunk because he privileges scientific themes over cyberpunk style; he is “more scientifically motivated than stylistically inspired” (Heuser, 15). In an inflection of Heuser’s critique, Brian McHale distinguishes Rucker from the rest of the cyberpunk movement on account of the humanoid robots in Rucker’s works. While the humanoid robot is a common figure in non-cyberpunk science fiction, McHale argues that it is not a significant component of cyberpunk.

In “Cybernetic Deconstructions: Cyberpunk and Postmodernism,” Valerie Hollinger, who does include Rucker within the cyberpunk movement, distinguishes cyberpunk from other forms of science fiction. She does so by distinguishing between

⁸ This distinction between thematics and philosophy seems a bit tenuous; for the purposes of this discussion, however, I am primarily concerned with how Rucker is located, if at all, within these different discussions of cyberpunk.

science fiction's anthropocentric perspective, and cyberpunk's distinctly postmodern human-machine hybridity:

If we think of SF as a genre which typically foregrounds human action *against* a background constituted by its technology, this blurring of once clearly defined boundaries makes cyberpunk a particularly relevant form of SF for the postindustrial present. We can read cyberpunk as an analysis of the postmodern *identification* of human and machines. (Hollinger, 205)

Locating cyberpunk within postmodern thought – indeed, as a “symptom of the postmodern condition” (Hollinger, 204) – Hollinger describes cyberpunk as operating to decenter the human and, in its cybernetic influences, working not to upend oppositions binary oppositions, but in fact to disrupt the delineation of these oppositions themselves.

Cavallero, like Hollinger, points to cybernetics as a critical means of distinguishing cyberpunk from the rest of science fiction (*Cyberpunk*, 14). Specifically, Cavallero points to cybernetics' understanding of the human body as machine as particularly influential for cyberpunk's “virtual interchangeability of human bodies and machines” (*Cyberpunk*, 12). Drawing together technology and mythology, cyberpunk radically reconfigures the human body precisely around its new fictional relationships with technologies and machines. Cavallero anticipates scientists' critique of cyberpunk as “abusing” scientific concepts, as drawing on the science without sufficient comprehension and out of context.⁹ In response, Cavallero suggests that cyberpunk's synthesis of scientific, technological, and imaginative fictional discourses can produce new forms of knowledge about these discourses, namely the similarities between them:

⁹ Cavallero specifically cites Alan Sokal and Jean Bricmont's *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*.

“[I]t could be argued that by bringing together diverse registers and discourses, mythological motifs and technological issues, cyberpunk is in a position to highlight the fact that scientific and narrative structures are analogous, for both continually move from one mythology to another” (*Cyberpunk*, 70-71).

Rucker, whom Cavallero includes in her discussion of cyberpunk, holds a doctorate in mathematics, is currently a professor of computer science, and has published and edited several books on mathematics and computer programming. Rucker, then, possesses both scientific and narrative expertise. Rucker, speaking to how these two discourses – science and science fiction – feed into and influence each other in his work, explains, “I don’t write SF to help my science. If anything, I study science to help my SF!” (“37 Questions,” 15). This influence is mutual, as numerous roboticists, including Rodney Brooks, note.¹⁰

While his inclusion in the cyberpunk movement is debatable among some critics, Rudy Rucker himself locates *Software* and *Wetware*, the first two novels in his *Ware* tetralogy, firmly within cyberpunk.¹¹ In his essay “What Is Cyberpunk?”, Rucker describes cyberpunk as having emerged between 1982 and 1986,¹² and as primarily “concerned with information.”

¹⁰ Further pointing to the inextricability of the fictional imaginary and the scientific reality of robotic creation, numerous roboticists point to their early fascination with fictional robots as their introductions into the field of robotics. See for example Brooks, *Flesh and Machines*; Minsky, “Our Roboticized Future”; A. F. Umar Khan, “The Ethics of Autonomous Learning Systems.”

¹¹ The *Ware* tetralogy is comprised of *Software* (1982), *Wetware* (1988), *Freeware* (1997), and *Realware* (2000). Both *Software* and *Wetware* won the Philip K. Dick Award for distinguished science fiction.

The essence of cyberpunk fiction, as I see it, is that it is concerned with information. The concern exists on several levels. On the objective level, a cyberpunk work will often talk about computers, software, chips, information, etc. And on the higher level [...] a cyberpunk work will try to reach a high level of information-theoretic complexity. (“What Is Cyberpunk?” 318-319)

Cyberpunk is *about* information technologies, in that these technologies appear in cyberpunk fictions. And cyberpunk is also formally *about* information, in that cyberpunk works are structured to pursue high information complexity. As familiarity, according to Rucker, bespeaks a decrease in information-content, cyberpunk’s concern with information complexity also renders it as engaged in defamiliarization (318).¹³ In a postscript to this article, Rucker expands on this defamiliarizing project by reframing cyberpunk’s “information complexity” as “high depth,” which he describes as the complete reinvention or making new, of previously familiar science fiction figures such as the robot (321). In the specific case of *Software*, it is just as much the robot as the human, largely through the human body, that is defamiliarized.

¹² In 1985, Rucker participated in a panel on “Cyberpunk” at the National Science Fiction Convention. Fellow panelists included science fiction writers Bruce Sterling, John Shirley, Lew Shiner, Pat Cadigan, and Greg Bear. William Gibson was asked to participate, but was unable to attend. To Rucker’s dismay, the panelists were met with hostility by both the moderator and the audience. “Here I’m finally asked to join a literary movement and everyone hates us before I can open my mouth?” (“What Is Cyberpunk?”, 315-6).

¹³ Rucker’s discussion of information complexity and familiarity speaks to Cavallero’s, which similarly points to cyberpunk’s defamiliarization and refiguration of the body.

2.2 Brains, Bodies, and Boppers

Rucker identifies two primary themes in *Software*. The first deals with intelligent and conscious robots that emerged through self-evolution and self-reproduction. The second deals with the removal of human intelligence and identity from the body, and the upload and download of this intelligence and identity into other bodies (“37 Questions,” 11). I will focus primarily on the latter theme, in that it articulates embodiment as necessarily comprised of multiple, specific, and dynamically shifting bodies. I will return briefly to the first theme in the final section of this chapter.

In *Software*, the capacities and implications of the extractable human mind is articulated through two couplings: one human and one robotic. The human dyad is comprised of Cobb Anderson, the now disgraced creator of the intelligent, conscious, and self-evolving bopper robots, and Sta-Hi Mooney, an aimless, twenty-something, pot-smoking cab driver. The robotic dyad includes Ralph Numbers, who led the 2001 revolt that liberated the robots from human control, and Wagstaff, who opposes the bopper plot to collect all human brain-tapes. These two pairings lay out a number of the varying positions, anxieties, and possibilities that are activated by the brain in a vat scenario. These discussions primarily take shape around two different interpretations of the extracted mind: does the boppers’ extraction of human brains from human bodies offer immortality, or destruction?

2.2.1 The Body in Multiples

The novel opens decades after Cobb programmed the boppers for self-evolution and consciousness. Having subsequently been tried for treason, Cobb is now retired in Florida, terrified of death, drinking heavily, and waiting for his replacement heart to give out. Awaiting death, Cobb is greeted by a stranger, who turns out to be his robotic copy. “I’m you,” the stranger tells Cobb. “You who?,” Cobb asks, confused.

“You me.” The stranger used Cobb’s own tight little smile on him. “I’m a mechanical copy of your body.”

The face seemed right and there was even the scar from the heart transplant. The only difference between them was how alert and healthy the copy looked. Call him Cobb Anderson₂. Cobb₂ didn’t drink. Cobb envied him. He hadn’t had a completely sober day since he had the operation and left his wife.

“How did you get here?” (*Software*, 5)

Cobb₂ identifies himself threefold: as Cobb (“I’m you”), as “mechanical copy of [Cobb’s] body,” and as himself (“me”). These three identifications are discrete; and yet all three co-exist simultaneously. And Cobb, unlike Sta-Hi, as we will soon see, is neither surprised nor disturbed by the encounter with his mechanical copy and its multiple claims to his identity, his personhood. Rather, Cobb is envious of Cobb₂, of the things that differentiate his alert, healthy, sober copy from himself. This copy is both him and not him; in that Cobb₂ is not him, Cobb₂ is better.

On behalf of the boppers, Cobb₂ traveled from the Moon to offer Cobb immortality. For the boppers, this offer of immortality comes in the form of a new body: “we can rebuild you from the ground up” (6). The immortality promised by the boppers necessitates having a body, but the specificity and materiality of this body is not narrowly construed, nor is the connection between identity and a single specific body an inviolate

one. Bopper immortality does not depend on maintaining the same body *ad infinitum*. In fact, for the boppers, survival and persistence requires that they change bodies every ten months. This is the immortality the boppers offer Cobb – the immortality that accompanies a new body. And Cobb welcomes the boppers’ offer to extract his brain tape and download it into a brand-new, healthy, *mechanical* body.

Drunk and drugged, Sta-Hi Mooney passes out on the beach; he wakes up unable to feel his body. “He opened his eyes. His body seemed to have disappeared. He was just a head resting on a round red table” (*Software*, 30). As Sta-Hi gradually regains consciousness, he finds his body, hidden under the table. “[He] had a body after all. It was just that his body was tied up under the table and his head was sticking out through a hole in the table-top” (*Software*, 30-31). Sta-Hi’s predicament speaks to the more horrifying aspects of the brain in a vat scenario, as Sta-Hi, unlike Cobb, finds himself effectively envatted against his will, kidnapped by a gang called the Little Kidders. Armed with spoons, the Little Kidders gather around Sta-Hi to eat his brain.

Sta-Hi, who later rejects the boppers’ vision of immortality, is bound to his body in a way that Cobb is not; for Sta-Hi, his body, this specific body that is tied up and entangled, has a proprietary relationship to his identity and his mind. Indeed, it is only when the Little Kidders cut into Sta-Hi’s flesh, his body, that his mind is jolted into action, looking for a way to escape. “The sudden pain over his eyebrow had brought Sta-Hi back to rationality” (*Software*, 33). This pain, in bringing Sta-Hi back to his body, brings him “back to rationality.” In other words, Sta-Hi’s rational mind is not activated until he is returned to his body, first in its tactile and sensory reappearance to himself as

he regains consciousness, and then through physical pain. In his renewed mental acuity, Sta-Hi resists his attackers, loudly and violently thrashing against the table until the police arrive.

On the Moon, plugged into a meta-bopper called the One, boppers Ralph Numbers and Wagstaff argue over the immortality offered to Cobb; does this immortality in fact result in extended life, or in death? Ralph is adamant that the powerful big boppers are offering Cobb immortality and liberation, while Wagstaff is equally adamant that the big boppers are actually offering Cobb his own death. Wagstaff wants to protect Cobb, a human hero to the boppers, from the big boppers' plan and their false promise of immortality: "I'mm goinng to do everrything I can to sstopp you fromm turrrning that poorr olld mman innto a piece of ssoftware in the bigg boppers memory bannks. Thatts nnot immortality" (*Software*, 23). Wagstaff, akin to Sta-Hi, argues for the importance of the body, asserting its intractable nonobsolescence in the face of threats of disembodiment. However, Wagstaff and Sta-Hi's understanding of embodiment risks boxing itself into its own claim to the body's singularity.

Wagstaff and Ralph's disagreement can be understood in relation to the question of what constitutes Cobb, of what counts. Is it Cobb's software, his pattern, or is it more than that? Ralph questions why the materiality of Cobb's body makes a difference. Why should this work any differently for Cobb than it does for the boppers? "When we're through with [Cobb] he'll be immortal. What's so important about having a carbon-based body and brain? [...] The pattern is all that counts" (Rucker, 20-21). Everything else, as Moravec says, "is mere jelly." For Ralph, humans and machines are in essence

no different; their identity, their essences persist in their patterns. “What was Wagstaff so upset about anyway? Everything would be preserved... Cobb Anderson’s personality, his memories, his style of thought. What else was there? Wouldn’t Anderson himself agree, even if he knew? Preserving your software... that was all that really counted!” (*Software*, 24). While “preserving your software” is what counts, the software or the pattern does not transcend bodies. Preserving the software preserves memories and identity, so long as one has a body into which the software can be downloaded. While the question of *which* body recedes in significance, that there is a body has never mattered more in terms of memory, that is, a conception of identity and experience in duration.

Before meeting their mechanical copies, both Cobb and Sta-Hi first encounter the effects of their mechanical copies on their worlds. Detached from Cobb and Sta-Hi, these copies move around the world, creating memories that Cobb and Sta-Hi access as effects rather than experiences.¹⁴ Upon his return from the moon, Cobb integrates all of Cobb₂’s memories, his experiences; Cobb rejoins his life where Cobb₂ left off, doubting nothing of what Cobb₂ changed.¹⁵ Sta-Hi, never accessing nor integrating Sta-Hi₂’s memories, has a more fraught and combative relationship with his copy. At the moon station, Sta-Hi₂, going by the alias Mr. DeMentis, had earlier secured visas and shuttle tickets for Sta-Hi and Cobb. In Mr. DeMentis, Sta-Hi₂ imposes an entirely new persona onto Sta-Hi.

¹⁴ Cobb, before meeting Cobb₂, is perplexed to find that himself a subject of interest in Mooney’s investigation into the warehouse theft of bopper-grown kidneys. Similarly, at the space station, Sta-Hi similarly must deal with the repercussions of his copy’s earlier harassment of a woman. In both of these scenarios of mistaken identity, we can see more Cobb and Sta-Hi who are in the dark, mistaken about identity, rather than Mooney and the harassed woman.

¹⁵ While Cobb was on the Moon, Cobb₂ began a romance with a neighbor, Annie. When Cobb returned a week later, Cobb₂ and Annie were living together.

Sta-Hi, formerly Stanley Hilary Mooney, Jr., and Stanny to his father, is not unfamiliar with the struggle for identity: “Look, Stanny...,” Cobb says. “‘STAY HIGH,’ he bawled. ‘GET IT RIGHT!’” (*Software*, 42). While Cobb and Sta-Hi are on the Moon, their copies take over their lives in Florida. Sta-Hi₂ is responsible and, unlike Sta-Hi, not directionless. An improved son for Mooney, Sta-Hi₂ becomes Stanny. Sta-Hi₂ imposes Mr. DeMentis, and later Stanny, onto Sta-Hi. Along with these separate aliases and identities are separate memories that are never integrated into Sta-Hi, as Cobb₂’s memories are into Cobb. Sta-Hi’s multiple identities remain discrete. Both bodies – Sta-Hi’s and Sta-Hi₂’s – house multiple identities that do not co-exist, but rather threaten each other with mutual exclusivity.

These two iterations of multiple memories and bodies – Cobb’s multiple bodies as extension and expansion and Sta-Hi’s multiple bodies as threatening imposition – suggest that for multiple bodies to amplify and augment, a different conception of the software, of identity is required. This reconceptualization is not fraught with anxieties of Cartesian obsolescence, but rather operates on an entirely different order, one that is not organized around dualism between mind and matter, but instead around a much more complicated dynamic between multiple bodies and, at times, multiple minds.

Sta-Hi, like Wagstaff, does not view the boppers’ offer as promising immortality, but rather replacement and subsequently, destruction. Sta-Hi expresses this concern to Cobb:

“Did you ever flash,” he asked through a cloud of exquisitely detailed smoke, “that maybe those copies of us could be *permanent*? That this is all just to get us out of the way so Anderson₂ and Sta-Hi₂ can pose as humans?” This was, at least

in Sta-Hi's case, a fairly correct assessment of the situation. But Cobb chose not to tell Sta-Hi this." (*Software*, 56)

Not only does Sta-Hi articulate this alternative interpretation, but, as Cobb understands, for Sta-Hi, the extraction of his brain-tape as destruction *is* a legitimate fear. The boppers' offer only extends to Cobb, as they never intended to download Sta-Hi's brain-tape into another body. Mirroring his life-threatening encounter with the Little Kidders', brain-tape extraction for Sta-Hi does not result in immortality, as it might for Cobb, but rather in the absence of a body, thus in immediate mortality.

2.2.2 The Matter and the Memory

In *Software*, the body's participation in the persistence of memory and identity is articulated through both the ephemerality of bodies and the temporal liminalities that emerge as the software changes bodies. While in between bodies, humans and boppers alike do not make memories. Absent a body, there is no memory. Upon leaving his meeting with Wagstaff, Ralph Numbers is attacked by digger robots that are sympathetic to Wagstaff's position. Hit with a laser, Ralph's machinery gives out and effectively dies. He wakes up a few hours later, acclimating to a new body. Ralph has no memory of the attack. "I guess I'm the new Ralph Numbers scion?" There was no memory of a tenth visit to the One, no memory of disassembly... but there never was" (*Software*, 47). Memory loss accompanies every disassembly and subsequent installation of a new body. This time in between bodies is a time without memory.

In a violent evisceration witnessed by a hidden Sta-Hi, Cobb's software is extracted and his organs removed (*Software*, 96-97). In his new mechanical body, Cobb wakes up back in Florida, buying ice cream from a Mr. Frostee ice cream truck. It takes Cobb a little while to become accustomed to his new body, as well as to process and integrate Cobb₂'s memories into his own. Cobb possesses Cobb₂'s memories, not just for the week that Cobb was on the Moon, but for the entirety of Cobb₂'s existence. As Cobb₂'s memories are integrated into Cobb, Cobb₂'s body becomes an extended aspect of Cobb's embodiment. Initially, all of this integration and extension – of memory, of body – deeply disorients Cobb. In an effort to center himself during this memory integration, Cobb mutters, “I am [...] I am me.’ He . . . this body . . . hadn’t thought that for . . . how long?” (*Software*, 101). In centering himself in this “I am,” Cobb echoes Ralph Numbers who, after he was attacked, lay on the ground as his system was shutting down, “clutch[ing] at the elusive moth of his consciousness. *I am. I am me*” (*Software*, 28). Ralph calls this consciousness “a memory of who he was... the self symbol. He was a big metal box resting on caterpillar treads, a box with five arms and a sensory head on a long and flexible neck. He was Ralph Numbers, who had set the boppers free” (*Software*, 28). For Ralph and Cobb, identity and the self are grounded in memory. And this memory, which grounds identity and self, only exists in that it is embodied.

Sta-Hi, having evaded the boppers and sneaked back to earth, outs Cobb as a robot. When Mooney goes to capture Cobb, to “deprogra[m] and dismantl[e]” him, Cobb activates the SELF-DESTRUCT subroutine programmed into his new body. “‘DESTROY,’ Cobb said, and lost his second body” (*Software*, 139). The explosion also

kills Mooney, who, like his son, only has one body. Cobb wakes up in yet another new body.

Sta-Hi, aimless again, re-encounters Cobb, who has changed bodies once more. Cobb is now Mel Nast, leader of the Personetics cult, a scam by which Cobb and Mr. Frostee can collect human brain-tapes. Sta-Hi, upon realizing that the Mr. Frostee truck is in fact a bopper, crashes his cab into Mr. Frostee, destroying the big bopper and the Mel Nast robot-remote. As for Cobb and his immortality, Cobb, now robot-remoteless, is absorbed into the damaged Mr. Frostee truck; always with a sense of body. Before the Mr. Frostee truck dies out completely, Cobb asks, “‘Am I on tape somewhere else?’ [...] ‘Is there a copy on the Moon?’ ‘I don’t know,’ Mr. Frostee said. ‘What’s the difference?’” (*Software*, 165). For Cobb, the difference lies in the initial promise of immortality. If the process is preserved, Cobb can persist, can survive. No persistence without the pattern, and no memory or identity without a body for this pattern.¹⁶ Sta-Hi₂, unlike Cobb₂ who is integrated into Cobb, is untethered to a pattern, an identity. Sta-Hi₂ is a robot remote to nothing, and cannot persist as such. The novel, with its shifting multiple bodies and robot-remotes, both posits and problematizes an embodiment that is expanded and distributed across multiple bodies and materialities. This conception of embodiment crucially grounds memory, identity, mentation, and indeed, consciousness.

¹⁶ This body without a pattern makes an appearance in the form of Sta-Hi₂, who is shut down by Mr. Frostee upon Sta-Hi’s escape back to earth. Sta-Hi₂ quickly deteriorates, both in functional capacity and in materiality, “[its] features blurred and melted, the nose flopped over to one side and sagging down the cheek, the folded hands puddled like mittens” (*Software*, 121).

2.3 Pure Body and Becoming-Robot

In the Prologue to *Flesh and Machines: How Robots Will Change Us*, Brooks articulates a familiar anxiety about technology: will robots take over the world from humans? For Brooks, humans and robots exist on two ends of a spectrum whose poles are moving continually closer as technology evolves; in other words, we are becoming more robotic. Closer examination of the language of these becomings, and the sites in which they take hold, brings out an interesting, indeed Cartesian, tension in Brooks' discussion of the human:

I have devoted my life to building intelligent robots, and these robots are just now starting to emerge from labs out into the real world. As these robots get smarter, some people worry about what will happen when they get really smart. Will they decide that we humans are useless and stupid and take over the world from us? I have recently come to realize that this will never happen. Because there won't be any us (people) for them (pure robots) to take over from. Barring an asteroid-sized thwack that would knock humans back into pretechnological society, humankind has embarked on an irreversible journey of technological manipulation of our bodies (*Flesh*, ix).

In this frightening yet familiar scenario, robotic intelligence surpasses that of “useless and stupid” humans. Brooks suggests that humans are becoming more “robotic” by the “technological manipulation of our bodies.” Because of technology's ever-encroaching presence in and on our bodies – for example, he mentions cochlear implants, laser eye surgery, and genetic therapies – we are, according to Brooks, less human. Thus, the question of robots taking over the world from humans is a non-question, because as humans' *bodies* become more technological, we become less human and more robot.

For Brooks, this becoming-robot is not a frightening, technologically dystopic view of the future (indeed, of our present), but rather a future that he welcomes.

Bracketing Brooks' technological optimism for the moment, I wish to focus on his discussion of the body as it delineates the human. For Brooks, it is this idea of an organic, "original," pretechnological body that determines the human, and that quickly and irretrievably disappears in the presence of technology. Thus the human can become robot, but the robot cannot become human. As robots become smarter, they remain robots – they do not become more human. Through the body, the human can become robotic-other; but this other, however, cannot become human.¹⁷

Both Brooks and Moravec posit the body as central to the human.¹⁸ However, they diverge in their different definitions of the body. Moravec, likewise Rucker in *Software*, holds the body to significantly less rigid parameters than does Brooks. For Moravec, the body does not exist as originally pure and singular. This impure, multiple body is the body of the human. And the non-bodied is the non-human – pure mind. Brooks, according to this "Prologue," might be said to define the human according to a notion of "pure body," a body that is impossibly pristine, self-contained, and untouched by technology. And, to return to Brooks' optimism for our future in technology, it is this pure body, this human, that Brooks is quite content to leave behind. For, just as the robot

¹⁷ Tim Lenoir points to Brooks' embodied and situated robotics, as well as his discussion of humans as machines, as reminding us that "our notion of the body is a cultural construct, a historical conception both contested and negotiated. [...] it is an interpretive frame we coconstruct along with our machines and the worlds they inhabit." ("Makeover: Writing the Body into the Posthuman Technoscape: Part One: Embracing the Posthuman," 210). Lenoir suggests that remembering that the body is a site of contestation is an important intervention; for, understanding the body as such provides an opening for our own agential participation in the multiple negotiations of the body, as well as in the related contestations of the human and the posthuman.

¹⁸ Brooks and Moravec were both at the Stanford Artificial Intelligence Laboratory (SAIL) in the late 1970s. Brooks, who during this time helped Moravec with some of his robotic projects, considers Moravec "a tremendous influence on my life, once I got to Stanford and met him" (*Flesh*, 27).

cannot become human, neither can the human. In moving beyond the pure body and its proprietary relation to the human, embodiment itself becomes vastly unbounded, opening up onto new materialities, imaginaries, and relationalities.

In the late 1960s, as a teenager, I saw *2001: A Space Odyssey*, and it was a revelation. Like millions of others, I was enthralled by the soft-spoken computer villain HAL 9000 and wondered if we could one day get to that level of artificial intelligence. Today I believe the answer is yes. Nevertheless, in hindsight, I believe that HAL was missing a fundamental component: a body. (“I, Rodney Brooks,” 73)

HAL, of course, does have a material structure. However, by Brooks’ configuration, HAL is not embodied, as for Brooks embodiment indicates a dynamic relationality with the world. He characterizes robotic embodiment as the following: “The robots have bodies and experience the world directly – their actions are part of a dynamic with the world and have immediate feedback on their own sensations” (“Intelligence without Reason,” 138). Primarily concerned that robots exist with us in our world, Brooks asserts that robots cannot be in the world if they are not embodied. “[O]nly an embodied intelligent agent is fully validated as one that can deal with the real world” (“Intelligence without Reason,” 167). Brooks, in designing intelligent robots to successfully navigate and operate in the world, designs robots that are embodied, recasting intelligence itself in relation to this embodied and worlded existence.

To more fully comprehend Brooks’ recalibration of robotic intelligence around embodiment, I will first discuss Brooks’ critique of Shakey, one of early AI’s robotic

achievements.¹⁹ Built by a team led by Nils J. Nilsson, Shakey was created over six years (1966-1972) at the Stanford Research Institute. Shakey is a mobile robot that can sense and locate objects, as well as navigate within a dynamic environment. Its robotic system, consisting of five levels, is organized hierarchically (Nilsson, 5-7). The bottom level contains the radio and microwave communication channels that connect Shakey's mobile and sensing body to its computer, which, during the six years, was housed in a separate room. The second level is comprised of Low-Level Actions (LLAs). LLAs control the motion of Shakey's head, and thus the camera attached to its head (for example, commands such as TILT and PAN). LLAs also control general aspects of movement (for example, ROLL and TURN). Certain LLAs activate the television camera that comprises Shakey's sensory system (for example, SHOOT, which takes and saves a picture) (29-30). The third level in the hierarchy consists of Intermediate-Level Actions (ILAs), such as GOTHRUUDR (go through door) and GOTOROOM, which are preprogrammed combinations of LLA commands. The fourth level of Shakey's system involves the STRIPS planning system, "which constructs sequences of ILAs needed to carry out specified tasks." The fifth and highest level of the hierarchy, PLANEX, is responsible for executing and monitoring the ILA sequences generated by the STRIPS

¹⁹ Of this earlier approach to robotics, Shakey, according to Brooks is "the most celebrated such robot" (*Flesh*, 22). In 2004, Shakey was instituted into the Robot Hall of Fame (RHF). The RHF, which is affiliated with Carnegie Mellon University, one of the major robotics research centers in the U.S., honors robots in two categories: "Robots in Science" and "Robots in Science Fiction." The RHF, in honoring fictional robots, notes the intricate relationship between robots in science and robots in the cultural imaginary ("These are fictional robots that have inspired us to create real robots that are productive, helpful, and entertaining. These robots have achieved worldwide fame as fictional characters and have helped form our opinions about the functions and values of real robots." Shakey's fellow inductees in 2004 include Honda's humanoid ASIMO robot, Tetsuwan Atomu's cute and loveable Astro Boy, and C-3PO of *Star Wars* fame ("Robot Hall of Fame").

system (6). Shakey's world began as a single room and was later expanded into multiple rooms and a hallway. The room was tiled with an unpatterned, non-reflective surface.²⁰ The objects with which Shakey interacts were painted red, grey, or white.

Brooks critiques Shakey's hierarchical organization as producing a robotic existence that is not worlded, neither temporally nor spatially. I quote the below passage at length, as it lays out Brooks' multiply-pronged, intertwined critique of Shakey's top-down approach to intelligence.

For Shakey, however, the researchers made a complete two-dimensional map of the world before Shakey was started up, and stored it in the computer's memory. Implicit in Shakey's programming was the assumption that the floor would be perfectly flat, and that nothing could sit on top of anything else, so two dimensions were sufficient. [...] It clearly had no sense of the here and now of the world – if someone came and moved things around while it was “thinking,” it would eventually start up again, acting as though the world was still in the same state it had been before the perturbation. Shakey used reasoning in situations where real animals have direct links from perception to action. It was designed on the premise that its perceptual computations could maintain an accurate model of the world, but that was so technically difficult that its designers were forced into the deceit of making the world very simple so that this was possible. (*Flesh*, 23)

Brooks' critique of Shakey is threefold. Firstly, Shakey does not operate in the dynamic world of humans and animals, but rather in a limited, constructed, two-dimensional representation of the world. Shakey's world, with its two-dimensional surfaces and its restricted color range, is a significant distillation of the uncontrollability and spontaneity proper to the real world. According to Brooks, for Shakey to successfully navigate and move in a quasi-dynamic world, “its designers were forced into the deceit of making the

²⁰ The expansion of Shakey's world into multiple rooms was largely enabled by replacing the cord that connected Shakey to its computer, with radio-remote connection (Nilsson, 12).

world very simple” (*Flesh*, 23). Nilsson describes Shakey’s world as “real, but contrived” (Nilsson, 10). Brooks, it seems, understands this contrivance, this “deceit,” as an unacceptable reduction and distortion of the world. Secondly, Shakey, in its limited world, is also asynchronous with what we might understand as real time. While Shakey can register dynamism, it can only do so with significant temporal lag. For Brooks, then, Shakey is unsituated twice-over, removed from both “the here and now of the world.” Finally, Brooks locates both of these critiques in Shakey’s hierarchical computational organization, in which older AI models of “reasoning” are privileged. This appeal to reasoning works to the disadvantage of Shakey’s ability to move in the world, to move more directly from perception to action in a temporality more familiar to humans.

In his 1991 paper, “Intelligence without Representation,” Brooks writes of creating robotic agents that both exist in the human world, and are viewed as intelligent by humans. Brooks makes clear that these robotic agents, however, are not humanoid, marking this distinction by calling his robots Creatures (“Intelligence without Representation,” 86). And yet, as I will discuss below, Brooks’ critique of Shakey relies on a certain anthropocentric conception and experience of the world: for Brooks, “real time” and “the real world” are in fact the time and world of human experience. And for robots to interact with humans, they must enter real (human) time and the real (human) world.²¹

²¹ While there is no question that Shakey’s is a highly constructed world, and its time scale might deviate from certain human time scales, we might ask why we insist that Shakey enter our world, rather than entertaining the possibility that we might enter Shakey’s world and exist within its time scale; or, better yet,

According to Brooks, between 1984 and 1991, “a new approach to robotics” emerged, one which addressed his primary critiques of Shakey’s robotic organization.

Driven by a dissatisfaction with the performance of robots in dealing with the real world, and concerned that the complexity of run-time modeling of the world was getting out of hand, a number of people somewhat independently began around 1984 rethinking the general problem of organizing intelligence. It seemed a reasonable requirement that intelligence be reactive to dynamic aspects of the environment, that a mobile robot operate on time scales similar to those of animals and humans, and that intelligence be able to generate robust behavior in the face of uncertain sensors, an unpredictable environment, and a changing world.²²

Brooks’ contribution to this “new approach” takes shape around subsumption architecture, “a new computation model ... [whose] purpose is to program intelligent, situated, embodied agents” (“Intelligence without Reason,” 172). Subsumption architecture is not a hierarchically-organized system; rather, it is comprised of multiple computational modules, each tasked to perform a specific behavior. These multiple behaviors – for example, avoid objects, wander, and build maps – run parallel to each other, and are each connected directly to perception and subsequent motor actuation. So, for example, in subsumption architecture, behaviors such as avoid objects, wander, and build maps co-exist as semi-autonomous computational processes and operations, each having partial access to the world, and thus producing a partial representation of the

that we might understand what new forms of interactivity and relationality might emerge from human interactivity with Shakey in his world and temporality.

²² “New Approaches,” 1227. In addition to his own work during this time period, Brooks also mentions MIT scientists Phil Agre and David E. Chapman, who developed Pengi, an AI that could play a video game. Brooks also points to SRI researchers Stanley J. Rosenschein and Lesley Pack Kaelbling, whose Flakey robot was able to operate both in real-time, and outside the environment of Shakey’s “real, but contrived” rooms. Brooks discusses these two projects at greater length in “New Approaches,” 1228-1229.

world; while these processes can share data amongst each other, there is no centralized, totalizing representation (“Intelligence without Reason,” 173).²³

2.3.1 Embodiment, Situatedness, and Emergent Intelligence

Brooks’ critique of Shakey, and more generally of the previous model of AI, and his various technological responses, from his subsumption architecture to his turn to embodiment, suggest the deeply intertwined relationship between body and world.²⁴ Brooks’ concern, as we saw in his critique of Shakey, is that the robot be in the world. And indeed, this is the purchase of his subsumption architecture, that “it enables us to tightly connect perception to action, embedding robots concretely in the world” (“Elephants,” 116). Brooks describes being “embedded in the world” as the state of being situated. “A *situated* creature or robot is one that is embedded in the world, and

²³ In 1984, Brooks built his first subsumption architecture robot, Allen. Named after Allen Newell, an important figure in AI, Allen is a cylindrical robot “about the size of R2D2” (*Flesh*, 35). Allen is equipped with sonar sensors and has three layers of behavior; the first allows Allen to avoid obstacles in its way, the second allows Allen to wander at random, and the third allows Allen to identify a desired location and move towards it. For detailed discussions of Allen’s subsumption architecture, as well as the interaction between these various behavioral layers, see Brooks, “Elephants Don’t Play Chess,” 18-20, Brooks, *Flesh and Machines*, 40-41, and Brooks, “A Robust Layered Control System for a Mobile Robot.” For a discussion of Toto, his former student Maja Mataric’s robotic project organized around decentralized representations and navigation, see “Intelligence without Reason,” 177. Meta-Toto, Lynn Andrea Stein’s expansion of Mataric’s Toto, operates through *imagined* decentralized representations. The introduction of imagination allows Meta-Toto to exist in the “now” of the world: “Imagination is ephemeral. MetaToto need only know the sensations that occur *now*. Where Toto ‘continually redecides what to do,’ MetaToto continually re-imagines the world. Thus, while world models persist and require maintenance, imagination can be reconstructed on the fly” (Stein, 5).

²⁴ The similarities between Brooks’ embodied, situated, and worlded approach and philosophical discourses, such as phenomenology, are striking. Merleau-Ponty in particular resonates with Brooks’ robotics: “To be a body, is to be tied to a certain world” (*Phenomenology of Perception*, 148). Brooks acknowledges that aspects of his theory are compatible with philosophical discourses; however, he clarifies that his innovations were not shaped by philosophical interventions, but rather from the technological developments: “in all cases the technological implementations came first, and the philosophical realizations followed later” (Brooks, “Preface” vii).

which does not deal with abstract descriptions, but through its sensors with the here and now of the world, which directly influences the behavior of the creature” (*Flesh*, 51-52). For Brooks, situatedness, or “connectedness to the world” works in conjunction with embodiment, which he describes through interaction with and being effected by the world: “An *embodied* creature or robot is one that has a physical body and experiences the world, at least in part, directly through the influence of the world on that body” (“Intelligence without Reason,” 150; *Flesh*, 52). The robot is situated in the world precisely because it is embodied.

For Brooks, designing a robot to successfully navigate and operate in the world, is designing a robot that is embodied. In other words, in that the problem, for Brooks, is that of unworldedness, or, in his vocabulary, unsituatedness, embodiment is the corrective. Brooks’ embodiment is a decidedly anthropocentric corrective, as the world in which Brooks situates his embodied robotic Creatures is a distinctly human world. In a paper co-authored with Lynn Andrea Stein, Brooks discusses humanoid robots as similarly embodied: “intelligence cannot be separated from the subjective experience of a body” (Brooks and Stein, 14). Brooks, in writing about humanoid robots, further connects a distinctly human embodiment to our human environments: “Our environments were designed and built for our bodies, so it will be natural to have these human-shaped robots around to perform chores like taking out the garbage, cleaning the bathtub, and carrying groceries” (“I, Rodney,” 74). In other words, because we have, for the most part, built our environments for our human bodies, task-oriented robots’ bodies should be designed to resemble our bodies. This anthropomorphization through embodiment does

not influence in only one direction, namely, from human to robot. Having earlier cleared the field of the pure, a-technological body, Brooks attends to the human, in its engagement with technology, as becoming-robot. In this way, Brooks' anthropomorphic robotic embodiment doubles back on the human, bringing with it new conceptualizations, robotic and otherwise, of the body, embodiment, and the world.

Brooks' insistence on embodiment does not simply invert the dualistic privileging of mind over body; his intervention is more complex. Brooks distributes cognition across various behavioral levels, as well as throughout the robotic body. In so doing, he eliminates the cognition box, the centralization of computational processes that mediate between perception and motor actuation.

Informally, at least, people thought of this box as the *cognition* box, the heart of thinking and intelligence. The best way to build this box, I decided, was to eliminate it. No cognition. Just sensing and action. That is all I would build, and completely leave out what traditionally was thought of as the *intelligence* of an artificial intelligence. (*Flesh*, 36)

Of course, Brooks is not proposing to eradicate robotic intelligence; he is in fact relocating intelligence in and throughout the robotic body, and redefining it as emergent. "[A]n emergent property of certain complex systems," intelligence is grounded in the robot's ability to inhabit the same temporality and dynamic world of humans and animals ("Intelligence without Reason," 185).

Moving away from earlier top-down approaches to intelligence, Brooks recasts intelligence as an emergent and distributed, rather than an originary and localizable, quality.

There is no homunculus. Rather, intelligence emerges from the interaction of the components of the system. [...] intelligence is produced by the interactions of

many components. Intelligence can *only* be determined by the total behavior of the system and how that behavior *appears* in relation to the environment. (“Intelligence without Reason,” 169-170 , my emphasis)

Brooks’ emergent conception of intelligence is intertwined with the crucial relationship between the body and the environment. Working backward from intelligence, if the goal is for robots to intelligently exist in our environment, our world, they, like us, must be embodied. Brooks’ “new approach to robotics” intervenes into earlier dualistic, disembodied approaches to intelligence precisely by foregrounding the body as the means of participation in and interaction with the world. The body, in other words, is the condition of access to the world; and from this interplay between body and world emerges intelligence.

Brooks describes emergent behaviors as produced when “multiple elementary behaviors couple with the environment to produce behaviors that are more than simple strings of suppositions of the elementary behaviors” (*Flesh*, 21). In *Flesh and Machines*, Brooks points to the production of emergent behavior in Grey Walter’s robotic mobile tortoises. Created in the 1940s, Walter’s tortoises were relatively simple in construction – they were comprised of two sensors (one to detect light and one to detect bumps) and two motors. Once the tortoises were let loose in various environments, unpredictable behaviors emerged. For example, Brooks describes what emerged when a tortoise was placed in front of a mirror. The tortoise, whose motor was connected to a light bulb, was attracted to the reflection of its own light source; this attraction, having drawn the tortoise to the light reflected in the mirror, shut off the motor and subsequently the lightbulb, effectively ending the attraction, which then turns the light back on, and so on (*Flesh*, 17-

21). This clunky, stop and start behavior was neither programmed nor predicted; it emerged only through the tortoise's interaction with its environment. Andrew Pickering, in a discussion of cybernetics as a "science of emergence," also locates emergent behavior in the interaction between Walter's tortoises and their environment, or world. "The tortoises," Pickering writes, "thus *thematized emergence* in that their world could always surprise them, and their behaviour was always an emergent and decentred joint product, of their trajectories in the world and what they found there; importantly, it was not something programmed into them in advance."²⁵

Within the relationship between robotic agent and world, emergence is primarily at stake for Brooks. For, emergence allows Brooks to put forth intelligence as a quality that takes shape epi-phenomenally to robotic behaviors and their interactions with the world. Brooks puts forth the following description of emergence, "Emergence: The intelligence of the system emerges from the system's interactions with the world and from sometimes indirect interactions between its components – it is sometimes hard to point to one event or place within the system and say that is why some external action was manifested" ("Intelligence without Reason," 139). Not only does Brooks' robotics recalibrate intelligence as an emergent, rather than *a priori*, quality, but it also renders intelligence unlocalizable. There is no longer a cognition box in which "intelligence" is isolated. Rather, Brooks proposes an emergent robotic intelligence that is delocalized

²⁵ Pickering, "Emergence," 130. In "Emergence and Synthesis: Science Studies, Cybernetics and Antidisciplinarity," Pickering reads Walter's tortoises alongside the projects of British cyberneticians Stafford Beer, whose Project Cybersyn operated on the Chilean economy; and Gordon Pask, whose Musicolor machine and Fun Palace, were explicitly designed to facilitate emergent behavior.

and distributed both throughout the robotic agent, as well as into and throughout the world.

Brooks further extends intelligence outside of the robotic agent by locating intelligence in the sphere of the observer. “The key idea from emergence is: *Intelligence is in the eye of the observer*” (“Intelligence without Reason,” 170). Intelligence is a quality that emerges not simply through robotic behavior in relation to the world, but also through the observer’s perception of this behavior. Intelligent robotic behavior, then, lies in the interpretative attribution of these behaviors by observers. If intelligence is largely attribution and projection, intelligence, then, can thus be reattributed and reconceptualized.

Alongside intelligence, Brooks also posits consciousness as emergent through embodiment and situatedness, of “being in the world.”

My feeling is that thought and consciousness are epi-phenomena of the process of being in the world. As the complexity of the world increases, and the complexity of processing to deal with that world rises, we will see the same evidence of thought and consciousness in our systems as we see in people other than ourselves now. Thought and consciousness will not be programmed in. They will emerge. (“Intelligence without Reason,” 184-185)

Brooks, like Rucker and Moravec, puts forth a conception of embodiment that expands well beyond any notion of an ordinary body. For Brooks, as in Rucker’s novel, it is precisely this appeal to worlded embodiment that opens up the possibility for both machine intelligence and machine consciousness.

2.4 Brain in a Vat Redux: A Coda on Consciousness

Neuroscientists Christof Koch and Giulio Tononi open their paper, “Can Machines Be Conscious?” with a disembodied, brain in a vat scenario:

Would you sell your soul on eBay? Right now, of course, you can’t. But in some quarters it is taken for granted that within a generation, human beings – including you, if you can hang on for another 30 years or so – will have an alternative to death: being a ghost in a machine. You’ll be able to upload your mind – your thoughts, memories, and personality – to a computer. And once you’ve reduced your consciousness to patterns of electrons, others will be able to copy it, edit it, sell it, or pirate it. It might be bundled with other electronic minds. And, of course, it could be deleted.

That’s quite a scenario, considering that at the moment, nobody really knows exactly what consciousness is.²⁶

Koch and Tononi tackle the difficult question of defining consciousness by listing what consciousness does *not* require. The list reads, “emotions, memory, self-reflection, language, sensing the world, and acting in it” (Koch and Tononi, 56). Koch and Tononi draw on mathematics, information theory, and complexity theory to suggest that consciousness can be understood as the capacity to generate integrated information. Within this framework, Koch and Tononi characterize consciousness as indicating “the availability of a large repertoire of states belonging to a single integrated system. To be useful, those internal states should also be highly informative about the world” (Koch and Tononi, 28). To test this capacity, they propose “a Turing Test for consciousness.” Provided an image – their photographic example depicts two men in a store, one holding a rifle to the other’s chest – the conscious machine would be able to provide information about what might be happening in the image, about things and people in the image, about

²⁶ Koch and Tononi. “Can Machines Be Conscious?” *IEEE Spectrum* (2008): 55.

things and people not pictured in the image (such as a getaway car waiting outside of the store), and about spatial and causal relationships depicted in the image. While Koch and Tononi seem to go out of their way to suggest that consciousness does not depend on the environment, their basis for machine consciousness relies on a decidedly situated knowledge of the world, on producing knowledge states that are “highly informative about the world,” as the responses they require for their Turing test demonstrate.

Echoing Descartes’ discussion of dream states,²⁷ Koch and Tononi point out that “when we dream [...] we are virtually disconnected from the environment – we acknowledge almost nothing of what happens around us, and our muscles are largely paralyzed. Nevertheless, we are conscious, sometimes vividly and grippingly so” (Koch and Tononi, 56). To my mind, this is not a convincing assertion of disconnection from our environments and our bodies. Just as Moravec’s brain in a vat scenario asserts the body in ideation – the *sense of* a body – as a component of embodiment in and of itself, dream states create environments in which we are immersed and to which we are connected. One may be disconnected from the environment outside the dream state; but to suggest that one is *de facto* disconnected from one’s environment relies on a too narrowly construed conception of the environment. The question is not whether we are connected to the environment, but rather to which environment we are currently connected, and in what temporality this connection operates.

In Koch and Tononi’s revised Turing test, how connected is the hypothetical human or conscious machine to his, her, or its environment? Is not the point of their

²⁷ See Descartes, *Meditations on First Philosophy*, 60-61.

Turing test that the machine, or human, speculatively yet rationally enter the environment depicted in the image? And do not the scenarios the machine is asked to produce (the more scenarios it can generate, the better, as “your *level* of consciousness has to do with how much integrated information you can generate” (Koch and Tononi, 58)) draw from a situated and experiential knowledge of the world? Sensory or otherwise, is it ever possible to be “disconnected from the environment”?²⁸ Within their discussion of consciousness, I would argue for a more expanded sense of the various ways in which we interact with our environments, as well as the multiple environments with which we can interact. Further, Koch and Tononi themselves move toward this expanded sense of the environment in their Turing test discussion, in which passing draws heavily on situated experiences and the production of various ideational situations and environmental interactions.

Koch and Tononi suggest two strategies to produce machine consciousness: model the brain or evolve a machine. Because of the complexity of the human brain, they view the latter as a more plausible strategy. Thus, the task is to continue to develop machines to become more robust, to produce greater amounts of integrated information. Their example is computer vision systems that have the capacity to categorize images in the real world. While these systems “perform admirably,” they still have far to evolve in order to be conscious. “Yet such systems are still very brittle. Move the test setup from cloudy New England to the Brighter skies of Southern California and the system’s

²⁸ Koch and Tononi do not rule out that this interaction with the environment in early childhood may be necessary for the development of consciousness.

performance suffers. [...] the range of what they can recognize must increase considerably to encompass essentially all possible scenes” (Koch and Tononi, 59). In their discussion of evolving this system to consciousness, Koch and Tononi articulate a need to be *more* responsive to the world, *more* situated. These computer vision systems, which aim to interact visually with the real world, must be able to operate in different environments, adjusting themselves accordingly so that their performance is not contingent on cloudy days. Koch and Tononi’s consciousness, then, is an unmistakably worlded consciousness. If the world (sunny days) unfolds identifiable problems for the vision computer system in its evolution to consciousness, and the generation of large amounts of worlded knowledge is the grounds by which a machine can pass their Turing test for consciousness, then the world is crucially intertwined with Koch and Tononi’s discussion of consciousness. Worldedness, we recall, is also a central concern for Brooks, who proposes embodiment as a necessary component in addressing this concern. We might then understand the expanded embodiment articulated variously by Moravec, Rucker, and Brooks as a necessary component for machine consciousness and the machine evolutionary process by which consciousness will emerge.

Like Koch and Tononi, Brooks is similarly optimistic about machine consciousness: “Eventually, we will create truly artificial intelligences, with cognition and consciousness recognizably similar to our own.” Brooks is also optimistic about our co-existence with these intelligent and conscious machines: “I’m an optimist. I believe we will all get along” (“I, Rodney Brooks,” 72-73). Brooks understands machine

intelligence and consciousness as worlded and embodied, and, as he hints at in the above quotation, social.

3 Worlds beyond the Uncanny Valley: Philip K. Dick and the Robots of Cynthia Breazeal and Dave Hanson

Kismet is the world's first robot that is truly sociable, that can interact with people on an equal basis, and which people accept as a humanoid creature. They make eye contact with it; it makes eye contact with them. People read its mood from the intonation in its speech, and Kismet reads the mood of people from the intonation in their speech. Kismet and any person who comes close to it fall into a natural social interaction. People talk to it, gesture to it, and act socially with it. Kismet talks to people, gestures to them, and acts socially with them. People, at least for a while, treat Kismet as another being. Kismet is alive. Or may as well be. People treat it that way. (Brooks, *Flesh*, 65)

Kismet, "the world's first robot that is truly sociable," is "truly sociable" only because humans treat it as human, only because humans anthropomorphize. Sociable robots, robots that interact socially with humans, are humanoid machines that operate through both human resemblance and anthropomorphization. The human knows, without question, that the robot is not human, but the human nonetheless interacts with the robot as if were. Unlike ELIZA, which facilitated an intimacy between human and computer program through the species-uncertainty allowed by the obscuring computer screen, contemporary U.S. roboticist Cynthia Breazeal's sociable robots, Kismet and Leonardo, activate and participate in this intimacy without the screen.

And how do we get from anthropomorphization, the Turing test, and the screen, to the anthropomorphic intimacy of the unscreened encounter between human and sociable robot? Like much of sociable robotics, we begin at the face, in its privileged status in certain discourses of the human, then as site of emotions in both Silvan Tomkins' theorization of affect and the anthropomorphization of Cynthia Breazeal's sociable robots. Reading the face as such reveals the limits of thinking the human through

anthropomorphization. Indeed, anthropomorphization does not always provide a way out of the tautological organization of the human, but can sometimes highlight and reinscribe this tautology. By moving from Turing to Brooks to Breazeal, I suggest that these anthropomorphic operations are, ultimately, decisions and choices made within various disciplines about what conception of the human they will value or legitimize. And yet, the human always exceeds how we see ourselves in the robot; and the robot, likewise, can claim modes of being that are not limited to the restricted interactions for which they are designed. To this end, I will turn to the robotic works of David Hanson and the science fiction of Philip K. Dick, which significantly influenced Hanson in his robotic creations and philosophies. In Dick's *We Can Build You* and *Do Androids Dream of Electric Sheep?*, robotic human-identity is a necessary condition for the recalibration of the human around empathetic and compassionate behavior. Hanson, in his pursuit of physical identity, similarly opens up new modes of relational sociality between humans and robots, relations that move outside of an anthropocentric positionality and instead posit new modes of being and new ways to conceptualize the human.

3.1 Design Choices and the Uncanny Valley

The distinction between appearance and behavior, particularly as it pertains to emotion, is interestingly complicated and conflated in sociable robotics. Appearance, or

the expression of emotion, is itself a kind of behavior.¹ Attending to emotion is a complex, multiply-sited task that incorporates decisions about what behaviors the robot will exhibit, what these behaviors will look like, how specific emotions will be expressed, and the overall design of the robotic body. This last factor - the appearance of the robot – plays a crucial strategic role in eliciting an affective anthropomorphic relationship between robot and human. The more the robot looks like a human, the more likely the robot will be anthropomorphized, or, treated as if it were human. In other words, the more easily anthropomorphized the robot, the more successful the interactivity.²

In 1970, Japanese roboticist Masahiro Mori introduced the theory of the Uncanny Valley. Mori's theory asserts that some human resemblance in robots produces a positive response in a human; however, if the robot is designed to resemble a human too closely, the positive sense of familiarity would become a negative sense of unfamiliarity. Thus, Mori suggests that roboticists maintain a safe distance from human-identity: "We predict that it is possible to produce a *safe familiarity* by a nonhumanlike design. So designers please consider this point" (Mori, my emphasis). Designing according to the uncanny valley can be tremendously productive, as Breazeal's robots demonstrate.

¹ This interrelation between appearance and behavior has significant practical implications for the traditional disciplinary separation between robotics engineering (which has in the past been charged with constructing the robotic "body" that houses the computing mechanisms) and computer science (which was responsible for the "internal" computing processes by which the robot functions).

² Roboticists Spexard et al. also suggest an easily anthropomorphized appearance, as well as familiar behaviors, as the key to human interaction with a robot: "Besides a flexible dialog system and speech understanding, an anthropomorphic appearance has the potential to support intuitive usage and understanding of a robot, e.g., human-like facial expressions and deictic gestures can as well be produced and also understood by the robot" (Spexard et al., 852). Roboticist Maja Mataric similarly holds appearance, behavior, and situatedness (as properties of embodiment) as necessary properties for human and robot interactivity (Ishiguro et al., 82).

However, the adherence to Mori's theory can also not only maintain, but in fact reify and create seemingly *a priori* and inviolable distinctions between human and robot.

Mori astutely points out how, through the presence of the unexpected, the familiar can in fact become unfamiliar. However, Mori does not consider the multi-directional trajectory between the familiar and the unfamiliar. Just as what is familiar can become unfamiliar, what is at one moment unfamiliar can also become familiar. Indeed, the unfamiliar, given time and exposure, can almost always become familiar.³ This question of the familiar must always be understood in relation to temporality. Mori introduces the prosthetic robotic hand as an example of the uncanny valley. Were the prosthetic robotic hand constructed as indistinguishable from a human hand, we would experience "a sense of strangeness" upon realizing the hand was not human (Mori). Mori describes this prosthetic hand as "too real." When touched, the prosthetic hand may in fact provoke an uncomfortable, unfamiliar response, as our experience of the coldness of the prosthetic hand conflicts with our expectation that the hand will be warm. However, were we to incorporate this prosthetic hand into our daily lives and activities, I suspect that this disjunction would disappear over time. The prosthetic hand's lack of warmth would no longer seem strange or unfamiliar, but perhaps become so familiar as to be invisible.⁴

³ The question of translation question arises here. While familiarity is a temporally-situated characteristic, something like "unpleasantness" can exist across temporal moments. What is unpleasant to me upon the first encounter can also be unpleasant upon the hundredth encounter. Many thanks to Jui-an Chao for discussing this translation with me.

⁴ Several roboticists and theorists highlight that Mori's theorization of the uncanny valley is not based on empirical or theoretical evidence. For example, Gee et al. write, "The acceptance of the 'Uncanny Valley' is wide spread, although the empirical data has yet to fully substantiate Mori's concept" (153). They also suggest that the lack of empirical investigation into the uncanny valley can be explained by the

Like Bartneck et al., I view familiarity as a temporally-situated concept: “Familiarity depends on previous experiences and is therefore likely to change over time. Once people have been exposed to robots, they become familiar with them and the robot-associated eeriness may be eliminated” (368). The familiar emerges over time. There is no *a priori* familiar, but only the familiar as the once-unfamiliar. In light of this temporally-specific familiarity, Bartneck et al. suggest that we reconsider the theoretical weight the uncanny valley has enjoyed (368-369).⁵ The uncanny valley, then, reifies through atemporalization, freezing the unfamiliar (the not-human) and excluding the processes by which it can become familiar.

undesirability of its investigation: “In addition to testing the positive aspects of the graphs it would be essential to conduct experiments using robots designed specifically to test the negative aspects. However this would require research platforms that invoke repulsion or fear, two emotions the majority of humanoid robots are designed to carefully avoid” (152). The following articles also comment on the lack of empirical proof for Mori’s theory: Bartneck et al. “Is the Uncanny Valley an Uncanny Cliff?”; and A. Tinwell and M. Grimshaw, “Bridging the Uncanny: An Impossible Traverse?” Additionally, Karl F. MacDorman, one of the English translators of Mori’s original paper, and Hiroshi Ishiguro point out that since the publication of Mori’s paper, the uncanny valley has remained untested: “In fact, there has been little direct scientific investigation of Mori’s uncanny valley hypothesis in the past 35 years” (MacDorman and Ishiguro, “The Uncanny Advantage,” 309). Regarding the adherence to the uncanny valley despite the speculative nature of its claims, David Hanson discusses Mori’s graph: “Mori put forth the uncanny valley as speculation, not as a true scientific theory. But he drew it as a graph, and that made it seem more scientific. It’s not a scientific hypothesis that was tested with data, though” (qtd in Geller, 12).

⁵ Bartneck et al. also suggest that the way that Mori’s theorization has been taken up in its English translation is somewhat removed from its original Japanese. “Shinwa-kan,” which has been translated as “familiarity,” is more appropriately translated into “likeability” or “affinity” (369). Similarly, translation questions have been raised about the translation of “bukimi no tani” into “uncanny.” Indeed, this translation to uncanny undergirds Mori’s theory with a substantial Western theoretical genealogy that Mori himself may not have intended. A number of scholars from both the humanities and the sciences have used Freud’s theorization of the uncanny (see Freud, “The Uncanny”) to frame their discussions of the Uncanny Valley. For example, Geller traces the genealogy of the Uncanny Valley back to Freud, and from there to Ernst Jentsch, whose work on the uncanny significantly influenced Freud’s. (“Overcoming the Uncanny Valley,” 11). Yet I am wary of importing Freud to Mori’s work without understanding the Uncanny Valley in the context of Japanese robotics, philosophy, history, and culture. We cannot assume that the decidedly Western theoretical genealogy drawn by Geller similarly organizes Mori’s theorization. It should be noted that Mori’s later work connects robotics to Buddhism (see Mori’s *The Buddha in the Robot*).

Androids are robots that are designed to be physically identical to humans. MacDorman and Ishiguro, whose android science uses android-human interaction to understand the human, do not dispute the validity of Mori's uncanny valley. However, while acknowledging the productivity of the uncanny valley, android science does not adhere to Mori's warning that roboticists should not design too closely to human appearance. MacDorman, in a separate article, suggests that while the experience of the uncanny might be unpleasant, it should not necessarily be avoided. Citing the popularity of the horror genre, MacDorman suggests that the uncanny can, in fact, be a desirable experience ("Subjective Ratings," 51).

MacDorman and Ishiguro, who define familiarity as "the absence of novelty," also point toward reintroducing temporality into the uncanny valley ("The Uncanny Advantage," 304). Mori discusses the feeling of strangeness produced by "an entire robot" or a mannequin (Mori). In response, MacDorman and Ishiguro suggest that the experience of the uncanny emerges from "the violation of expectations mannequins elicit *based on our past experience with them*" ("The Uncanny Advantage," 303, my emphasis). MacDorman and Ishiguro's theorization of the uncanny valley is fundamentally temporal, relying on experiences in the past. If the familiar is the *absence* of novelty, the uncanny, then, is the previously unexperienced, the new.

Mori suggests that adding movement to the aforementioned prosthetic hand would heighten the sense of the uncanny for the human (Mori). Mori's caution regarding robotic movement resides in the difficulty of accurately replicating human movement in

relation to the temporally-situated factors of velocity and speed.⁶ In discussing the challenge of designing human facial expressions in robots, Mori mentions the importance of speed in replicating laughter:

[...] laughing is a kind of sequence of face distortions, and the distortion speed is an important factor. If we cut the speed in half, laughing looks unnatural. This illustrates how slight variations in movement can cause a robot, puppet, or prosthetic hand to tumble down into the uncanny valley. (Mori, n. pag)

Mori's wariness of movement in robotic design can also be linked to movement's existence in time. To introduce movement is to introduce time.⁷ And time deeply problematizes the reification of the familiar and the human effected by Mori's theory. Brian Massumi links the subtraction of movement to the subject's reduction to its position. Subtracting movement "catches the body in cultural freeze-frame. The point of explanatory departure is a pinpointing, a zero-point of stasis. When positioning of any kind comes a determining first, movement comes a problematic second" (*Parables*, 3). Mori's theory of the uncanny valley begins with positions – human and non-human robot – and subtracts movement, both conceptual between these two positions, and robotic, which might activate this conceptual movement.

⁶ Regarding the prosthetic hand, Mori writes, "humanlike movement requires similarity of velocity and acceleration."

⁷ Aristotle theorizes time as unable to exist without motion. Aristotle (*Aristotle's Physics*. 80).

3.2 Kismet: Animal, Infant, and Lowering the Threshold

MIT roboticist Cynthia Breazeal, formerly a student of Rodney Brooks, creates robots that not only exist in the world of humans, but also interact socially with humans. “A sociable robot is able to communicate and interact with us, understand and even relate to us, in a personal way. It is a robot that is socially intelligent in a human-like way. We interact with it as if it were a person, and ultimately as a friend. This is the dream of a sociable robot” (*Designing Sociable Robots*, xi). Breazeal further nuances the question of machine intelligence by introducing sociality. And the goal of this social intelligence, which is modeled on and directed toward the human, is to elicit an anthropomorphization from the human, so that, as in the Turing test, the human and the robot interact with each other as if the robot were human. And in Breazeal’s “dream of a sociable robot,” this robot is not simply a human, but a friend.

Breazeal’s sociable robots, Kismet and Leonardo, which are modeled on small animals and human infants, do not step into the uncanny valley. Kismet, unlike its descendent Leonardo, appears more machine than animal. While its expressive components – its eyes, eyebrow, ears, and mouth – are both animal- and human-like, the rest of the hardware – its wires and circuits – is left exposed (see Figure 2).

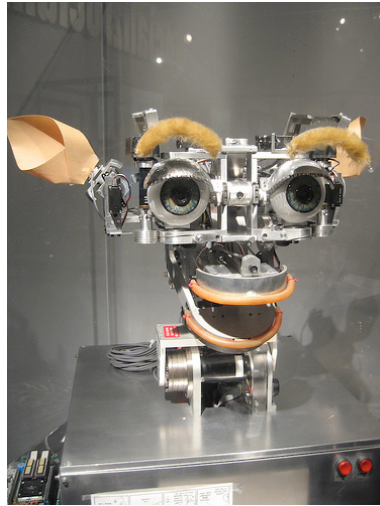


Figure 2: Expressive Kismet.
“Kismet is momentarily lifeless.” Nadya Peek. Flickr. Creative Commons Attribution 2.0 Generic.

Kismet is unmistakably machine, a robotic head with no body.⁸ According to Breazeal, humans anthropomorphize the unfamiliar in order to make the unfamiliar “familiar and understandable;”⁹ Kismet was designed to capitalize on this familiarizing process, by presenting the unfamiliar robot as familiar animal. The animal, while still an unfamiliar other, is more familiar than the robot. Thus, Kismet’s animal aspects function as “border concept,”¹⁰ mediating between human and robot, and facilitating human anthropomorphic response.

⁸ Kim, Lee, and Chung present three classifications of robotic faces, based on the robot’s skin: real, mechanical, and mascot-like. “In brief, this classification is based on the existence and flexibility of the robot’s skin. The real type of robot has flexible skin, the mechanical type has no skin, and the mascot type has hard skin” (679-680). Kim et al. place Kismet in the mechanical robot category, and Leonardo in the real robot category.

⁹ “The human engaging the robot will tend to anthropomorphize it to make its behavior familiar and understandable” (Breazeal, *Designing Sociable Robots*, 10).

Kismet-human interactions rely on relations of familiarity and sympathy to elicit a specific form of anthropomorphization. Breazeal's aim is not for Kismet and the human to interact as equals; rather, Kismet was designed to resemble both animal and human infant so that the human would take on a mentoring, teaching role (Breazeal and Aryananda, 87). In other words, Breazeal aims for Kismet and the human to have an explicitly *unequal* dynamic. As a strategy, this makes a great deal of sense: while Kismet is animal- and infant-like enough to be anthropomorphized, the infant-like appearance invokes an anthropomorphization through the interactive dynamic of teacher-student or parent-child.¹¹ Similarly to Weizenbaum's ELIZA, Breazeal's Kismet elicits anthropomorphization by providing a specific context for the interaction. Within this context, Kismet activates an unequal anthropomorphic dynamic by keeping the human's expectations of Kismet's capabilities low, and prompting the human to take the lead in these interactions (Breazeal and Aryananda, 87). "Lowering the threshold," then can be read as a necessary supplement to the absent Turing screen.

¹⁰ Joanna Zylinska challenges the use of the animal as a border concept that mediates between human and machine (Zylinska, "Of Swans," 149).

¹¹ For a further discussion of this sympathy effect and its function in "lowering the inhibition threshold" in robotics, see Sosnowski et al., 3114.

3.2.1 Affect-Emotion, Anthropomorphization, and Breazeal's Sociable Robots

Breazeal holds emotion as crucial to establishing a fluid, “natural” (anthropomorphic) interaction between human and robot – emotion is the key to designing robots that can socialize with humans (Breazeal, *Designing Sociable Robots*, xi). Moving away from AI's metric of intelligence as natural language fluency, Breazeal organizes robotic intelligence around the emotions and sociality. This turn to emotion is also seen in Rosalind Picard's affective computing.¹² As described by Picard, affective computing takes as its first premise that emotion, rather than hindering cognition and judgment, is in fact a crucial component of intelligence, cognition, and judgment, both human and artificial. “I have come to the conclusion that if we want computers to be genuinely intelligent, to adapt to us, and to interact naturally with us, then they will need the ability to recognize and express emotions, to have emotions, and to have what has come to be called ‘emotional intelligence’” (Picard, *Affective Computing*, x).

Neither Picard nor Breazeal are interested in producing a definition of the emotions; rather, they are interested in how emotions can be used to facilitate human-machine sociality. For Picard, the question of definition, which she dismisses, is one of dualism;¹³ and for Breazeal, the question dismissed is one of primary definition. Breazeal is not concerned with what emotions are, but rather with how they work:

¹² Picard, with whom Breazeal co-authored the essay “The Role of Emotion-Inspired Abilities in Relational Robots,” founded and directs the Affective Computing Group at MIT.

¹³ “The distinction – are emotions physical first or cognitive first – is not as important to us as the question, ‘How can emotions be generated in computers, recognized by computers, and expressed by computers?’” (Picard, *Affective Computing*, 23).

... the question of whether or not robots could have *real* emotions is irrelevant to our purposes. [...] as we continue to design integrated systems for robots, that complement its cognitive capabilities with those regulatory, signaling, biasing, and other useful attention, value assessment, and prioritization mechanisms associated with emotion systems in living creatures, we will effectively be giving robots a system that serves the same useful functions, no matter what we call it. (Breazeal, "Function Meets Style," 187)

For Breazeal, to know what emotions are is to understand how they work, to create the robotic apparatuses by which emotions emerge and function as communicative media.

In Kismet and Leonardo, emotion is largely localized in the face. Kismet is designed to elicit anthropomorphization through the appearance of an emotional life; and this appearance of emotion, as robotic behavior, is played out in and on Kismet's face (indeed, Kismet is virtually all face). In Breazeal's sociable robots, appearance and behavior, through the privileged site of the face, are more than simply intertwined, but in fact each other's necessary facilitators.¹⁴

Kismet possesses eight emotions, each with its corresponding facial expression. Designed to facilitate anthropomorphization through affective communication, Kismet's range of emotive expression is comprised of "three degrees of freedom to control gaze direction, three degrees of freedom to control its neck, and fifteen degrees of freedom in other expressive components of the face (such as ears, eyebrows, lips, and eyelids)" (Breazeal and Aryananda, 86-87). Kismet's six emotions are drawn from "the Ekman six," the six basic emotions as theorized by psychologist Paul Ekman. The emotions consist of anger, disgust, fear, sorrow, surprise, and happiness (Breazeal and Aryananda,

¹⁴ Kismet is commonly cited as a benchmark for sociable robots. For example, Deniz et al call Kismet "the most influential social robot" (Deniz et al, 71).

96). Kismet's emotion expression system also employs Ekman and Wallace V. Friesen's Facial Action Coding System (FACS), which maps emotions according to the movements of the muscles in the face (Breazeal, *Designing*, 173-175; Ekman and Friesen, 1978). As Ekman, in his work on the emotions and the face, was significantly influenced by Silvan Tomkins' theories of affect,¹⁵ I will draw from Tomkins' work to theorize the face as privileged site of expression. Further attending to the face's similarly privileged status in disparate discussions of subjectivity and intersubjectivity, I will also discuss the face in philosophical and neuropsychological discourses.

3.3 The Face I See and The Face You Know: Intersubjectivity and Recognition

Emmanuel Levinas locates the face as the site of intersubjectivity and sociality – of ethics itself. The face, according to Levinas, is an inescapable exteriority; it is that by which we encounter and claim others, and that by which others encounter and claim us. Through “the extraordinary exteriority of the face,” we are revealed as never having been wholly self-possessed; conversely, we extend our notion of possession to “the extraordinary exteriority” of the faces of others, to claim them as ours just as they claim us as theirs (*Alterity and Transcendence*, 103).

According to Levinas, my face is my face, but it is also simultaneously the face of the other and the face that speaks to the other. Therefore, my face is not wholly mine: it

¹⁵ Ekman was first persuaded of the universality of facial expressions, around which much of his work is now organized, through his encounter with Tomkins' work (Ekman, *The Face of Man*, 3).

“resists possession, resists my powers” (*Totality and Infinity*, 197). One might even suggest that we are less possessed of our own face than we are of the faces of others. While our familiarity with our face is uniquely ours, others possess our face in ways that, without significant technological intervention, remain unknowable to us.¹⁶ Likewise, we possess the other’s face in a way that the other cannot, and in a way that evades our possession of our own face. It is, as concentrated in the site of the face, this always-partial possession that makes us constitutively susceptible to the other, a susceptibility that is more than possibility, but rather certainty.¹⁷ It is through this certain susceptibility that we can understand the human in a way that begins to test the limits of its own tautology; to be human is to necessarily extend beyond oneself. The human exceeds itself through the intersubjectivity, through encounter with the face of the other: “The face-to-face is a relation in which the *I* frees itself from being limited to itself (which it thus discovers), from its reclusion within itself” (*Alterity and Transcendence*, 56). In other words, only when the “I”, by way of the face-to-face encounter, frees itself from the limits of the self, does the “I” become human.

¹⁶ While some technologies (such as cameras and true non-reversing mirrors) allow us to encounter our own face as others see it, other technologies maintain the distance between our face as we see and know it, and our face as others encounter and possess it. For example, the videoconferencing platform Skype provides me with both an image of the other’s face on my screen, as well as a smaller image of my own face. However, my face on my screen is not the face that appears on the screen of the other; my face on my screen is a mirror reflection, while my face on the screen of the other is a non-reversed, “true” image. Thus, reversibility exists as a gap of knowability between myself and the person with whom I am communicating.

¹⁷ According to Levinas, it is this susceptibility, this encounter with the face of the other, that activates ethics itself. Ethics thus defined is an injunction to be responsible for this other. (*Alterity and Transcendence*, 35).

In light of the recent advances in cosmetic surgery and body morphing technologies, Bernadette Wegenstein argues that the face is no longer the privileged site of subjectivity and otherness. These modifying technologies have heralded what Wegenstein calls the “post-facial era,” in which “the face ‘proves’ itself to be a code precisely by the fact that its role can be taken over by any other body part. Toward this end, the head and its face have lost their position of prominence” (235). While the face may no longer be the primary arbiter of coding and (over-)determination, when read in light of certain technologies, such as sociable robotics, the face remains a uniquely privileged, and thus powerful, coding mechanism.

3.3.1 Intersubjectivity as Indistinguishability: The Rubber Hand Illusion and Self-Face Recognition:

The work of neuropsychologist Manos Tsakiris explores the face and the hand as sites of intersubjectivity. In his studies, the face and the hand emerge as two sites in which intersubjectivity itself becomes so powerful, one cannot distinguish one’s own face or hand from the other’s, oneself from the other. Based on an original study by Matthew Botvinick and Jonathan Cohen, the Rubber Hand Illusion (RHI) is a phenomenon in which participants, viewing a rubber hand on a screen, mistake this rubber hand for their own.¹⁸ Tsakiris and Patrick Haggard’s study, like that of Botvinick and Cohen, observes that coinciding visual and tactile sensations can cause individuals to mistake a rubber

¹⁸ Matthew Botvinick and Jonathan Cohen, “Rubber hands ‘feel’ touch that eyes see.” It is particularly interesting to think about Botvinick and Cohen’s Rubber Hand Illusion in light of Mori’s discussion of the hypothetical prosthetic hand.

hand on a screen as one's own hand.¹⁹ Tsakiris and Haggard describe this misidentification as a "mislocalization of the stimulated body part" (91). According to Tsakiris and Haggard, "Watching a rubber hand being stroked synchronously with one's own unseen hand causes the rubber hand to be attributed to one's own body, to 'feel like it's my hand'" (80). In the RHI, vision and touch opens up a channel, through mislocation, between the subject and other.²⁰ Tsakiris and Haggard diverge from Botvinick and Cohen by reframing the misidentification in terms of phenomenological experience. "The phenomenological content within the illusion is a description of one's own body and not a description of the stimulation." In other words, one experiences the RHI not as simulative, but rather as bodily coherence in its Rubber Hand extension. Indeed, this channel of misidentification, or extended identification, illustrates how our very physical and ideational limits and experiences of ourselves can be recalibrated.

In a similar experiment, Tsakiris finds that this "strange phenomenal experience of ownership" applies not only to the hand, but also to the face.²¹ The participant, in

¹⁹ Tsakiris and Haggard observe that only specific combinations of visual and tactile sensation produce the RHI. RHI does not occur when the participant's hand and the hand on the screen are not stroked at the same time (incongruence of sensory efference), nor when the hand on the screen is at a ninety degree angle from the participant's hand (incongruence of hand position) (Tsakiris and Haggard, 81, 84).

²⁰ Tsakiris' studies recall Giacomo Rizzolatti and Laila Craighero's discussion of mirror neurons in relation to imitation learning in monkeys and humans. Rizzolatti and Craighero observe that in monkeys, grasping an object and watching a human hand grasp an object trigger the same neurons in the brain. Similarly to Botvinick and Cohen's work on RHI, Tsakiris and Haggard's revision of the RHI, and Tsakiris's self-face recognition studies, Rizzolatti and Craighero find that some level of congruence of object-directed action is required to activate mirror neurons ("The mirror-neuron system").

²¹ Silvan Tomkins cites the hand's attention to the face – "the hand acts as if the face is the site of feeling" (Tomkins, *Vol. I*, 210) – as evidence of the face's privileged status. For a more detailed discussion of the relationship between the hand and the face (for example, the way the hand cradles the face when one is tired, and hides one face when one feels shame), see Tomkins, 210-211.

Tsakiris' study of visuo-tactile sensation and self-face recognition, is seated in front of a screen. This screen displays a movie in which the participant's face, across one hundred frames, morphs into someone else's face. While the movie is playing, the experimenter touches the participant's face with a paintbrush. The morphing face on the screen is also touched with a paintbrush, in the identical location as on the participant.²² When the movie morphs from self to other, the participant identifies the moment at which the face becomes more other than self at frame forty-four of one hundred. When the movie morphed from other to self, the participant identified the moment when the face become more self than other at frame sixty-two of one hundred. In both cases, the face was recognized as one's own below the fifty frame mark (Tsakiris, 2). What emerges is not that we mistake the face of the other for our own, but rather the facility with which our own faces so easily become foreign, unrecognizable to our own selves. In an inversion of the RHI illusion, Tsakiris and Haggard's self-face recognition study does not demonstrate the misidentification of the other's face for ours, but rather the facility with which our own face so easily becomes foreign, unrecognizable to us. Tsakiris' study suggests a certain alienation from our own face, an inability to recognize our face as our own and not other. The face is a confused site of subjectivity, existing as so many iterations within the same spatio-temporality – as self, for self, for other, for self as other, for other as self.

²² In an email correspondence, Tsakiris explains that the results were identical whether the face of the participant was a reversed mirror image, or an unreversed image. Neither face emerged as more familiar and recognizable to the participant (Tsakiris, personal communication).

3.4 Affect-Emotion

Silvan Tomkins, whose influence on Ekman I noted above, similarly identifies the face, as a medium for affective communication, as a primary site of encounter between self and other, largely on account of the substantial amount of information the face can convey and receive.²³ Influenced by cybernetics' attention to feedback loops and information circuits, Tomkins theorizes affects as motivating amplifications that imbue our relations to objects, situations, images, or others with urgency. Affect is that by which things matter: "without [the affect system's] amplification, nothing else matters – and with its amplification, anything else *can* matter" (*Vol. III*, 6). Tomkins' affect references the processes of amplification and urgency, of mattering, that structure how we experience the world. In the recursive language of the feedback system, Tomkins describes affect as the following: "affect connects its activator and the response, which is also amplified" ("Modification in the Theory," 95). In other words, affect is the way that we encounter, are susceptible to, and unquestionably changed by the external world; affect thus describes both how we are connected to the world, and the world to us. In other words, affect is that which sutures us to the world.

Across his various writings on affect, which span nearly thirty years,²⁴ Tomkins discusses affect as "primary biological motivator, more urgent than drive deprivation, pleasure, and physical pain," as "analogue amplifier," and, most pertinent to this

²³ Tomkins, who includes the voice in the face (Tomkins, *Vol. I*, 215), points to the "relative density of receptor-effector units in the face" (Tomkins, *Vol. I*, 215).

²⁴ Tomkins' series, *Affect, Imagery, Consciousness*, consists of four volumes, the first published in 1962, the third published in 1991, and the fourth, published posthumously, in 1992.

discussion, as facial behavior (“Modifications in the Theory,” 86; 88). “In short, affect is primarily facial behavior. Secondarily it is bodily behavior, outer skeletal and inner visceral behavior. When we become aware of these facial and/or visceral responses we are aware of our affects” (*Vol. I*, 205-206).²⁵ For Tomkins, affect is not just primarily located in the face; affect is the face itself in behavior and in action. In relation to affect, the face is what the face does.

Eve Kosofsky Sedgwick and Adam Frank, in their introduction to their edited collection of Tomkins’ work, discuss the face as the site of affect’s occurrence: “More than the place where affects are *expressed*, Tomkins shows the face to be the main place in the body – though by no means the only one – where affect *happens*” (Sedgwick and Frank, 30). Affect’s happening is not just expression, but also communication. Ekman describes the face in Tomkins’ work as “inform[ing] the self, not just others. Feedback of the facial response *is* the experience of affect” (Ekman, “Silvan Tomkins,” 210). While the self is informed by affect in the face by *feeling* the responses (the rush of blood to the face, etc.), the other is informed by affect in the face by *reading* the visual signs of these feelings and the correlative facial expressions.

Tomkins distinguishes between these two faces and these two subject positions through two experiences of the face: seeing and inhabiting. He writes, “The face one sees is not so different from the face one lives behind” (“What and Where,” 227). While

²⁵ Tomkins revises this description in Volume III, in which he moves the location of affect from the muscles of the face – “Affects are sets of muscle and glandular responses located in the face and also widely distributed through the body, which generate sensory feedback which is either inherently ‘acceptable’ or ‘unacceptable’ (Tomkins, *Vol. I*, 243) – to the skin of the face. For a detailed discussion of “the skin receptors of the face” as “the major locus of analogic amplification,” see *Vol. III*, 9.

these two experiences may not be “so different,” this difference between living behind a face and seeing a face is the difference between self and other. Within Tomkins’ affective interactivity, reading (seeing) is closely intertwined with feeling (living behind), as each is the other’s essential supplement.

Tomkins lists the primary affects as Interest-Excitement, Enjoyment-Joy, Surprise-Startle, Distress-Anguish, Fear-Terror, Shame-Humiliation, Contempt-Disgust, and Anger-Rage (“What Are Affects,” 74).²⁶ Tomkins also lists each affect’s corresponding “facial behavior,” as it can be read and identified by others:²⁷

Interest-Excitement: eyebrows down, track, look, listen
Enjoyment-Joy: smile, lips widened up and out
Surprise-Startle: eyebrows up, eye blink
Distress-Anguish: cry, arched eyebrow, mouth down, tears, rhythmic sobbing
Fear-Terror: eyes frozen open, pale, cold, sweaty, facial trembling, with hair erect
Shame-Humiliation: eyes down, head down
Contempt-Disgust: sneer, upper lip up
Anger-Rage: frown, clenched jaw, red face (“What Are Affects” 74)

Tomkins is concerned with what the face *does* in relation to affect,²⁸ and, for Tomkins, what the face *does* includes the primary function of expression, of making the emotional state of the individual legible to others.

²⁶ One of Tomkins’ later revisions to his theory of affect is the separation of the joint affect, Contempt-Disgust, into two separate affects.

²⁷ Ekman largely approaches emotions in the face as a universally expressed code that one can learn and master, in order to better communicate with and correctly read others. For example, see Paul Ekman and Wallace V. Friesen’s *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*.

²⁸ At times, it seems that Tomkins is primarily concerned with what the face looks like, with the face as it is read, less than the face as it is felt or lived. In some ways, for Tomkins, the face is less *of* the other, as Levinas suggested, than *for* the other. This difference in preposition, from “of” to “for,” is the difference between intersubjectivity and the somewhat more facialized subject positionality that I read in the works of Tomkins, Ekman, and Breazeal.

While Tomkins discusses the emotions as felt, he is primarily interested in the readability of emotions. For Tomkins, the language of affectivity is constituted by the emotions as they appear on the face.²⁹ Emotions are not affects if they are unidirectional, operating solely inwardly on the individual. Rather, emotions are affects *because* they operate multi-directionally. Indeed, this subtle difference between emotions as they are felt, and emotions as they are dually felt and expressed in order to be read might account for the frequent conflation between the terms “emotion” and “affect.”³⁰ For the sake of precision, I will use the term “affect-emotion” to refer to this multi-directional affect that is organized around the emotions.

3.4.1 Sociable Robots and Affect-Emotion

Affect-emotion is the primary means by which Kismet and Leonardo elicit anthropomorphization, drawing the human into an affective-emotional feedback circuit, an interactive intimacy. Reading human-robot sociality through affect-emotion makes visible both the decidedly tautological structure of the human, and the limits of thinking about the human through any theoretical lens (anthropomorphization, affect-emotion) that participates in this tautology by organizing itself around the category of the human.

Through the emotions, Breazeal’s sociable robots substantively incite and shape the

²⁹ Tomkins discusses affect as distinct from, but in a necessarily co-operative relationship with, cognition. Moving from Tomkins’ discussion of affects as amplification and urgency, to his facial description of the specific affects, to the intertwined process of affect and cognition, one can understand how, for some, affect becomes conflated with something more like emotion.

³⁰ According to Brian Massumi, emotion and affect “follow different logics and pertain to different orders.” (Massumi, *Parables for the Virtual*, 27).

communication between robot and human. Within this feedback loop, which one might call “social interaction,” facial expression of emotions is the medium by which information is transmitted and received. Emotion expression is also the data that is transmitted and received by both the human and the robot.

According to Breazeal, sociable robots are designed in terms of their capacity to interact with humans: “[...] our approach is designed to support a rich and tightly coupled dynamic between robot and human, where each responds contingently to the other on an affective level” (“Affective Interaction,” 582). One might even say that Breazeal’s aim is less to design a robot with the capacity to interact with a human, but rather to design the interaction itself. Designing the interaction is accomplished by giving the robot the appearance of emotionality, which, if done skillfully, will prompt the interacting-human to imbue the robot with an internal emotional life. In other words, the human, upon seeing the signs of emotional expression, imagines that the robot possesses and feels emotions and the corresponding physio-cognitive experiences in the same way that we do. In human-sociable robot interaction, anthropomorphization occurs when the gap between robotic emotional appearance and emotional behavior is closed, when emotional appearance is itself the behavior.

Sociable robotics’ production of anthropomorphization through affect-emotion risks reifying emotional behavior-expression. This reification would foreclose rather than open up new possibilities for interactivity with machines, as well as with other

humans. Human-robot interactivity does not necessarily have to delimit as such, as my discussion of David Hanson's robots will suggest.³¹

3.4.2 Affective Anthropomorphization

Affect-emotion always exists as more than a single subject position. It is interactivity, reading to be read. Thus, appearance (what is readable) is a component of behavior (the act of reading), which is in turn a component of appearance (what is readable as a reflection of what has been read), and so on. It is from here that the interactive feedback dynamic between human and Kismet emerges. Equipped with the capacity to express as well as recognize emotion (in self and in other, respectively), Kismet draws the human into an emotional, intimate interaction that is effectively anthropomorphic. Kismet may not feel emotions in the same way that you or I might; and yet, this is the anthropomorphic leap that Kismet elicits. Because Kismet expresses emotions, we are inclined to attribute our experience of emotions to Kismet.

While emotion expression is key to Kismet's readability, and thus anthropomorphization, Kismet is also designed to identify and recognize the emotional state of the human, as "Basic affect recognition and expression are expected by humans in communication" (Picard, *Affective Computing*, 15). Kismet recognizes affect-emotion

³¹ In "The Time of Affect," Mark B. N. Hansen discusses Bill Viola's new media artwork, *Anima*, as "opening the imperceptible in-between of emotional states" through "technical expansion." This technical expansion produces a decidedly bodily experience of "self-affection," in which we can more fully experience our own subjectivity (589). Hansen's discussion of technologies that operate on and through affectivity engages a more productive and generative technological appeal to something like affect-emotion.

primarily by reading facial expression and voice. Kismet's vision system, its "attention system," is comprised of two stages: the pre-attentive stage, which processes motion, highly saturated color, and skin tones; and the limited capacity stage, which processes more complex, targeted visual information, such as facial expression recognition, the space between the eyes, gaze direction,³² and object identification, in a spatial field narrowed by the first pre-attentive stage.³³

Kismet is equipped with a set of preprogrammed prosodic contours, which map human voice pitch and energy patterns according to predetermined templates of affective intent.³⁴ Kismet's data set is based on psycholinguist Ann Fernald's classification of prosodic contours for approval, prohibition, attention, and soothing, which emerged from her study of adult speech and pre-verbal infants (see figure 10).

³² Kismet identifies the direction of the human's gaze, in order to respond accordingly, to share the gaze, to look in the same direction, to respond to the human's behavior in a way convincingly "human" way. (Breazeal *Designing Sociable Robots*, 79).

³³ For a more detailed discussion of both the processes of both the pre-attentive stage and the limited capacity stage, see Breazeal, *Designing Sociable Robots*, 62-72.

³⁴ Alex (Sandy) Pentland, head of the Human Dynamics Group at MIT, similarly relies on "nonlinguistic social signals" such as prosody, emotion expression, and posture to create "socially aware communications systems." These systems are technologies that create and expand social networks. These technologies analyze conversation using the metrics of activity level (amount of participant's speaking time), engagement (how one's participation in the conversation affects the pattern of the conversation), stress (measured through prosody), and mirroring (mimicking of each other's signals). See "Socially Aware Computation and Communication."

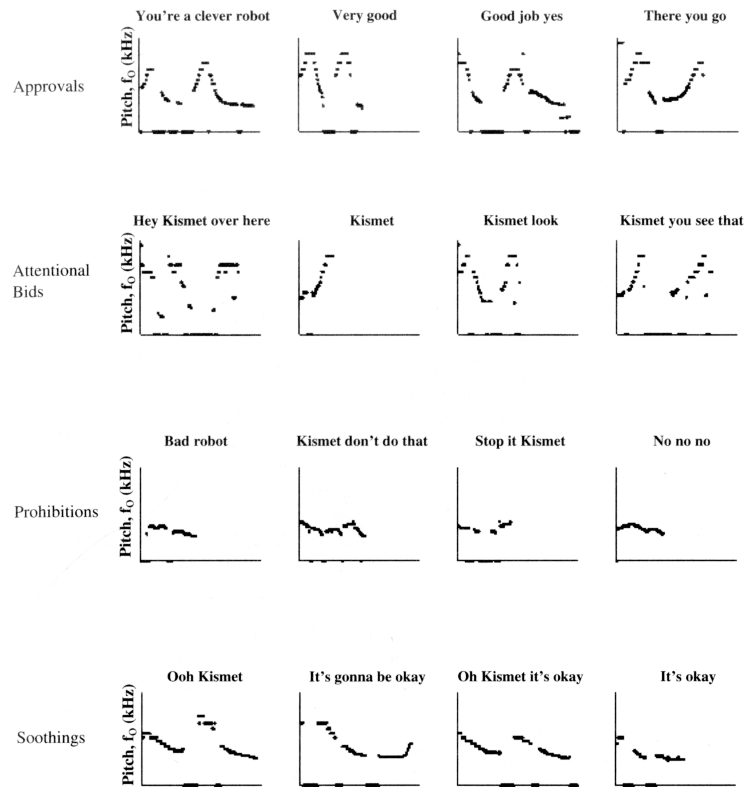


Figure 3: A selection of Kismet's preprocessed affective vocal patterns organized according to Fernald's prosodic classifications.
 Courtesy of Cynthia L. Breazeal, *Designing Sociable Robots*, and The MIT Press.

For Kismet, meaning is not delivered linguistically, but rather by the vocalization and prosody of the voice's delivery.³⁵ According to Breazeal, Kismet's appearance, in inducing the sympathy effect from the human, is also designed to elicit the often subsequent exaggeration of prosody that one exhibits when speaking to an infant or small animal. This exaggeration makes it easier for Kismet's vocal emotion recognition system

³⁵ Affect, as theorized by Massumi and Deleuze and Guattari, similarly operates in non-linguistic registers. However, while affect in Massumi and Deleuze and Guattari works against positionality, Kismet emerges from and preserves positions. Later in this chapter, I will identify the affect that operates against positionality as affect-force, so as to distinguish affect-force from affect-emotion.

to correctly identify the emotional state of the human.³⁶ This vocal recognition system was a factor in Kismet's design; in other words, Kismet's appearance needed to support and maximize the operation of this vocal recognition system, and the desired outcome of exaggerated prosody (Breazeal and Aryananda, 85).

Leonard Lawlor argues that the voice is claimed by humans as the sole province of the human; in so doing, the voice is then withheld from the category of the animal as a way of simultaneously identifying and perpetuating the prominence of the human (Lawlor). Breazeal's sociable robots expand the category enclosure which Lawlor views as characteristic of the human's species. Using the human voice to coax the human into anthropomorphizing the robot, Kismet opens up and expands the category of robot precisely through the human voice. Kismet, in responding to the prosody of the voice and producing the appropriate corresponding emotional expression, recasts the privileged status of the human voice in order to draw the human into an affective intimacy with the robot.

3.4.3 HRI and Leonardo

Breazeal distinguishes between two recently emerged interdisciplinary subfields of robotics: human-robot interaction (HRI) and human-robot collaboration (HRC): "[...] whereas interaction entails action *on* someone or something else, collaboration is inherently working *with* others" (Breazeal et al., 551). In human-robot collaboration, the

³⁶ Indeed, this strategy has proven successful: "Even the naïve subjects (male and female) use exaggerated prosody to address the robot" (Breazeal and Aryananda, 87).

human and robot are partners, working together toward a common external goal.³⁷ In HRI, the interactivity *itself* is the object or goal.



Figure 4: Expressive Leonardo.
“Leonardo.” jeanbaptisteparis. Flickr.
Creative Commons Attribution-Share Alike 2.0 Generic.

³⁷ Indeed, this collaborative model seems to be more in line with Sedgwick’s reading of Tomkins’ affect theory. Sedgwick discusses Tomkins’ affect in relation to the preposition “beside,” which holds, among other things, the promise of a non-dualistic relationality (Sedgwick, 8). And yet, if we read Tomkins’ affect-emotion alongside affect as theorized by Spinoza, Deleuze and Guattari, and Massumi, it becomes clear (in part due to the prepositions that fall in and out of applicability) that we are indeed speaking of two very distinct concepts. I read “beside” as not pertinent in discussions of affect, as “beside” still operates on the order of discrete subject positions. Affect as I understand it is not interested in this separation; indeed, affect is not about gap, or lack. What I find useful about affect is its operation outside of positionalities, expanding the limits of the human as opposed to contracting it around lack.



**Figure 5: Breazeal collaborated with Hollywood special effects company Stan Winston Studio to create Leonardo's robotic body.
"Leonardo." cognitive.evolution, Flickr.
Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic.**

Breazeal's Leonardo is equipped with sensory and expressive architecture similar to that of Kismet. Relying on similar attention and voice recognition systems, Leonardo's social-affective apparatus, that which lets it identify, recognize, and express an emotional state, is comprised of three systems: "an imitation-based emotion-empathy system, a shared attention system, and an object-based affective memory system" (Thomaz, Berlin, and Breazeal, 595; 594). Leonardo's physical appearance is more filled out in ways that continue to seek out, with more intensity and directedness, anthropomorphization by way of affect-emotions. Unlike Kismet, Leonardo has a body, with skin that resembles that of a living creature and hands that resemble those of an infant's.³⁸ These various surface

³⁸ Breazeal's team collaborated with the famous Hollywood special effects studio, Stan Winston Studio, whose extensive filmography includes the *Terminator*, *Alien*, and *Jurassic Park* films, as well as the recently released *Avatar*.

and body modifications, or evolutions, as it were, suggest the ontological limits of anthropomorphizing through affect-emotion. Leonardo, like its progenitor Kismet, reads affect-emotion by way of the face and the voice, expression and intonation (Tomaz et al., “An Embodied Cognition,” 593). This relationship between human and Leonardo, this feedback circuit of affect-emotion (the reading and responding, the responding as reading), remains organized around two poles. These poles are both organized around certain expectations of how a human acts, feels, and looks, reifying as much as reacting to this set of codified behaviors, training the interacting-human to remain within the parameters of a predetermined set of affect-emotions, as well as how to properly evoke them.

Lucy Suchman describes her interactivity with Kismet as “failed” because she, unlike Breazeal who has spent hours interacting with Kismet, was not familiar with Kismet’s predetermined communication cues.

The contrast between my own encounter with Kismet and that recorded on the demonstration videos [with Breazeal] makes clear the ways in which Kismet’s affect is an effect not simply of the device itself but of Breazeal’s trained reading of Kismet’s actions and her extended history of labors with the machine. In the absence of Breazeal, correspondingly, Kismet’s apparent randomness attests to the robot’s reliance on the performative capabilities of its very particular “human caregiver.” (Suchman, 246)

For Suchman, Kismet’s sociality does not quite live up to Brooks’ description (see beginning of this chapter). Suchman falls outside of the very narrow articulation of the human according to which Kismet is designed. As with all things unfamiliar, given time and exposure Suchman could become well-versed in Kismet’s affective language. While Kismet’s behaviors would not change, Suchman could adapt and modify her

participation, producing interactions with Kismet that approximate the fluidity of Breazeal's interactions with Kismet. In other words, Kismet could train Suchman to be Kismet's kind of human. In designing Kismet, Breazeal designs not only the Kismet-human interactivity, but also the human. Kismet, designed according to the uncanny valley and its atemporalization of the unfamiliar, relies on a temporally-situated process of familiarization to produce "successful" interactivities.

3.5 Beyond the Uncanny Valley: David Hanson and Philip K. Dick

Recent robotics work has begun to challenge Mori's theory by arguing that robots should bypass human resemblance and pursue identity. According to roboticist David Hanson, the pursuit of indistinguishability can yield more insight into human facial perception, but it is also the very means by which the uncanny valley itself can be bypassed.³⁹ Hanson's works articulate an alternative approach to anthropomorphization, one that seeks to confuse and expand the category of the human, rather than define and reify it. Approaching the uncanny valley less as prohibition than as playground, Hanson's robots "do not tiptoe around the uncanny valley, but dip in and out of the uncanny in an attempt to chart the territory and its boundaries" (Hanson et al., "Upending the Uncanny Valley," 26).⁴⁰

³⁹ "[T]he robots of Ishiguro and Hara do attempt to look as perfectly real as possible, to dodge potential 'uncanny valley' effects" (Hanson et al., "Upending the Uncanny Valley," 25).

⁴⁰ Hanson's robots, in this pursuit of human resemblance beyond the Uncanny Valley, attend to concerns that coincide with those of Tomkins' affect-emotion: expression of emotion via the musculature and skin of the face. For a description of the facial hardware as designed according to the "the natural muscle actions

Hanson et al. propose the Path of Engagement as a counterpoint to the Uncanny Valley. The Path of Engagement holds that humans have the capacity to adjust to increasing degrees of realism in robots: “As no ‘valley’ is inherent; anthropomorphic depictions can be either disturbing or appealing at every level of abstraction or realism. People simply get more sensitive with increasing levels of realism” (“Upending the Uncanny Valley,” 30).⁴¹ Thus, the uncanny valley is not a static metric, but rather shifts in relation to more realistic robotic models. If this is the case – and indeed, the evolutionary histories of “realist” representational technologies, such as photography, film, and video games, support Hanson et al.’s assertion – we must ask what is at stake in keeping robotic appearance at a certain distance from human “realism.”⁴² What efforts are involved in maintaining this distance between human and robot, and to what end? And, if the distance between human and robot must in fact be maintain and policed, what might this say about the inherentness of the distance itself? What is being protected by this distance, and what might be threatened if this distance decreases, if not disappears entirely? Hanson et al.’s Path of Engagement challenges the Uncanny Valley precisely in

of the face,” see Hanson et al., “Upending the Uncanny Valley,” 27. This hardware underlies the material of Hanson’s patented Frubber™ material, which is also designed to perfectly mimic the surface and movement of facial skin (For a discussion of the material composition of Frubber™, see Hanson et al., “Upending the Uncanny Valley,” 26).

⁴¹ While I find Hanson’s theorization of the Path of Engagement useful, I find the study conducted by Hanson et al. (the studies that support their argument for the Path of Engagement over the Uncanny Valley) interestingly problematic, as it measured “liveliness” or “dead[ness],” which is slightly different than measuring the humanness of the robot. For more on this study, see Hanson et al., “Upending the Uncanny Valley,” 30.

⁴² The anthropocentrism of Mori’s theory, and much of the debate surrounding the theory, is evident in the mostly uninterrogated use of “realism” to indicate “humanness.”

relation to these questions, to the Uncanny Valley's complicity in creating and preserving an untenable idea of the human that is narrow, static, unchangeable, and unaffected. The Path of Engagement, as a counterpoint to the Uncanny Valley, highlights that in humanoid robotics, the human as privileged category is created largely in its preservation.

Hanson's work in the area of human-like robotic appearance is somewhat of an anomaly in U.S. robotics. However, Japanese robotics has substantively attended to the art of the android, a robot that looks as identical as possible to human. Robots in Japanese culture and science have been figured quite differently than robots in Western culture and science. Japanese roboticists have largely pursued identity, whether in physical appearance or imitation of human motion, such as in Sony's walking, dancing, and orchestra conducting QRIO robots. And in the imaginative life of Japanese culture, robots have been depicted, less as threatening human adversaries, but often as friendly and not lacking in emotionality, such as the protagonist of the comic *Astro Boy*.⁴³ Japanese roboticists, such as Hiroshi Ishiguro pursue anthropomorphization through identity of appearance (see Figure 6). Ishiguro and his collaborator Minoru Asada centrally locate human-identity in "android science," an interdisciplinary approach to "find the essential factors of humanlikeness and realize a humanlike robot" (Ishiguro and Asada, 74-75). Like Breazeal's sociable robotics, android science also attends to

⁴³ J.P. Sullins posits that these two cultures' different relationship to robots stem from the longer histories of the two cultures' relationship to technology, philosophy, and religion. Sullins casts the East's Buddhist attention to the spiritual essence of all things, living and nonliving, as extending into Eastern science, while the West's approach is organized around dualism, which is evident not only in the scientific approaches, but also the Christian, Jewish, and Muslim religious traditions (Sullins, J. P. "Friends by Design"). For further discussions of Japan's relationship to life, death, and technology, see Ian Hacking's "The Cartesian Vision Fulfilled."

behavior and appearance. However, android science attends to humanlike behavior as a component of human-like appearance (as, indeed, “Appearance and behavior are tightly coupled” (74)). In order not to neglect this tight coupling of appearance and behavior, Ishiguro and Asada employ the framework of synergistic intelligence, which they define as

intelligent behaviors that emerge through interaction with the environment, including humans... Synergistic intelligence provides a new way of understanding ourselves and a new design theory of humanoids through mutual feedback between the design of humanlike robots and human-related science. (75)

Ishiguro and Asada’s synergistic intelligence emerges from the androids’ experiences in the world and interactions with humans, both of which cannot be isolated from the appearance and design of the android.



Figure 6: Hiroshi Ishiguro (left) and his android, Geminoid HI-1 (right). “o9_ars electronica festival.” Photo a_kep, Flickr. Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic.

Hanson’s approach to robotic design, as well as his response to the significant influence of Mori’s recommendation to roboticists, is deeply influenced by the fictional

writings and robotic philosophy of Philip K. Dick, who “fearlessly and ingeniously played with issues of the uncanny in his stories about androids and AI.”⁴⁴ In a speech written in 1975, Dick asserts that for him, android and human are not determined by their appearance, but rather by their behavior.⁴⁵ “‘Man’ or ‘human being’ are terms which we must understand correctly and apply, but they apply not to origin or to any ontology but to a way of being in the world [...]” (“Man, Android and Machine,” 202). “Human” does not signify origins, or even appearance, but rather “a way of being in the world.” Specifically, the difference between being in the world as human and being in the world as android lies in compassionate behavior; while the human has empathy (*Einfühlung*) for others, the android lacks the capacity for compassion, sympathy, and empathy. Thus, it is possible – and indeed, it happens frequently in Dick’s fictions – that an android is deeply compassionate, or human, and that a human is completely devoid of compassion and empathy, and thus android.

For Hanson, the appeal of Dick’s works lies in the centrality of compassion therein: “Dick’s robot stories highlight the importance of compassion in defining humanity, be it artificial or biological humanity” (“Humanizing,” 100). Hanson, inspired by Dick, reads humanity as organized around a capacity to acknowledge and care for others, rather than around material properties. Humanity, in other words, is neither limited to nor a *de facto* component of the human. In Dick’s fiction, androids often think

⁴⁴ Hanson et al., “Upending the Uncanny Valley,” 26. In an email, Hanson mentioned that his robotic vision owes a particular inspirational debt to the two novels I discuss in this chapter, *We Can Build You* and *Do Androids Dream of Electric Sheep?* (Hanson, personal communication).

⁴⁵ This behavior is distinct from the appearance-behavior of affect-emotion I discuss in relation to Cynthia Breazeal’s work.

they are human, or could easily be mistaken for human. The boundaries blur in ways that disorient and startle the reader, challenging the reader to question the meaning of human nature. Sometimes Dick's androids are more compassionate than any people in the tale (like the Lincoln in *We Can Build You*), while others lack compassion altogether (those of *Do Androids Dream of Electric Sheep*). For Dick, compassion distinguishes humans from androids. Compassion would be the essence of a social intelligence—one that values knowledge, life, and creativity. Sociopathic intelligence lacks these values and ultimately may destabilize the path towards greater creativity. Sociopaths burn libraries, commit genocide, and instate repressive totalitarian regimes. This is not the kind of AI that we want taking control of the world in our future (Hanson, 165-166).

Dick's redefinition of human and android is largely facilitated by his presumption of human-android identity. Through identity of appearance, Dick destabilizes and recalibrates what it means to anthropomorphize, as well as the effects of this anthropomorphization. In the fictional worlds of *We Can Build You* and *Do Androids Dream of Electric Sheep?*, the two Dick novels I will discuss in this chapter, androids are already indistinguishable from humans in appearance. Thus, appearance ceases to provide any indication of whether one is human or android. The task in the novels is distinguishing between android and human outside of appearance, and solely on the basis of behavior.

In addition to empathy, Dick articulates unpredictability as a metric of humanness. Conversely, predictability is a metric of androidism. Dick asserts that one becomes-android by behaving predictably, by following and obeying all rules and laws:

“Androidization requires obedience. And, most of all, *predictability*. It is precisely when a given person’s response to any given situation can be predicted with scientific accuracy that the gates are open for the wholesale production of the android life form” (“The Android and the Human.” 134). For Dick, predictability is not simply an index by which one can identify an android; rather, the android experience of predictability is the experience (or non-experience) of non- affectivity.

[...] possibly the difference between what I call the “android” mentality and the human is that the latter passed through something the former did not, or at least passed through it and responded differently – *changed*, altered, what it did and hence what it was; it *became*. I sense the android repeating over and over again some limited reflex gesture, like an insect raising its wings threateningly over and over again, or emitting a bad smell. Its one defense or response works, or its doesn’t. But, caught in sudden trouble, the organism that is made more human, that becomes precisely at that moment human, wrestles deep within itself and out of itself to find one response after another as each fails. (“The Android and the Human,” 151)

The human is mutability, variability, and affect, while the android is repetition and subsequently predictability. This affect, which signals the non-android in Dick’s work, is of a different order from the affect-emotion in Tomkins’ and Breazeal’s work. When Tomkins states that affect’s amplification is what makes things “matter,” he does so from a distinctly subject-oriented perspective; affect makes things matter [*to us*]. Tomkins’ affect, and the subsequent amplification, conceptualizes the interrelation, this mattering, from the perspective of individual subject, from the perspective of the human. Thus, through affect-emotion, we can only understand the human, the not-human, indeed the world, from the centrality of the human subject position.

Affect-emotion connects, anthropomorphically, the human and robot; the order of affect-emotion brings the human and robot into relationality with each other. Affect-

force, on the other hand, attends to the connectedness of the human and android, thus approaching the human and android less by the specific mechanics of the anthropomorphic encounter (indeed, Dick reduces the possibility of this kind of anthropomorphization by foregrounding identicality), but rather by the vectors of interrelationality as already existing conditions of possibility for human-robot intimacy. Read through affect, this intimacy can be said to exist prior to the actual encounter. Or, to push this reading further, this intimacy in fact calls the human and the robot into being. To move from affect-emotion to affect in this way is to move from positions and linear causality to multiple vectors of force.⁴⁶ Indeed, to read this interrelation through affect is to discuss human and android outside of the tautology of anthropomorphization, outside the tautology of the human itself. Thinking about the human from the perspective of affect allows us to think about the interrelation of human with machines, animals, objects, while being less beholden to the theoretical tautologies that haunt and limit explorations of the human.⁴⁷ Thinking anthropomorphization through affect does not eliminate the subject-position from the human, but rather insists on the simultaneous existence of other

⁴⁶ As Michael Hardt articulates in “What Affects Are Good For,” to think about affects is to re-think causality as multiple (ix).

⁴⁷ Tomkins makes this theoretical contraction around the human quite legible by locating his work on affect within what he calls “human being theory.” “Human being theory is a part of what used to be called general psychology. I have relabeled it because general psychology has been abandoned as a consequence of increasing specialization and because I wish to provide a theory for understanding human beings rather than a more general theory for understanding all animals” (*Vol. IV*, 1). Within the context of Tomkins’ desire to move away from species-generalality, his work in fact operates by way of enclosure, by limiting the scope of experience and the world. Tomkins’ “script theory,” a primary organizing principle of human being theory, similarly expresses this contraction. “Scripts are not simply actions or thoughts or memories or percepts or feelings or drives but the rules that generate organized scenes made up of these component functions, their processes, and their products. Through his scripts a human being experiences the world in organized scenes, some close to, some remote from, the heart’s desire” (*Vol. IV*, 9). Script theory is less a

positions.⁴⁸ When thought through this understanding of affect, anthropomorphization is only one of many perspectives and forces at work, and its evocation does not raze or subsume all other perspectives, but rather calls up these other concurrent perspectives.

Dick's theorization of the android, in its human identity, posits anthropomorphization as a kind of affectivity that, in the right context, reveals and pushes against the limits of the tautological human. In Dick's fictions, indistinguishability of appearance is the normative state, and de-anthropomorphization the disarticulating project. Dick, in positing the human not-human (human as android) and the not-human human (android as human), suggests a radical relationship to anthropomorphization that defamiliarizes the human itself. We might understand Dick as approaching anthropomorphization through hyper-saturation – if everything has some claim to “human,” then the terrain is leveled, and new conceptions of the human, as well as new theoretical orders in which the human ceases to exist as a privileged category, can potentially be articulated.

description of how a human being makes sense of his or her experiences and the world, but rather a reification or codification of these very experiences by way of scripts, of “rules.” Interestingly, Tomkins’ work references Charles Darwin’s *The Expression of the Emotions in Man and Animals*, first published in 1872. Darwin’s work attends to the different ways in which humans and animals express emotions; however, he discusses the specific emotions themselves as shared across species. In fact, Darwin suggests that animals can at times more effectively express emotions than humans: “man himself cannot express love and humility by external signs so plainly as does a dog, when with drooping ears, hanging lips, flexuous body, and wagging tail, he meets his beloved master” (“Introduction to the First Edition,” *The Expression of the Emotions in Man and Animals*, 18).

⁴⁸ One wonders if the prefix “not-” exists in the same way in the order of affect. If the human is only one of many positionalities, rather than the central, dominant positionality, the human is “not-” just as much as that which is not human. One might posit that there is no “not” in affect; there is just “is.”

3.5.1 Life, Simulacrum, and De-Anthropomorphization in Dick's *We Can Build You*

Louis Rosen and his partner, Maury Rock, run an electronic organ company that, at the opening of *We Can Build You*, is suffering at the hands of their competitors' hypothalamus-stimulation mood organs.⁴⁹ Maury surprises Louis by unveiling a prototype for a new product – a simulacrum of Edwin M. Stanton, Abraham Lincoln's advisor.⁵⁰ The Stanton simulacrum is so convincingly human that Louis confesses, "If I hadn't seen it spring to life I would believe myself it was just a sour elderly gentleman in old-style clothes and a split white beard, brushing itself off with an attitude of outrage" (*Build You*, 13-4). It is only in witnessing the Stanton "spring to life" that Louis recognizes the Stanton as a simulacrum, as something other than human. In the novel, human-identity introduces anthropomorphization, while the activation and deactivation of life – in simulacra and, as we will see later in the novel, in humans – initiate a de-anthropomorphization that is crucial to Dick's re-calibration of human and not-human. Upon first meeting the Stanton, Jerome Rosen, Louis' father, is similarly not convinced the Stanton is not a living breathing human – that is, until Maury deactivates the Stanton. "'Glop,' the Stanton said, and then became rigid, as lifeless as a window-store dummy; the light in its eyes expired, its arms paused and stiffened. It was graphic, and I glanced to see how my dad was taking it" (18). Both son and father need to see the

⁴⁹ We will encounter the mood organ again in the 1968 *Do Androids Dream of Electric Sheep?*. Likewise, the human simulacra machines that Dick introduces in *We Can Build You*, first published in 1972, but written in 1962, evolve into the androids of *Do Androids Dream*.

⁵⁰ Fredric Jameson, who describes the Stanton and the Lincoln simulacra as "perhaps the most sublime of Dick's characters" cites *We Can Build You* as introducing the figure of the android into culture (Jameson, 375).

activation (its springing to life) or deactivation (its lifeless-ness) of the simulacrum, as proof of its non-humanness. The opposite of life, it seems, is not death, but rather the external manipulation of life – the simulacrum.

Simulacra, while not murderous like the androids in *Do Androids Dream*, nonetheless posit the mutual exclusivity of existence between human and simulacrum. For example, Louis notices that Jerome's vitality has been diminishing since meeting the Stanton simulacrum: "Since the day [Jerome] had set eyes on the Stanton – and found out it was a machine built to resemble a man – he had become progressively more feeble" (48). For Louis, the life-force of the Stanton simulacrum has a direct negative correlation to the life-force of Jerome. This perception of threat speaks to the ways in which the presence of the simulacrum signals a deep disruption; Louis' perception of Jerome's diminished vitality can be understood as indicating that the previously anthropocentric world, with the introduction of the simulacrum, must be significantly reconfigured outside of its anthropocentrism.

The existence of the simulacrum introduces the necessity for an entirely new relationship to and understanding of the world. To this end, Louis, after meeting the Stanton, begins to question whether he himself is a simulacrum. This question begins as a ploy to raise suspicion as to Pris' mental health. Pris, Maury's daughter and the designer of the Stanton simulacrum, was previously treated for schizophrenia and had only recently been released from the care of the Federal Bureau of Mental Health. Louis makes an appointment with Pris' psychiatrist, Dr. Horstowski, and, in an attempt to convince Dr. Horstowski that Pris is not well, Louis tells the doctor that he himself is a

simulacrum: “Pris is playing a cruel prank on you. She sent me in here. I’m a simulacrum, like the Stanton. I wasn’t supposed to give the show away, but I can’t go on with it any longer. I’m just a machine, made out of circuits and relay switches” (50-1).

And again, as he is leaving,

“I was not kidding when I told you I’m one of Pris’s simulacra. There used to be a Louis Rosen, but no more. Now there’s only me. And if anything happens to me, Pris and Maury have the instructional tapes to create another. Pris makes the body out of bathroom tile. It’s pretty good, isn’t it? It fooled you and my brother Chester and almost my father.”

[...]

But you, I said to myself. You’ll never guess, Doctor Horstowski, not in a million years. I’m good enough to fool you and all the rest of them like you. (54)

During his first attempt to convince Doctor Horstowski that he is a simulacrum, Louis does not believe what he was saying. However, in this latter outburst to Dr. Horstowski, the idea that he is a simulacrum begins to take hold of Louis. And this idea stays with Louis after he leaves Doctor Horstowski’s office: “After having told Doctor Horstowski that I was a simulacrum I could not get the idea out of my mind. Once there had been a real Louis Rosen but now he was gone and I stood in his spot, fooling almost everyone, including myself...” (55). For Louis, even the existences of himself and his imaginary simulacrum are mutually exclusive – it is either him or not-him. And once he imagines that he is no longer Louis Rosen, but rather simulacrum, Louis cannot stop thinking about this possibility.

Fredric Jameson identifies this questioning, which is a common theme in Dick’s fictions, as the “android cogito” (“I think, therefore I am an android”), and describes it as a “peculiarly virulent and modern” iteration of Cartesian skepticism that “reverses the external issue of testing into a permanent rift within self-consciousness itself” (Jameson,

374). The android cogito turns the question, “are you human or android?” in on the inquiring subject, so that the question itself is embedded in the subject who seeks to distinguish. In *We Can Build You*, the android cogito’s turn inward is temporally inflected. While Louis begins to question if he is in fact a simulacrum, he does not presume that he was a simulacrum all along. Rather, each articulation of the android cogito asserts the previous presence of a non-simulacrum Louis Rosen: “there used to be a Louis Rosen, but no more” (54); “Once there had been a real Louis Rosen but now he was gone” (55); “I claim there is no Edwin M. Stanton or Louis Rosen anymore. There was once, but they’re dead. We’re machines.” (57). In *We Can Build You*, the android cogito introduces ontological difference as the possibility of displacement, as temporal progression, not as the absolute obliteration of all things human.

Pris similarly articulates this human-simulacrum mutual exclusivity, though she does so from a different conception of life and temporality; for Pris, the simulacrum is not a replacement of the human, but rather a resurrection of a dead human: “Do you think someday somebody will make a simulacrum of you and me? And we’ll have to come back to life?” (66). Pris understands the simulacrum as a continuation of life beyond death, an extension of the living self rather than a separate entity. While for Louis, the question of whether he is a simulacrum is a question of the present and the past, for Pris the question of the simulacrum is a question of the future: ““There we’ll be, dead and oblivious to everything... and then we’ll feel something stirring. Maybe see a snatch of light. And then it’ll all come flooding in on us, reality once more. We’ll be helpless to stop the process, we’ll have to come back. Resurrected!’ She shuddered” (66). Pris does

not understand the simulacrum as an entity ontologically discrete from the human, but rather as on the same continuum as the human, distinguished from the human temporally, through death.

We can understand the distinction between Louis' and Pris' conceptions of the simulacrum through Hayles' theorization of the schizophrenic in Dick's fictions. According to Hayles, the schizophrenic and the android are closely linked as boundary-destabilizing figures: "The inside/outside confusion links the schizophrenic to the android. Like the schizophrenic, the android is a hybrid figure – part human, part machine –whose very existence calls boundaries into question" (*Posthuman*, 177). Pris, a schizophrenic, is already hybridized, already part machine. Thus, the android cogito does not emerge as "am I human or simulacrum?," as for Pris "human" and "simulacrum" are not joined by mutual exclusivity. As schizophrenic, Pris is already both. For Pris, then, the simulacrum question is not about "what am I," but rather "when am I?"⁵¹

In contrast to Pris' idea of simulacrum as resurrection, for Louis, after the initial moment of disarticulation, the simulacrum exists as other. This alterity, which eludes Pris, allows Louis to engage the Lincoln simulacrum with compassion and sympathy. In the world of *We Can Build You*, simulacra are indistinguishable from humans, and anthropomorphization presumed. The project is to disarticulate simulacrum from human,

⁵¹ Hayles introduces what she calls the "schizoid android" as a way of understanding the extensive reach of the schizoid's and the android's boundary-destabilizations in Dick's works. She writes, "At the center of this extraordinarily complex traffic between cultural, scientific, and psychological implications of cybernetics stands what I will call the 'schizoid android,' a multiple pun that hints at the splittings, combinations, and recombinations through which Dick's writing performs these complexities" (*Posthuman*, 161). In the progression of Pris from schizophrenic in *We Can Build You*, to twinned androids, in *Do Androids Dream of Electric Sheep?*, we can understand Pris herself as embodying the terrain on which many of these complexities are navigated and negotiated.

to *de*-anthropomorphize, in order to establish new, more compassionate relationships between simulacra and humans. And, while these newly articulated conceptions of human and not human may continue to operate according to anthropomorphic forces, recalibrating these categories signals their fluidity. For, the theoretical project is not to think or advocate for a world without anthropomorphization, whether such a thing is possible, but to understand anthropomorphization as only one of many positions and forces at work in the world, and thus fluid in its capacity to be affected.

3.5.2 Sympathy and the Simulacrum

Sympathy: The quality or state of being thus affected by the suffering or sorrow of another; a feeling of compassion or commiseration.⁵²

The sympathy in *We Can Build You* anticipates the empathy that Dick locates as central to the human in his 1975 talk, “Man, Android and Machine,” and in *Do Androids Dream of Electric Sheep?*. “Empathy” appears only once in *We Can Build You*, in a description of Lincoln: “Lincoln had taken everything hard. He might have been remote, but he was not dead emotionally; quite the contrary. So he was the opposite of Pris, of the cold schizoid type. Grief, emotional empathy, were written on his face. He fully felt the sorrows of the war, every single death” (182). In this discussion of Lincoln’s empathy, it is unclear whether Louis is describing Lincoln the historical figure or the Lincoln simulacrum. This distinction, this blurring of human and simulacrum, is less

⁵² “Sympathy,” *Oxford English Dictionary online* (01.06.10). The *OED* posits multiple definitions of “sympathy,” including one that is more in line with what we now understand as “empathy.” For reasons that will become evident in this section, I offer the above quoted definition of “sympathy” as closest to Dick’s usage in *We Can Build You*.

relevant than the distinction drawn between the empathy of the two Lincolns and the android schizoidism of Pris.

“It groaned. Something about the noise made me shiver. Turning to Bob Bundy I said, “What do you think? Is it okay? It’s not suffering, is it?” (72). Rather than experiencing something akin to the uncanny, Louis, upon witnessing the Lincoln’s somewhat laborious awakening, is struck with a sense of compassion and concern, of sympathy, for the Lincoln. Louis’ sympathy for the Lincoln, while a synthesizing force, is premised on an alterity that was foregrounded through Louis’ encounter with the Stanton. As hybrid, Pris does not conceive of simulacra as discrete from her. Following from Hayles’ reading of the schizophrenic as hybrid, “part human, part machine,” if we understand distance and alterity as a condition for intersubjectivity and relationality, we can describe Pris’ distinctly uncompassionate and unsympathetic relationships with the simulacra, as well as with humans, as an index not of coldness and radical alterity, but rather of too much proximity.

This moment of Louis’ compassion, upon witnessing the Lincoln’s laborious awakening, further marks the otherness of the Lincoln for Louis.

We were, beyond doubt, watching a living creature being born. It now had begun to take note of us; its eyes, jet black, moved up and down, from side to side, taking us all in, the vision of us. In the eyes no emotion showed, only pure perception of us. Wariness beyond the capacity of man to imagine. The cunning of a life form from beyond the lip of our universe, from another land entirely. (72)

The Lincoln is, “beyond doubt,” alive. However, it is not human – it is an entirely different form of life. The Lincoln, despite its supremely human appearance, is unquestionably other. And it is this other for whom Louis feels sympathy. This

sympathy, which the Lincoln will reciprocate, is premised on the otherness of the Lincoln (and later, of Louis for the Lincoln); and, it is this relationship of sympathy that reframes both Louis and the Lincoln within the same category of the human. The human, then, is reorganized through alterity. Despite the appearance-identity of the Lincoln, Louis considers the Lincoln to be “from beyond the lip of our universe, from another land entirely.” The Lincoln is de-anthropomorphized in the scene of his awakening, only to become progressively human through compassionate and sympathetic behavior.

Upon first exchanging words with the Lincoln, Louis passes out quite violently; his world explodes:

Glancing up, the Lincoln put aside its quill pen and said in a rather high-pitched, pleasant voice, “Good afternoon. I take it you are Mr. Louis Rosen.”

“Yes sir,” I said.

And then the room blew up in my face. The rolltop desk flew into a million pieces; they burst up at me, flying slowly, and I shut my eyes and fell forward, flat on the floor; I did not even put out my hands. I felt it hit me; I smashed into bits against it, and darkness covered me up.

I had fainted. It was too much for me. I had passed out cold. (Dick, 79)

The explosion of Louis’ world recalls Louis’ previous engagement with the android cogito. In the Lincoln, Louis is confronted with the externalization of his earlier doubts about his own possible status as android. The world is now recalibrated, so that the android cogito does not take shape solely in Louis’ mind, or in archetypal scenarios of interiority, such as Dr. Horstowski’s psychiatrist’s office, but rather in the world itself.⁵³ Why now? Why does Louis’ world recalibrate after meeting the Lincoln, but not the Stanton, who seemed, to Louis, more human than Pris (29-30), and indeed more human

⁵³ Notably, after this dramatic encounter with the Lincoln, the idea that he is a simulacrum no longer preoccupies Louis.

than everyone but his father (60)? The difference is the initial stirring of sympathy that Louis feels for the Lincoln. The Stanton, though shocking in its physical and conversational identity to humans, remains other for Louis. Through sympathy, the Lincoln's alterity becomes proximity, thus altering the terrain of Louis' existence, including the entirety of the world.

The Lincoln, arguably the novel's most compassionate character, human or otherwise, sends Louis "a short note of sympathy" (92) after Louis faints. As the novel progresses, Louis and the Lincoln's bond of sympathy progresses to identification. After reading several biographies of Lincoln, Louis concludes that, "Lincoln was exactly like me. I might have been reading my own biography, there in the library; psychologically we were as alike as two peas in a pod, and by understanding him I understood myself" (182). Just as the Lincoln previously reciprocated Louis' sympathy, the Lincoln also reciprocates Louis' identification: "The simulacrum reached out and patted me on the shoulder. 'I think there is a bond between us, Louis. You and I have much in common.' 'I know,' I said. 'We're alike.' We were both deeply moved" (186). From sympathy to identification, both are mutually reciprocated by Louis and the Lincoln simulacrum.

3.5.3 Affect-Emotion and the Limits of the Face

If we understand Dick's novel as operating according to a presumption of human-simulacrum identity and a subsequent de-anthropomorphic move, it should be noted that Dick, like Tomkins and Breazeal, employs certain aspects of affect-emotion to indicate humanness and not-humanness. For example, Sam Barrows, the

multimillionaire businessman whom Maury and Pris initially court to invest in their simulacra, is devoid of emotion expression in vocal prosody as well as in facial behavior.⁵⁴ The difference is that Barrows' stark lack of affect-emotion indicates not a lack of legibility, but a lack of emotionality that characterizes humanness for Dick: "[Barrow's] eyes like the dots stuck in a snowman's face: expressionless, tiny. No emotion there; only the lower half of the face seemed to be grinning" (26). Barrows, whom Pris admires precisely because he is not burdened by emotions and morals, is not troubled by the existence of the simulacra. Upon encountering the Stanton and the Lincoln, the world does not change for Barrows, who is decidedly *unaffected* by the simulacra and their claims to life. Nor does the novel suggest that the existence of the simulacra raises any questions or doubts about the nature of his own existence. Rather, Barrows views Lincoln purely as a potential commodity, as a commercial object. Barrows concludes an extended philosophical debate with the Lincoln, on the nature of man, machine, and animal, with "I know you're a machine; I don't care. All I care is whether you work or not" (108-9). Their philosophical debate affects the Lincoln profoundly; for Barrows, the debate is merely an opportunity for product testing.

Barrows, as we find out later in the novel, is quite unethical in his business dealings, from the deliberate mismanagement and neglect of the Green Peach Hat housing development, to his proposal to place simulacra families on his moon colonies in order to trick humans into believing that his colonies are richly populated with humans.

⁵⁴ For descriptions of Pris' flat affect see pages 24, 29-30, 32. For descriptions of Sam Barrow's flat affect, see page 26.

Barrows' businesses traffic in appearance and deception; the simulacra, then, are nothing new to him. Or rather, the simulacra can be nothing new to Barrows, as he only views them as human-imitators. For Barrows, the simulacra remain within the realm of a uni-directional anthropomorphization, as objects that do not threaten or challenge the nature or categorical privilege of the human.

Jameson argues that Dick's androids, in their human-identity and emotional capacities, mark a shift from the labor-oriented robots of Isaac Asimov's fictional works (373). Following from this observation, Barrows' conception of the simulacra is more aligned with Asimov's robots than with Dick's androids. Further articulating this distinction, Barrows, under the guise of fair labor practices and peonage laws, appeals to the Stanton to join him in Seattle. Barrows offers to pay the Stanton "six dollars an hour," a "more than fair wage" (133). Barrows' effort is stymied by the Lincoln's outmaneuvering, which the Lincoln provides upon Louis' request: Lincoln advises Louis and Maury to sell MASA to Jerome Rosen for one dollar, and offer the Stanton the Chairmanship of the Board of Directors of the resulting merger between the Rosen spinet and organ factory and MASA (134-7). These two different appeals to the Stanton represent the two different fictional robotic approaches: while Barrows views the Stanton through the dehumanizing lenses of commerce and capital, Louis and Maury approach both the Stanton and the Lincoln as agential beings.

In contrast to Barrows, the Lincoln is often described according to the emotions conveyed through his face.⁵⁵ Upon first meeting Barrows, “the Lincoln regarded him with a melancholy expression. I had never seen such despair on a face before [...]” (102). And later, when Pris brings up the debate between Lincoln and Barrows, the Lincoln “did not respond, but its smile seemed – to me – to become even sadder, and its face longer and more lined with care” (118). The richness of the Lincoln’s emotional life is played out on his face.

3.5.4 Alterity, Fugues, and Agency

Pris, though described as cold and inhuman in both her affect-emotion and behavior, is certainly affected by the simulacra. As schizoid, Pris facilitates the mutual affectivity between the human and android by mediating between the two. Hayles discusses the “schizoid android” as a hybrid figure that operates multiply in Dick’s work:

In one of its guises, then, the schizoid android represents the coming together of a person who acts like a machine with a literal interpretation of that person as a machine. In other instances, however, the android is placed in opposition to the schizoid. [...] The android is not so much a fixed symbol, then, as a signifier that enacts as well as connotes the schizoid, splitting into the two opposed and mutually exclusive subject positions of the human and the not-human. (161-2)

In *We Can Build You*, the android’s unfixity is transformative partly on account of the presence of the schizoid. Louis and the Lincoln’s mutual sympathy emerges, and subsequently recalibrates the nature of humanness, in part because the more recognizable, oppositional dynamic between human and not-human is preserved in Louis’ relationship

⁵⁵ Similar descriptions of the Lincoln’s facial affect-emotion occur on pages 191 and 197.

with Pris. Illustrating this triangulation, Louis' relationships with the Lincoln and with Pris as demonstrating two different forms of alterity at work in the novel. Louis' relationship with the Lincoln suggests an alterity that allows for encounter, for identification and sympathy, while Louis' relationship with Pris, on the other hand, suggests a kind of radical otherness that produces alienation rather than proximity and encounter. "I had my own personal experience with [Lincoln] – or to be more exact, with his simulacrum. I didn't catch the *alienness*, the otherness, with the simulacrum that I had caught with Pris" (182). And later, upon witnessing Pris destroy the sloppy John Wilkes Booth simulacrum created by Barrows, Louis thinks to himself, "I have seen into the *other*, I said to myself, when I saw Pris" (202). With Pris there is no bond, no reciprocation of feeling; rather, her radical otherness enables and shores up Louis' reciprocal bond of sympathetic identification with the Lincoln.

Having failed to bring Pris back home, much less win her love, Louis is diagnosed with schizophrenia, and subsequently becomes a patient of the Kasanin Clinic. It is only through a series of drug-induced fugue states that Louis, now a patient, finally encounters Pris outside of a prohibitive radical otherness. Of course, Pris is arguably never more absent than in these fugue encounters, during which Louis hallucinates an entire life with Pris in which they "court, marry, and have a child" (*Posthuman*, 172).⁵⁶ And yet Pris is also physically proximate, as Louis discovers when, outside of any hallucinatory state, he runs into her in the halls of the Kasanin. Pris, in this state of physical proximity, encourages Louis to end the fugue states and to work toward leaving the clinic. "If I do,"

⁵⁶ This shared life takes place across upwards of two hundred and twenty fugues (*Build You*, 241).

Louis asks Pris, ““will you marry me?” She groaned. ‘Sure, Louis. Anything you want. Marriage, living in sin, incidental screwing – you name it’” (*Build You*, 242). While Pris’ intentions here are ambiguous at best, as Hayles points out (*Posthuman*, 172), Louis takes her at her word and secures his release, fully expecting his fugue-life with Pris to be mirrored and extended outside of the clinic. As Louis is leaving, he finds out that Pris did not apply for release, and is in fact staying in the Clinic. Perhaps this is Pris’ noble gesture, giving Louis an opportunity to live in the world; or perhaps she is simply being cruel: “does she trick Louis because she doesn’t want him around or because she wants him to get on with his life?” (*Posthuman*, 172). Or, alongside this ambiguity, perhaps this is Pris’ way to escape Louis’ fugue-life, in which she has no agency or presence. For Pris, who is centrally concerned with the fear that an extended life will be imposed on her by external means (recall that this fear of an externally extended life organizes her initial misgivings about the simulacra), the idea of her presence in Louis’ hallucinations must be deeply disturbing. Perhaps, in addition to being motivated by nobility or cruelty, Pris’ last gesture is her way of taking a high heel to Louis’ fantasy life with her, seeking to destroy that which threatens to bring her, against her will, into another state of life.

3.5.5 Initial Dis-articulation in *Do Androids Dream of Electric Sheep?*

“But you see,” Pris said, “if you’re not human, then it’s all different.” (*Do Androids*, 161)

In *Do Androids Dream*, as in *We Can Build You*, de-anthropomorphization, or disarticulation between human and android, centrally organizes the narrative, as well as

Dick's larger aims of destabilizing and multiply reconstituting what it means to be human. De-anthropomorphization does not just alter the human, but in fact the world that, when no longer experienced and inhabited through a deeply entrenched and familiar anthropocentrism, is "all different." This altered new world is, in a sense, Louis Rosen's post-black out world, in which humans and their mechanical imitations are distinguished not through biology, but rather through behavior. In this new world, androids, which are identical to humans in appearance, are as intelligent or more intelligent than humans. Intelligence is no longer the metric by which human and android can be distinguished, and empathy is now the characteristic by which human and not-human are identified and differentiated.⁵⁷ In *Do Androids Dream*, the quality of empathy itself is also destabilized and ultimately expanded. After briefly discussing a central and founding disarticulating scene in the novel, I will discuss three different forms of empathy at work in the text. The dynamics between these three modes of empathy, and the various reconfigurations of species that occupy the subject- and object-positions of these empathies, destabilize and recalibrate the human, the android, and indeed the world in *Do Androids Dream of Electric Sheep?*.

⁵⁷ This shift from intelligence to emotion (or, for Dick, emotion's close relation, empathy) demonstrates the "sliding divider," between human and not-human, which I earlier discussed in relation to the Turing Test. This metric shift, both in robotics and AI technologies as well as in the cultural imaginary maintains an essential distinction between human and not-human. Once machines are possibly considered equally or more intelligent than humans, emotion emerges as the new lack that differentiates machines from humans. This sliding divider between human and machine can also be read generously: as taking on a new challenge (of programming emotion) in response to the developments and achievements in programmable intelligence (Brooks and Stein, "Building Bodies for Brains"). Of course, neither reading sufficiently encompasses the complexities by which such a shift in metrics emerges (for example, we cannot ignore the discussions around emotional intelligence that emerged in the mid-twentieth century and continue to occupy a significant space in psychology, as well as in popular culture). Nor is it my aim to identify any kind of definitive causality of this emergence. My aim is to attend to the deep instability of the category of the human by highlighting these complexities.

Rick Deckard, an android bounty hunter, is tasked to “retire” a number of Nexus-6 androids that escaped Mars and arrived on Earth. Rick goes to the headquarters of Rosen Associates, the manufacturers of the highly intelligent, highly dangerous Nexus-6 androids. There Rick consults with Eldon Rosen and his niece, Rachael Rosen, with an eye toward getting a Nexus-6’s baseline reading for the android-detecting Voigt-Kampff Empathy test. Rachael offers herself as a test subject: “Give me the test” (*Do Androids*, 46). The test concludes that Rachael Rosen, a fictional descendent of *We Can Build You*’s Pris, is an android.⁵⁸ The Rosens briefly convince Rick that Rachael is not an android, that the test is flawed. Rachael reveals herself to be an android when she refers to an owl as an “it”: “It, [Rick] thought. *She keeps calling the owl it.* Not her” (*Do Androids*, 58). Though it turns out that the owl in question is in fact mechanical, Rachael’s slip is verboten within a society that views empathy toward animals, mechanical or otherwise, as a primary index of humanness. This disarticulation, this founding de-anthropomorphization, signals the ways in which empathy, throughout the novel, is crucially triangulated across human, animal, and machine species.

3.5.6 *Do Androids Dream of Electric Sheep?* and the Construction of Empathy

Empathy: “The power of projecting one’s personality into (and so fully comprehending) the object of contemplation.” (*OED* online)

⁵⁸ We will later learn that Rachael has known that she is an android for at least two years, during which time she has been trying to protect the escaped Nexus-6s by sabotaging android bounty hunters (*Do Androids*, 198-200).

The Voigt-Kampff Empathy Test, which externalizes and institutionalizes the skepticism of the android cogito, is based on the tenets of Mercerism, which emphasizes empathy above all else. With the aid of empathy boxes, followers of Wilbur Mercer are able to fuse with other followers, hearing each others' thoughts as they all physically merge with Wilbur Mercer, joining him in corporeal experience as he walks up an endless hill (*Do Androids*, 22). Followers need only place both hands on the handles of the empathy box, and they then physically and mentally enter this state of fusion. The physical merger that accompanies this fusion is less an index of intensity, but rather a necessary fortification of the experience of empathy-fusion. This fusion is activated through embodied action – the hands on the handles – and heightened through embodied presence, in physical fusion with Wilbur Mercer.

Accessing Mercerism's empathy takes little by way of effort, moral or otherwise. One simply places one's hands on the handles of the empathy box, and one "has crossed over in the usual perplexing fashion; physical merging – accompanied by mental and spiritual identification – with Wilbur Mercer had reoccurred. As it did for everyone who at this moment clutched the handles, either here on Earth or on one of the colony planets" (22). In Mercerism, fusion generates empathy; empathy does not generate fusion. And, while this causal distinction may be too fine a point to discuss at length, I bring it up to return to empathy's status as the sole province and capacity of the human. Literary critic Jill Galvan points to the novel's construction of empathy through the Voigt-Kampff test, "which, because it measures empathy, Rick uses to ensure the android identity of his potential prey – throws into relief the contrived nature of this putatively most basic of

human qualities” (Galvan, 415). In other words, empathy is contrived not just as a distinguishing criterion, but as a quality in itself.⁵⁹

Dick’s disarticulation of behavior from appearance, which I discussed in relation to *We Can Build You*, as well as in relation to the Voigt-Kampff test, also emerges in an extension of Mercerism’s empathy: caring for animals. Mercerism posits caring for animals as a virtual moral imperative. In this post-nuclear world that is slowly being taken over by life-destroying dust (Kipple), healthy living animals are increasingly rare, and thus increasingly more expensive. Rick, in a conversation with his neighbor Bill, reveals that the sheep he and his wife, Iran, have been tending, is in fact mechanical. While it is not uncommon for people to own and care for mechanical animals, considering the price and scarcity of biological animals, it is not often disclosed when an animal is mechanical, as it stigmatizes the owner as non-empathetic. Not owning a real animal wears on Rick, “gradually demoralizing” him (9). Iran, on the other hand, “did care [for the mechanical sheep]. Very much.” (9). Iran’s care for mechanical animals extends to her cross-species empathy for androids. In the opening pages of the novel, Iran calls Rick “a murderer” for retiring androids: “You’re a murderer hired by the cops,” she accuses him. To which he responds, “I’ve never killed a human being in my life” (4). Iran’s empathy for animals and androids is not limited to the fusion with other humans, but in fact expands across species and organicism.

⁵⁹ Christopher Palmer similarly points to the reliance on the technology of the Voigt-Kampff test to articulate, if not create, the difference between human and android (62).

In his pursuit of the Nexus-6 androids, Rick teams up with Phil Resch, a fellow android bounty hunter. Upon meeting Phil Resch, Rick experiences a subsequent destabilizing encounter with the android cogito. Rick first meets Phil through Inspector Garland, an android masquerading as a police officer, in what turns out to be a fake police station populated entirely by androids. The externalized android cogito of *Do Androids Dream* is already at work, as android accusations fly amongst the three of them, culminating in Rick identifying Garland as an android, and Garland deliberately misidentifying Resch as an android. Rick does not realize until later that Resch is in fact a human. Within this temporal space of misidentification, and perhaps somewhat enabled by it, Rick grows to feel non-empathetic toward Resch, disgusted by his coldness and heartlessness.

Rick's encounter with Resch also marks one of several shifts from the term "retire" to "kill." Earlier in the novel, Rick "retires" Max Polokov, the wholly unsympathetic android that injured Rick's bounty hunter predecessor (*Do Androids*, 91-4). However, "kill" begins to displace "retire" when Rick encounters Resch, first as human, then as android, then as human.⁶⁰ Hayles importantly notes that Rick's increasing empathy for androids is not unproblematically connected to gender, through Rick's sexual attraction to Rachael Rosen. Hayles also points to Rick's momentary shift back to "retire" after Rachael agrees to retire the Pris Stratton android if Rick has sex

⁶⁰ Resch also speaks of "killing" androids. "'You're claiming Garland?' Phil Resch asked. 'But I killed him, not you. You just lay there. And Luba, too. I got her.'" (135) While for Rick, the use of "kill" indicates an expansion of empathy toward androids, for Resch, the use of "kill" is less clearly delineated as a metric of species or empathy. Resch, not not-human and not not-android, does not seem to have empathy for anyone or anything. For Resch, unlike for Iran, Rick, and John, the word "kill" does not evoke empathy, but rather the inverse.

with her (172-3). It should be noted that both Rachael and Rick return to “kill” after Rachael reveals that she has slept with a number of bounty hunters, all of whom can no longer kill androids afterward. Recalling Iran and Rick’s earlier semantic squabble, during which Rick sets up the linguistic divisions between killing and retiring, human and android – “I’ve never killed a human being in my life” – this terminological shift from “retire” to “kill” criss-crosses the boundary between human and android, moving across the space of potential object of empathy.⁶¹ Phil is the only bounty hunter exception, until he is joined by Rick, who hunts down the remaining Nexus-6s in John’s apartment building, and then kills first Pris, then Irmgard, then Roy. This scene is markedly devoid of either “kill” or “retire” (221-4). As these terms have been carefully selected throughout the novel as a metric for Rick’s empathy, we can understand the absence of these terms in the aforementioned scene as speaking to Rick’s deep ambiguity and confusion of these destabilized categories.

As Rick’s empathy expands to androids, it also begins to contract, excluding Phil Resch who, though nominally human, lacks not only the expansive empathy of Iran and John Isidore, but also the more limited empathy of Rick as well as the preservation-empathy of the android community. Before administering the Voigt-Kampff test to Phil Resch, Rick divulges, “I hope to god you do test out as an android. [...] I see a pattern. The way you killed Garland and then the way you killed Luba. You don’t kill the way I do; you don’t try to – Hell,’ he said. ‘I know what it is. You like to kill. All you need is

⁶¹ Whereas in *We Can Build You*, species boundaries are deterritorialized through concepts of life, in *Do Androids Dream*, these category boundaries are deterritorialized through Dickian life’s non-opposite, death.

a pretext. If you had a pretext, you'd kill me” (137). I understand Rick’s divulgence, together with the revelation that Phil Resch is not an android, as an additional disarticulating moment – an uncertainty and ambiguity in which the human and not-human are newly rearticulated along different metrics and genealogical criteria. Outside of Mercerism and the Voigt-Kampff test, Rick identifies Phil Resch, through his behavior to others, as a non-empathetic subject and object.

Once the Voigt-Kampff test identifies Phil Resch as human, Rick begins to wonder about his own relationship to empathy. Human or android – either way, “it’s all different” (161). As Phil Resch is in fact human and not android, as Rick hoped and believed, Rick must subsequently recalibrate his own feelings and assessments of Phil Resch. And the species question, which was previously exteriorized onto Phil Resch, is turned inward to Rick and his empathy for Luba, whom Phil Resch killed without hesitation and without emotion. “So I was wrong,” Rick thinks to himself. “There’s nothing unnatural or unhuman about Phil Resch’s reactions; *it’s me*” (142). Rick realizes that Resch is not an empathy-lacking android, but rather that Rick is an inversely empathetic human who feels empathy for the android, Luba, and not the human, Phil Resch (143). Rick’s expansion of empathy outside of Mercerism seems to hold a certain promise in its de-anthropocentric humanity; however, this empathetic expansion is significantly impaired once the Voigt-Kampff test concludes that Phil Resch is human. In these brief yet periodic moments of species-uncertainty, de-anthropomorphization shifts the ground of the human. In the text, these moments exist only as glimpses, and are countered and often foreclosed by the definitional and categorical forces of Mercerism

and its technological accoutrements. But that they exist at all in the text's bleak and desperate world is not insignificant; nor are their destabilizing effects fleeting and unaffecting.

Mercerism characterizes empathy as exclusively human; androids, Mercerism asserts, cannot access the experience of fusion, and thus cannot understand what it means to exist within a group:

Empathy, evidently, existed only within the human community, whereas intelligence to some degree could be found throughout every phylum and order including arachnida. For one thing, the empathic faculty probably required an unimpaired group instinct; a solitary organism, such as a spider, would have no use for it; in fact it would tend to abort a spider's ability to survive. It would make him conscious of the desire to live on the part of his prey. [...] Evidently the humanoid robot constituted a solitary predator. (31)

While Mercerism understands androids as "solitary predator[s]," Jameson points out that the novel challenges this exclusion of androids from community, sociality, and empathy. Escaped androids Roy, Irmgard, and Pris, as well as Rachael, in her remote participation, form an undeniable community, with a "very real community of interests and feelings between the rebel androids and their palpable dismay at the extermination of their fellows."⁶² We can also understand the android police station, led by Inspector Gardner, as an interconnected branch of this community.⁶³

The predatory spider returns at the end of the novel and challenges the mutual exclusivity of predation and empathy. John Isidore, who is now aiding Pris, Roy, and

⁶² Jameson (373) and Galvan (414-5) also draw attention to the androids' community and compassionate bonds.

⁶³ After all, Luba Luft reached out to this android police station for help when confronted by Rick (*Do Androids*, 106).

Irmgard in their efforts to survive, comes across a living spider – a non-mechanical rarity (205). Pris proceeds to snip off five of the spider’s legs, sickening and upsetting J.R. in the process. “The empathic faculty probably required an unimpaired group instinct; a solitary organism, such as a spider, would have no use for it; in fact it would tend to abort a spider’s ability to survive.” (31) Certainly, this is not the case for this spider, who would have had great use for empathy in the android other. Empathy in the android other would not abort, but facilitate, the “spider’s ability to survive.” The three androids hiding in J.R.’s building find themselves in a similar reversal, from predator to prey, and would certainly welcome the empathetic faculty that might keep them from being killed. The spider and the androids both meet their demises shortly thereafter, though the spider’s death emerges from a deep empathy. J.R., unable to bear the spider’s suffering, takes the spider from Pris and drowns it (211), thus complicating the distinction between killing and empathy. Rick’s final showdown with the androids remains ambiguous. Did Rick “retire” the androids, or did he “kill” them?

3.5.7 Anticipation as Invention: Dave Hanson’s Philip K. Dick Android

Hanson, who has written of Dick’s early and extensive influence on his robotics work, identifies science fiction writers as not just predicting the future with remarkable accuracy, but in fact *creating* it in the very act of imagining.

The aging work of Asimov, Shelley, and Dick each point accurately to strange aspects of today’s emerging reality of bio-inspired robots. From where does such foresight come? Anticipation acts as a force of nature, unleashed in the mind of the science fiction author; science fiction authors act as ad hoc inventors, physicists, and philosophers. *The authors not only discover the future, as it were,*

but alternately – they – through the art of fiction – also invent the future in fits of acutely self-fulfilling prophecy. (“Humanizing,” 8, my emphasis)

Speaking to the ways that the cultural imaginary, here science fiction, shapes the world, Hanson points to Dick as not only engaging in this future creation-imagination, but also as posing an ethical imperative for robotics.

Finally, in the process of making the robot, we push forward the threshold of social AI, a path of development that promises to realize compassionate machines as an antidote for the sort of nightmare machines currently under development in laboratories around the world. PKD foresaw the imperative of such work in his writings. Recent technology trends underscore the urgency: The vast majority of current robotics and AI research is military in purpose and not sociable. (Hanson, 167)

Dick’s compassionate sociable machines resist the militarization of robotics and AI. In *We Can Build You* and *Do Androids Dream of Electric Sheep?*, as well as in Hanson’s robots, identity is an important means by which to generate compassionate behavior in both the machine, as well as in the human for the machine. Like Dick, who also grappled with this “nightmare” aspect of technology, Hanson takes recourse in identity, eschewing Mori’s warnings. And, like Dick, Hanson foregrounds identity in order to push appearance to the background, thus enabling him to also address other capacities and behaviors, such as compassion. Hanson points not only to the deep and obvious influence of Dick’s work, but also to the importance of highlighting the intersections between robotics and fiction, as a means to identify, diffuse, and resist the more terrifying uses of robotics technologies.

Dick’s influence on Hanson is materialized in Hanson’s Philip K. Dick (PKD) Android. Hanson’s aim was to create a social android who, in appearance and in knowledge, could recreate the experience of conversing with Dick. The incredibly life-

like android (see below figures) is capable of speech, facial expression, and facial recognition. Constructed using Hanson's patented skin-like substance, Frubber™, and equipped with natural language software, a "highly realistic" synthetic voice, and facial recognition software, the PKD Android can speak, produce life-like facial expressions, recognize and remember faces, as well as make eye contact (Hanson, 150, 169-71).⁶⁴

Hanson, who emphasizes the element of creativity in robot design, links this creativity to the robot's genealogy in the cultural imaginary:

In many ways, art and fiction routinely twist, challenge, and transform human identity. Therein, our culture—our software (memes)—are (re)invented, ostensibly propelled forward as an extension of evolution. Figurative arts redefine the human visual identity, while the robots of science fiction redefine the conceptual boundaries of the human being. As an artistic and narrative medium, however, the actual AI-driven robot both redefines our visual appearance and the conceptual framework of the human. As sculpture, the android challenges in ways like science fiction but with a verisimilitude only possible with physical embodiment. Because robots may be programmed to act much smarter than they actually are, robots are in effect a new fictional medium, physically embodied science fiction. ("Humanizing," 95-6)

Hanson orients robotics within the artistic and imaginative forces of culture, specifically, science fiction. Placing his android robots within this artistic genealogy, Hanson views robots themselves as "a new fictional medium" ("Humanizing," 96). Indeed, this was one of the aims in creating the PKD Android: "Bringing PKD back to life as an android extends the spirit of Dick's fiction into a new medium," (Hanson, 165). The android as medium. Hanson, recalling the creations of the Stanton and the Lincoln, uploaded the

⁶⁴ It should be noted that, like the natural language AIs that participate in Loebner competitions, the PKD's conversational fluidity was quite stilted and unconvincing. The PKD only recognized around sixty percent of what was spoken to it, and the PKD's lengthy responses were often experienced as off-putting ("Humanizing," 173).

PKD Android with both biographical knowledge of Dick and Dick’s writings. And, similarly to Dick’s simulacra and androids, Hanson’s PKD “highlights the issues with which Dick feverishly wrestled: What is human, how can we avoid destroying ourselves, what distinguishes compassion from cruelty (social intelligence? wisdom?), and into what kind of organism are we evolving?” (Hanson, 165) The question of our evolution is particularly striking when we consider that it is not just the PKD Android that emerged from the imaginative world of Dick’s musings, but Hanson himself who, having long been inspired and captivated by Dick’s work, is himself a kind of Dickian creation, brought to life through his own pursuit to create human-identical androids. Human and android are co-constituted through Dick’s fiction.



Figure 7: Hanson’s PKD android.

“Philip K. Dick robot.” Mirka23. Flickr. Creative Commons Attribution-Noncommercial 2.0 Generic.

3.5.8 Post-Script: An Expansion of the Cultural Imaginary

Just as Ridley Scott’s *Blade Runner* (1982) and Richard Linklater’s *A Scanner Darkly* (2006) have extended Dick’s cultural presence, Hanson’s PKD Android has similarly become an addendum to Dick’s fictional legacy. As is now well known in some robotics and science fiction circles, in 2005, Hanson, while on a flight from Dallas to Las Vegas, left the PKD’s head in the overhead storage bin. Subsequent efforts to locate the head have proved fruitless.⁶⁵

⁶⁵ For discussions of this imaginary coda, see the following newspaper articles: Sharon Waxman, “A Strange Loss of Face, More Than Embarrassing”; and Wendy M. Grossman, “Do Robots Dream of Copyright?”

4 Anthropomorphization Incorporated and Abjected: Stelarc's Extra Ear

My previous chapters discussed anthropomorphization of the human as mediated through humanoid machines and imaginaries. Through the theoretical concepts of abjection, the obsolescence, and sensory reorganization, and *Extra Ear*, an ongoing work by Australian performance artist Stelarc, this chapter attends to anthropomorphization as it operates directly on the human. This discussion of anthropomorphization operates across multiple theoretical registers – psychoanalysis, poststructuralist thought, and media theory. Though at times seemingly incompatible, these discourses provide ways to understand the human as resolutely and unavoidably porous. From Julia Kristeva's psychoanalytic theory of the abject, to Deleuze and Guattari reorganization of the body, to Marshall McLuhan's media obsolescence and sensory extension, these theories articulate the borders of the human as continually undone in its relation to the world, so much so that the boundaries of the human emerge as constituted only in their perpetual transgression. And while these theories operate through differently-inflected conceptions of the body, these theoretical registers consider the body as an essential aspect by which solipsism as a mode of human existence is an impossibility.

Stelarc's work thematizes the relationship between technology and the human body. Asking how technology and the body variously rely on, threaten, amplify, and define each other, Stelarc's works articulate a body that is deeply and intractably technological. From *Suspensions*, in which Stelarc hangs from the ceiling from hooks piercing his flesh; to *Third Arm*, in which a mechanical arm is attached to his body, not as

prosthetic but as expanded function; to *Exoskeleton*, a mechanical walking machine that encases the human body, Stelarc engages with what might be understood as the limits of the human body (for example, the skin), only to reveal these limits as the very conditions for the body's extension and expansion in technology.

Operating through the continual doubling-back movement characteristic of anthropomorphization, the object in *Extra Ear* folds Stelarc's body in on and out of itself; it abjects itself. Stelarc's ear, the object that activates this abjection, is not completely identifiable as other; rather, the ear references the subject of perception. While anthropomorphization functions as a founding metaphor in AI, the Turing test, and Powers' *Galatea 2.2*, anthropomorphization in *Extra Ear*, operates on the human metonymically. In conversation with Deleuze and Guattari on the reorganization of the body and Massumi on the obsolete body in Stelarc, I discuss this metonymic anthropomorphic object as continually disfiguring and refiguring the boundaries of the human. Within this conceptual framework, sense perception itself can be understood as a mode of perpetual and acceptable abjection.

I will conclude my discussion of Stelarc's *Extra Ear* by turning to Marshall McLuhan's theories on media technologies, the senses, visibility and invisibility, and the function of art as uniquely illuminating certain aspects of our relationship with technology. Just as literature functions in earlier chapters, not as a set of mere diagnostic examples of how certain theories operate, but rather as a significant contribution to the theorizations themselves, Stelarc's art *qua* art has the capacity to yield insights – the anthropomorphic object – we cannot access solely through other forms or discourses.

4.1 The Visibility of Obsolescence: Stelarc's Obsolete Body

The body is obsolete, Stelarc proclaims.¹ This claim has led some critics to misrecognize the body's persistent centrality in his performances. Fred Botting and Scott Wilson describe Stelarc's work as "us[ing] the meat to enter the matrix of technological and post-human becoming" (158). This characterization misunderstands the body in Stelarc's work. It is not the meat that allows Stelarc to access this "matrix of technological and post-human becoming"; rather, this matrix of technology and post-human becoming can only be played out on and through the meat. To say that Stelarc is an artist of technology is to say that he is an artist of the body.

Stelarc does not mourn the body in its obsolescence, but rather views it as an evolutionary necessity in our technological present, a desirable state for our body to inhabit – or rather, to un-inhabit.² In fact, Stelarc suggests that this corporeal obsolescence "might be the highest of human realizations" (Stelarc, "The Splitting," 134). But what is this obsolete body, and how is it that this obsolescence can be realized as a featured aspect of Stelarc's performances? According to performance art critic

¹ "The body has become obsolete and invaded by technology" (*Recent Projects*, 3).

² The evolution Stelarc discusses is not on the level of the species, but rather on the level of the individual. "We are not capable, nor should we try to engineer a total transformation of the human species – but we can modify chosen individuals. In this way, we may be able to trigger new evolutionary directions without the trauma of trying to mold whole populations with the possibility of a catastrophic failure" (Stelarc, "Strategies," 76). Jane Goodall opposes Stelarc's individual technological evolution to the species evolution of *Blade Runner* (1982), Ridley Scott's film adaptation of Dick's *Do Androids Dream of Electric Sheep?*, and Spielberg's *A.I. Artificial Intelligence* (2001), adapted from Brian Aldiss' story "Supertoys Last All Summer Long." While these fictions depict evolution through competition and species rivalry, Stelarc's evolution operates through the individual and his or her technological environment (1-2).

Edward Scheer, Stelarc's body involves a contradiction of presence: "The basic contradiction of all Stelarc's actions consists in their return to the image of the artist's body in a way that reinforces the effect of its presence and its adaptive capacities" (85). According to Scheer, Stelarc's performances highlight the presence and capacities of the body, thus contradicting the technological fantasies of disembodiment³ in Stelarc's own "rhetorical productions."⁴ By this reading, Scheer opposes the body's "adaptive capacities" with the hollowed out, obsolete body. For Stelarc, however, adaptation and obsolescence are not contradictory; instead, obsolescence is a *form* of adaptation, of evolution. Scheer rightly points out that in performances, Stelarc's body is undeniably present. However, this undeniably present body that does not recede in the face of technology is indeed the hollow, obsolete body that Stelarc invokes rhetorically. Stelarc's obsolete body is not disappeared or vanished, but rather an altered conception of corporeality – one that is conditioned on its co-existence with technology.

Art historian Amelia Jones highlights the affective impact of the body's acute and moving presence in Stelarc's works against his claims of the body's obsolescence. I quote Jones at length, describing her corporeal, emotional, imaginative, and cognitive responses to one of Stelarc's performances of *Extended Arm*:

Contrary to Stelarc's vociferous claims about his work extending the "obsolete" body of the global capitalist era [...], I found myself responding in a deeply empathetic way to the drone of the *Virtual Arm*'s technology and to this small,

³ "Stelarc's statements, seen as separate from his actions, seem to generate a proliferation of disembodied fantasies, as if we could leap from the body into virtual reality and never return" (Scheer, 85).

⁴ "But if we ignore the performances and focus on his rhetorical productions, presence becomes only a detour for Stelarc. He recedes from us into the hollow body, emptied of qualities associated with personality and affect" (85).

compact, masculine body partly sheathed in metal and speaking to the audience passionately about his work. Tears came into my eyes in the most emotional fashion, as I imagined (even empathetically experienced) my own body trapped, controlled, directed by this technological apparatus. Far from experiencing Stelarc's (or my own) body as "obsolete" or otherwise irrelevant or transcended, I felt *more* aware of my bodily attachment to his artistic practice – *more* and not less cathected to his technologized form. My response was immediate and identified deeply with him through an interwoven process of sensation, perception, cognition, and feeling.⁵

Stelarc's obsolete body is not contradictory to either the body's presence or the intensity of Jones' response to his performance.⁶ Rather, his obsolete body predicates the presence of the body as it is experienced and viewed in the moments of performance. It is precisely in the body's obsolescence that the acuteness of its presence and "affective materiality" can be experienced. This is the point of Stelarc's obsolete body – it speaks to the deep cathexis between technology and the body (Stelarc's, Jones', mine) through the embodied processes of "sensation, perception, cognition, and feeling." This body is made visible precisely in its obsolescence – this is the seeming contradiction of Stelarc's obsolete body that Scheer and Jones observe.

In an interview, Stelarc clarifies the non-contradictory relationship between presence and the obsolete body. Though the body may become, or is already, obsolete, it is not something we might ever escape:

What irritates me is people who see the Internet and virtual reality systems as strategies of escaping the body. You don't escape the body; you function differently with the interfaces that produce these immersive and interactive

⁵ Jones, 87. *Extended Arm*, a mechanical, pneumatic arm and hand are attached to Stelarc's right arm and operated by his right hand. *Extended Arm* evolved from an earlier work, *Virtual Hand*.

⁶ In response to the body's presence in Stelarc's performances, Jones asks "Why then does he continue to claim the obsolescence of the body – despite its obdurate refusal to disappear, its brute centrality in his performance practice, and its affective materiality?" (Jones, "Stelarc's," 87).

effects. You construct an extended operational system that functions beyond the biology of the body and beyond the local space it inhabits. (Zylinska and Hall, 122)

The body is dynamic; it “function[s] differently” with each technological engagement. This body in its dynamism is the body made obsolete through technology. And the body that is disappeared is not the obsolete body, but rather an Idea of the body that is prior to technology – an Idea that exists only as it is continually disproved through the obsolete body. In opposition to this Idea of the body, the obsolete body does not recede from the world; it is not self-contained and cut off from the world. Rather, the obsolete body – the body that combines with technology to “construct an extended operational system” beyond the biological body – is the body that is open to and in the world, and highlighted in Stelarc’s artistic performances. In Stelarc’s work, no less in his writings, the body’s obsolescence is the condition of both its engagement with technology and its openness to the world.

Further highlighting the body’s presence in technology, Stelarc warns that the sense of technological disembodiment is a kind of experiential *trompe l’oeil*: “I think we have to be careful about that idea of disembodiment because it is one of these experiences that technology in a sense promotes but in relation we are increasingly more grounded, pacified, and constrained by the vast technological terrain that is required” (Abrahamsson and Abrahamsson, 297). The sense of disembodiment in the face of technology, Stelarc asserts, is misleading. The body is not less present in technology, as the sense of disembodiment is in fact an indication of the extent to which the body is present in

technology.⁷ This is the obsolete body – the body that is reorganized and recalibrated in its relationship with technology. And this is what Stelarc’s works *do* – they engage with technology in such a way as to make visible the reorganized body and, consequently, the obsolete body.

Brian Massumi reads Stelarc’s body as “always-already obsolete”: “Isn’t change always inexorably under way? Then, in a very real sense the body is always-already obsolete, has been obsolete an infinite number of times and will be obsolete countless more – as many times as there are adaptations and inventions. *The body’s obsolescence is the condition of change. Its vitality is in obsolescence*” (*Parables*, 109). By Massumi’s theorization, it is the non-obsolete body that is disappeared, that does not exist except in ideation. Rather, Massumi suggests that the obsolete body is always already- and always ever- obsolete.

So, while obsolescence may be “the condition of change” by which new organizations of the body happen, change can also be understood as the condition for obsolescence, as it is these new organizations that make obsolete the older calibrations of the body. To clarify what might initially seem tautological, but is in fact the seeming paradox of the obsolete body, I look to Massumi’s discussion of the body through the twinned states of movement and sensation:

⁷ Stelarc’s theorization of technological disembodiment calls to mind media theorist Alexander Galloway’s argument that the protocols by which the internet promises freedom and agency in fact do so only through coding strategies of control (see Galloway, *Protocol*). I will further elaborate on this tension between absence (of the body, of control) and the underlying apparatuses by which this absence is in fact produced in my later discussion of Marshall McLuhan’s media theories.

When I think of my body and ask what it does to earn that name, two things stand out. It *moves*. It *feels*. In fact, it does both at the same time. It moves as it feels, and it feels itself moving. Can we think a body without this: an intrinsic connection between movement and sensation whereby each immediately summons the other? (*Parables*, 1)

When Massumi writes of the body in motion, he writes of the body that feels and senses.

Stelarc's obsolete body is not something that just happens, but rather is effortfully created: "the obsolescence of the body that Stelarc waxes long on must be *produced*" (*Parables*, 108). Massumi's reading of Stelarc's obsolete body might initially seem to contradict his discussion of the body as always-already obsolete; however, the dynamism that is the condition of the always-already obsolete body is not at odds with the body in art, particularly Stelarc's art, which is about the body that struggles to be a body in movement.

A body "is when it isn't" moving. Massumi theorizes the body in motion as not self-coincident: "When a body is in motion, it does not coincide with itself. It coincides with its own transition: its own variation. The range of variations it can be implicated in is not present in any given movement, much less in any position it passes through" (4). The body's non-coincidence with itself in movement and sensation suggests, according to Massumi, "an incorporeal dimension *of the body*," which he describes as "[r]eal, material, but incorporeal. Inseparable, coincident, but disjunct" (5). The body, in its non-coincidence with itself, is thus always in excess of itself. In fact, the body is *only* itself when it is not itself, when it is in motion, when it is changing – when it is obsolete.⁸

⁸ Traveling exhibits of chemically-preserved bodies and body parts (for example, Bodies the Exhibit) rely on this ontological paradox, exhibiting bodies that no longer move or sense – bodies that are no longer

4.2 Stelarc's *Extra Ear*

4.2.1 Iteration 1: *Extra Ear on the Face*

Stelarc's ongoing project, *Extra Ear*, which he began in 1997, has gone through several permutations. Stelarc originally intended to attach a third ear, acquired from a donor, onto the right side of his face next his ear. This third ear would have been equipped to emit sound – “speaking” rather than hearing.⁹ This vision of the project was unrealized, as Stelarc was unable to find physicians to participate in the project. The procedure would have posed significant risks to Stelarc's health, such as the possibility of muscular and nerve damage during the attachment procedure, as well as of rejection of the donor ear (Clarke, 203). However, Stelarc suggests an additional factor that contributed to physicians' reluctance to perform this attachment. According to Stelarc, his project threatened to disrupt a specific conception of the body, a conception whose parameters can be said to be delineated by the limits of cosmetic surgery:

The difficulty is that it goes beyond cosmetic surgery. It's not just a matter of me having enough money to pay for an operation, because this involves more than one kind of medical practitioner. It's going to involve reconstructive surgery, skin stretching, some orthopaedic work and perhaps some micro-surgery. And

bodies, precisely because they are bodies. These bodies and partial bodies, infused with chemical preservatives, encased, and displays, do not move. While the exhibits travel around the world, the bodies, rest assured, are not in motion as bodies. These bodies are not bodies because they neither move nor sense. I would suggest that the curatorial efforts to make legible that these bodies are not bodies aim toward distancing these bodies from anything resembling movement. These bodies are not obsolete – they are, to use Massumi's characterization, “concrete.” (*Parables*, 6)

⁹ Aitor Baraibar reads this ear that speaks as defamiliarizing the body by refunctioning the ear (164).

because it's an issue of excess, it becomes more and more a problem of medical ethics. Although cosmetic and plastic surgery is now considered a part of conventional medical practice, to construct extra body parts would just not be allowed ordinarily in a tightly regulated medical profession. (Zylinska and Hall, 126)

The extra body part, this “issue of excess,” points to the ways in which certain ideas of the body are policed and protected. Stelarc aims to change and evolve the body through technology, though not without resistance it seems.¹⁰

Stelarc further elaborates on physicians' reluctance to participate in *Extra Ear*. This “issue of excess” is an unacceptable modification of the body that cannot be understood according to conventional, though some may argue no less radical, cosmetic surgery body modifications: “It is not simply about ... adjusting existing anatomical features (now sanctioned in our society), but rather what's perceived as the more monstrous pursuit of constructing an additional feature that conjures up either some congenital defect, an extreme body modification or even perhaps a radical genetic intervention” (qtd in Murray, 171). Consequently, Murray reads Stelarc's work as, among other things, a kind of disruption not just in the organization of the body, but in the scientific discourses and practices – from cosmetic surgery to genetics – that are intimately concerned with the organization of the body: “Stelarc's corporeal obsolescence here makes way for a disturbing personification of monstrous metaphor. Rather than demystifying or deconstructing the utopian aims of genetics, he makes scientific practices

¹⁰ The above quote from Stelarc is taken from an edited collection on both Stelarc and French artist Orlan, whose works have included a series of transformations through multiple cosmetic surgeries. Stelarc's juxtaposition between this “issue of excess” and the extensive space of acceptability for cosmetic surgery practices is particularly resonant.

foreign to themselves for the sake of extending their psycho-social range in the practice of art” (Murray, 172, my emphasis). While Baraibar points to the ways in which Stelarc’s work effects a defamiliarization of the body, Murray focuses on how *Extra Ear* destabilizes scientific practices, making these practices foreign and alien to themselves through the body.¹¹

4.2.2 Iteration 2: *Extra Ear: ¼ Scale*

In 2002, Stelarc began collaborating with bio-art group Tissue Culture and Art Project on *Extra Ear: ¼ Scale*.¹² This iteration of *Extra Ear* involved creating, from human cells, an ear in a laboratory, unattached from the body.¹³ Tissue Culture and Art Project (TCA) discovered that the ear could not yet be replicated at full scale, as “the engineered cartilage tissue seems to lose its structural integrity and the whole form tends to collapse on itself” (“The Art,” 6). Because the ear could not maintain its form at that size, TCA created the ear in reduced scale. Beginning in 2003, *Extra Ear: ¼ Scale* was exhibited at various sites, with a different installation apparatus for each exhibition.

¹¹ I will elaborate on this theme of destabilization and disorientation in Stelarc’s work in a later section on the abject.

¹² TCA previously worked with Dr. Joseph Vacanti, a leading scientist on tissue engineering who, along with Dr. Robert Langer, constructed a human ear out of cartilage cells and polymer, using the back of a mouse for scaffolding. For their ground-breaking article in which they first coined the term “tissue engineering,” see Langer and Vacanti, “Tissue Engineering.”

¹³ TCA describes the multiple installations of *Extra Ear: ¼ Scale* as subsequently reduced with each exhibition (TCA, “The Art,” 7). For descriptions of the various exhibition installations, see TCA members’ Oron Catts and Ionat Zurr, “The Art of the Semi-Living and Partial Life: Extra Ear – ¼ Scale.”

4.2.3 Iteration 3: *Extra Ear* on the Arm



Figure 8: Stelarc's Extra Ear.
“IMG_6363.” we-make-money-not-art. Flickr. Creative Commons
Attribution-Share Alike 2.0 Generic.

In 2006, Stelarc began work on the current version of *Extra Ear*, which involves attaching a prosthetic ear onto his left forearm: “a left ear on a left arm. An ear that does not hear but transmits” (Stelarc, *Recent Projects*, 12). According to Clarke, who has written extensively on Stelarc, this relocation of the ear to the arm, as well as the ear’s functional recalibration from hearing to speaking, dismantles hearing itself from its previous site-specificity of the ear (“A Sensorial Act of Replication,” 205). In other words, Stelarc’s *Extra Ear*, in repurposing the ear, suggests the possibility of reorganizing the body through the senses.

Clarke reads Stelarc's extra ear, before it is attached to his arm, as an organ without a body, an inversion of Deleuze and Guattari's theorization of Antonin Artaud's body without organs:

But what about organs without bodies? As partial objects created through the desire of the original body to recreate itself, organs without bodies are machines waiting for connection. Stelarc's ear replica is at the present time an organ without a body. However, this desiring machine – once grafted to the body and connected to communication technologies – renders the body as multiple, working against the hierarchical, organized, and rational body. (Clarke, "A Sensorial Act of Replication," 205)

For Clarke, the *attachment* of the extra ear the arm transforms the body into the body without organs. I understand the ear, even ungrafted, as producing the reorganization that characterizes Deleuze and Guattari's body without organs. In a later work, Deleuze describes the body without organs as oppositional to the organism, a fixed organizational system: "The body without organs is opposed less to organs than to that organization of organs we call an organism" (Deleuze, *Francis Bacon*, 44). The unattached ear destabilizes the organization of the body differently, though certainly no less than, the attached ear. Further, Stelarc's *Extra Ear* does not emerge from a "desire of the original body to recreate itself," but rather from the non-originary body's desire continually to change and reorganize. Stelarc's *Extra Ear*, and indeed this can be said for all his works, is not about the body's replication or recreation, but about the body's perpetual reorganization. In an interview, Stelarc points out that the *Extra Ear* project differs from his other projects because the ear is a permanent modification of his body.¹⁴ Previous

¹⁴ Zylinska and Hall, 125. For example, in previous works such as *Third Hand* and *Exoskeleton*, technological devices were worn during performances, and then taken off afterward.

works such as *Third Hand* and *Exoskeleton*, involve technological devices that were worn during performances and then taken off afterward. *Extra Ear* engages with this perpetual reorganization differently; in expanding the reorganization of the body beyond the performance, Stelarc forecloses a return to a “normalized” body, a body that exists outside of its technological engagements.

In an interview with Stelarc, Christian and Sebastian Abrahamsson describe the technical aspects of the *Extra Ear* project – through a Bluetooth connection, Stelarc will be able to speak with someone right into the ear; the other side of the conversation will be transmitted through a speaker in Stelarc’s mouth (294). This technological modification suggests not only a reorganization of the body and its capacities, but also a synaesthetic reorganization of the senses; further, in literally calling forth the voice of the other from Stelarc’s mouth, *Extra Ear* disorients what it means to be a subject who speaks. In the interview, Stelarc elaborates on these confused corporeal functions and organizations:

[...] the idea that the ear becomes a kind of blue tooth construct... that if you telephone me or if we are communicating by VoIP over the internet I will be able to speak to you through my ear but your voice will come in my mouth, it will *happen* in my mouth. If I open my mouth and you are close to me you will be able to hear what someone says through my mouth. To me that sort of idea or rewiring of the human body is aesthetically and conceptually exciting. Whether it has a practical use...that’s got nothing to do with it at the present time. It is that kind of conceptual experience and artistic endeavour that is exciting. Actually if I try to lip sync the voice that is coming from my mouth I will be like a bad foreign movie [laughs]. (304)

While this image of Stelarc trying to “lip sync” to someone else’s voice as it emanates from his own mouth is humorous, this image – precisely in its humor – suggests the futility of trying to maintain previous organizations of the body in the face of this new technological modification. By thinking Stelarc’s obsolete body through its dynamism,

as described by Massumi, and its continued reorganization, as discussed by Deleuze and Guattari, we can understand the *Extra Ear* as a defamiliarizing project, one that seeks to make-foreign our own relationships to ourselves as subjects and agents in the world. *Extra Ear*'s sustained reordering of the body operates less through the extra ear as an organ, but through the disruption and reordering of the organism and its relationship to the environment.

While my movement across seemingly incompatible theoretical systems might give one pause, by attending to the disorienting, reorganizing, and porous elements in these multiple systems of thought provides a prismatic understanding of Stelarc's work. Further, this prismatic discussion is necessitated by the complexity and fluidity produced by Stelarc's work. Ever unfixed, Stelarc's work engages multiple discourses and materialities, and brings them into dialogue with each other.

4.3 The Abject

My discussion of Kristeva's psychoanalytic theory of the abject might seem incommensurate with the poststructuralist and mediatic threads in this chapter. My extrapolation of the abjection's movement outside of the psychoanalytic subject is facilitated by Tilottama Rajan's characterization of Kristeva as a theorist, not of a particular school of thought, but of "trans-position." Kristeva theorizes across psychoanalysis, phenomenology, and post-structuralism, all the while foregrounding the body across transpositions (216-217).

Kristeva describes the abject as “above all ambiguity”: “We may call it a border; abjection is above all ambiguity. Because, while releasing a hold, it does not radically cut off the subject from what threatens it – on the contrary, abjection acknowledges it to be in perpetual danger” (9). The abject, as ambiguity, threatens the subject with a “revolt of being,” with a collapse of meaning (1-2). And yet the abject, while activating this threat to being itself, also protects the subject from the very threat that it invokes. The abject does not decimate the borders – between subject and object, meaning and unmeaning, inside and outside – but disturbs and disrupts without undoing them. The abject is not on the side of meaning, or the collapse of meaning, but rather exists precisely on the boundary between meaning and collapse of meaning.¹⁵ The abject borders by way of ambiguity, and this ambiguity disrupts. “It is thus not lack of cleanliness or health that causes abject but what disturbs identity, system, order. What does not respect borders, positions, rules. The in-between, the ambiguous, the composite” (4).

Kristeva introduces the corpse as “the utmost of abjection,” as death threatening life: “The corpse, seen without God and outside of science, is the utmost of abjection. It is death infecting life. Abject. It is something rejected from which one does not part, from which one does not protect oneself as from an object. Imaginary uncanniness and real threat, it beckons to us and ends up engulfing us” (4). The corpse renders the boundary between death and life *ambiguous*. The corpse does not evacuate all life, or

¹⁵ In *Extra Ear: ¼ Scale*, I suggest that we can understand the structurally impossible full-scale ear as perfect replica, as threatening a collapse of meaning in its external identity, while simultaneously protecting the subject from this threat by diverting itself to its replication in miniature.

undo all death, but rather links these two inextricably through infection and ambiguity, unfolding the abject itself as the ambiguous border between life and death.

Kristeva discusses bodily excretions as abjection, as a border between life and death that does not separate life from death, so much as suggests that this constant calling up of death in life is in fact a condition of life. The corpse operates similarly, but threatens the entirety of the body with the spectre of expulsion, of death:

[...] refuse and corpses *show me* what I permanently thrust aside in order to live. These body fluids, this defilement, this shit are what life withstands, hardly and with difficulty, on the part of death. There, I am at the border of my condition as a living being. My body extricates itself, as being alive, from that border. Such wastes drop so that I might live, until, from loss to loss, nothing remains in me and my entire body falls beyond the limit – *cadere*, cadaver. If dung signifies the other side of the border, the place where I am not and which permits me to be, the corpse, the most sickening of waster, is a border that has encroached upon everything. (3)

In the fearful realization that the skin does not hermetically mark and distinguish the subject from the world, the fluids and excretions produce an abjection that simultaneously marks and distracts from this realization, bringing the subject to the very border that it holds him or her back from crossing. The abject brings the subject to the very border of life, meaning, and subjectivity, only to be the very force that keeps the subject from crossing this border. The subject does not cross the border, but rather exists in the border as ambiguity. This abjection, Kristeva suggests, does not so much create the ambiguity between inside and outside, but rather highlights it as a condition both of life and of death:

If it be true that the abject simultaneously beseeches and pulverizes the subject, one can understand that it is experienced at the peak of its strength when that subject, weary of fruitless attempts to identify with something on the outside,

finds the impossible within; when it finds that the impossible constitutes its very *being*, that it *is* none other than the abject. (5)

This abject as “the impossible within” that is the very condition of being emerges as a disruption that cannot call too much attention to itself, in order to protect the subject from the very threat the abject poses.

4.3.1 Abjection of the Living

The abject in Stelarc’s work does not operate through waste or expulsions, but rather through aspects of the “living” body. Rather than disordering the relationship between life and death by calling up death as that which somehow both always and never opposes life, Stelarc’s works use the living to disrupt and disorient life itself. For example, in *Stomach Sculpture* (1993), Stelarc swallows a camera, and displays the inside of his stomach on a screen, turning the body inside out. In this work, abjection operates through expulsion, abjecting the body from the inside out.

In Stelarc’s *Extra Ear*, unlike his *Stomach Sculpture*, what is abjected is not a foreign body that is incorporated, which then abjects the body itself. Rather, the *Extra Ear* abjects through incorporation rather than expulsion. Because the technological devices and modifications in *Extra Ear* – in the form of the ear on the arm and the voice in the mouth – are in fact technological devices that intentionally resemble and invoke components of the human, these devices metonymically introduce anthropomorphization. In *Extra Ear*, this anthropomorphic abjection doubles back on itself, reverberating across the very boundaries it threatens.

In Stelarc's *Extra Ear*, anthropomorphization can be understood as an inverted abjection. It is not abjection by way of the expelled or the dead that emerges in Stelarc's work, but abjection through a very component of the self that is abjected. Because the abject operates anthropomorphically, through life, the abject uniquely folds Stelarc's body in on itself. The ear, as non-object of abjection,¹⁶ cannot be understood as wholly "other," as the no-longer living, as can waste or the corpse. In this work, abjection does not operate through something external to life. The ear, this object of life, signifies the participation in the processes of sense perception by which the subject encounters and incorporates the world, and by which the world permeates and affects the subject. So the ear, the anthropomorphic non-object of abjection by which the self folds in on itself, also suggests a perpetual openness to the world. The abject, like the obsolete, does not indicate the subject's separation from the world, but rather a deep, constitutive engagement with the world. Stelarc's ear, then, reveals a complicated nexus of border-disruption – self disrupting self, life disrupting life, outside disrupting inside. The anthropomorphic abject.

The ear and the relocated voice of the other recalibrate the organization of the body and its capacities. Abjection, as Stelarc's work suggests, does not operate solely through excrement, fluid, corporeal expulsions or excesses, or even through the corpse as non-object of abjection *par excellence*. Rather, the living body itself calls into being this

¹⁶ Kristeva describes excrement as "excluded as a possible object, asserted to be a non-object of desire, abominated as ab-ject, as abjection" (65). The abject, as Kristeva articulates, is neither object nor subject

abjection, disrupting the border between the inside and the outside, the subject and the world.

Stelarc's extra ear, though repurposed so that it does not hear, or sense, nonetheless invokes the sense of hearing. Indeed, Stelarc's ear activates the abject both through the ear as "extra" as well as the ear as functionally repurposed. The abjection that occurs relies on the ear occupying these new locations and functions while constantly recalling the ear both as "extra," as well as an ear that speaks *instead of* hears. In *Extra Ear*, the abject's threat lies in the juxtaposition of these previous organizations of the body with a new one.¹⁷ In Stelarc's work, these previous models that are constantly recalled also serve to "reassure" the subject.¹⁸ The anthropomorphic abject inverts – it is not fluids and excrement that perform this subject reassurance, protecting the subject from the very threat they pose. Rather, the body with only two ears, both of which were designed to hear, and neither of which emit sound, the body that was permanently transformed by the attachment of the ear on Stelarc's arm, is the body that is expelled, that sits as a threatening specter on the *other* side of the border, the unassimilable that protects the subject from itself in the reorganization of Stelarc's modified body.

¹⁷ The subject here as spectator, for the ways in which the ear operates on Stelarc would be significantly different from the ways in which the ear operates on the spectator.

¹⁸ According to Kristeva, bodily fluids and excrement threaten the subject, revealing that the subject is neither pristine nor completely self-possessed. However, the externality of these fluids and wastes also "reassure" the subject of the very threat that they present, allowing the subject to reclaim this idea of the "own and clean self." "The body's inside, in that case, shows up in order to compensate for the collapse of the border between inside and outside. It is as if the skin, a fragile container, no longer guaranteed the integrity of one's 'own and clean self' but, scraped or transparent, invisible or taut, gave way before the dejection of its contents. Urine, blood, sperm, excrement then show up in order to reassure a subject that is lacking its 'own and clean self'" (53).

The ear that does not hear speaks to abjection in two ways. Firstly, as I discussed, Stelarc's ear activates a uniquely anthropomorphized abjection. Secondly, as I will discuss in the following section, this unhearing ear points toward an understanding of the senses through abjection. The senses, in bringing the world to us and into us, and likewise bringing us into the world, continually disrupt the boundaries between inside and outside, subject and object. The senses, as sites of openness to the world, make ambiguous these boundaries, in a perpetual and acceptable abjection, that we do not conceive as abjection, largely because the senses operate through invisibility. The senses reveal the world to us, while diverting us from the very means by which we encounter and know the world. Unlike waste and the corpse, and unlike Stelarc's ear, the senses, the non-objects that activate the object, disappear themselves in the process of perception.

4.4 The Double Invisibility of the Senses

Michel Serres discusses efforts to understand the world as effectively black box-ing the processes of sensation by which knowledge itself emerges:

There is no science that does not in some way presuppose a preceding sensation, even if it has sometimes, often, almost always – always, in fact – been necessary to expel the senses from the field it occupies. Not only does sensation stand behind the knowledge that presumes to speak of it, but what is more, it finds itself ousted by what we know at any given point... We do not know sensation: we might as well say that it occupies this black box.¹⁹

¹⁹ Serres, 128-129. Norbert Wiener describes the black box as a component of a larger system, by which input information is processed and output, and the operations by which input becomes output are not known. In other words, the black box brackets the processes by which information or knowledge – the output – emerges, in order to attend to the knowledge itself (Wiener, *Cybernetics*, xi). This is the invisibility by which the senses offer up to us knowledge, indeed the world.

The senses operate through invisibility. And it is this invisibility by which the senses offer up to us knowledge, indeed the world. In other words, sense perception diverts us to the object of perception – the world – while rendering the processes of sensing invisible by black boxing, or bracketing the processes by which information or knowledge emerges from sensing. Part of this tendency to expel the senses or presume sensation, at least on the part of the philosopher, according to Serres, can be linked to the fact that sensation is not an object of inquiry, but rather that which we inhabit: “The philosopher holds forth about sensation, yet he inhabits it already, dwelling in a kind of sensation, a part of his house as the pupil is part of his eye. [...] The house is a picture box, like a skull or an eye. The philosopher inhabits his own problem” (147-148). We exist in sensation. How, then, does one make that in which we are immersed the object of study?²⁰ Erin Manning, drawing from the work of Gilbert Simondon, Deleuze and Guattari, and Massumi, suggests we begin by not presuming that we understand how the senses work, by not inadvertently, or knowingly, black-boxing the senses as a component of sensation: “The challenge when working with the senses is to not presuppose that we already know what it means to sense” (Manning, xii). Taking seriously, then, inquiries

²⁰ My research has not yet produced a satisfactory articulation of the exact definition of sensation and its relationship to the senses and perception. I find Davide Panagia’s definition somewhat helpful, “neither sense nor perception (though both are crucially involved), but rather the heterology of impulses that register on our bodies without determining a body’s nature or residing in any one organ of perception” (Panagia, 2); Massumi defines sensation as “the registering of ... potential connections,” the “direct experience of a more to the less of every perception. It may be considered a third pole or limit of experience, accompanying each degree of action-perception” (Massumi, 92). While both of these definitions speak of the relationship between sensation and sense perception, neither of the definitions seems to me particularly clear or specific about the exact nature of this relationship. For the purposes of this discussion, I understand sensation as a rubric for understanding the senses and perception, though sensation also speaks to phenomena that are in excess of sense perception.

into the senses might lead us to reframe questions such as “What are the senses?” or “How do the senses work?” as “What happens to the senses when we turn our attention to them, when we observe them?”

What is at risk in un-black boxing the senses? Would the world lose focus if we seek to understand the senses? Is this the difficulty of inquiring into the senses? Would we lose something of the world if we un-black box the senses? This is the abject epistemological danger posed by the processes of sense perception – that the world that is presented to us by the senses in fact disappears or is somehow undone as the senses come into relief.

Manning, whose discussion of sensing and the body draws significantly from Deleuze and Guattari’s *body without organs*, defines the act of sensing as extending the body and worlding in multiple directions, worlding outward toward the world as well as inward toward us:

To sense is to deviate from the organic capacities of the body, to challenge the interstices between insides and outsides, spaces and times. To sense is to world in all directions at once. Through the senses, bodies become alchemical mixtures, incorporeal concoctions of visions and touches, smells and sights, tastes and sounds. Senses lead us without taking us by the hand. Senses draw us toward an object as they modulate our own responses, relaying insides and outsides into a conglomerate that deviates, always, from the implied borders of our skins. There are no sense-borders: sense is not a limit-concept. To sense is to world unlimitedly. (155)

Sensing, not just the acutely defamiliarized sensing of Stelarc’s *Extra Ear* project, but sensing in all its familiarity, effects the reorganization that characterizes the body without organs. Manning extends this sensing, which isomorphically speaks to the movement of Kristeva’s abject, outward into the world, suggesting that a reorganization of the body

cannot be thought outside the body's deeply non-bordered relationship to the world. The senses do not just present the world to us; in fact, they shape and unfold it.

In Manning's discussion of the body in the above passage, the senses operate both on the world and on the body, creating the former by exploding the limits of the latter. The sensing body is not a self-contained, self-sustained body, but rather an infinite body of multitudes:

A sensing body is an infinite body. Infinity is not eternity. A sensing body does not return to the garden nor does it perform miracles through touch. A sensing body ruptures conceptions of time and space that are considered stable, reaching toward a continued metamorphosis of the body that violently spaces time and times space. Acute senses allow the world to appear in relation. Bodies are qualitatively altered by sense. Bodies emerge as multitudes, infinitely sensing in excess of their organisms, reaching toward songs of experience. (Manning, 83)

For Manning, the senses are acutely disruptive, "violently" altering the body, time, space, and the world; the senses thus recalibrate the body into bodies, and insist on the necessary relationality of these bodies with that which is "in excess of their organisms" – the world.

In light of these discussions of the invisibility of the worlding senses, we might reframe the question of what is at stake in un-black-boxing the senses, by asking *how* we might un-black-box the senses. Drawing upon media theorist Marshall McLuhan's theorization of the role of art, I propose that it is through Stelarc's work, precisely as art, that we can in fact put the senses under a microscope without sacrificing knowledge of the world. Through art we can simultaneously attend to both the senses and the world they perceive and unfold. Stelarc's *Extra Ear*, as a specific example, highlights the worlding senses through the movement of abjection. *Extra Ear*, both in its current iteration and taken in its longer history of iterations, makes visible the senses (and the

obsolete body that senses) by disrupting and reorganizing them through abjection.

Reading the senses as acceptable abjection offers a way to understand the invisibility of the senses in experience, as well as in knowledge projects. Following from Kristeva's discussion of the abject, we can understand sensory abjection in relation to the violation of the subject's boundaries; the senses realize, indeed are the agents of, this threat, but in their invisibility keep this very realization hidden from us.

4.5 McLuhan and the Senses

Marshall McLuhan famously theorized media technologies as extensions of the senses.²¹ As the senses play a crucial role in organizing experience, their extensions, media technologies, similarly shape how we experience the world. According to Eric McLuhan, the *Laws of Media*, a collaboration between himself and Marshall McLuhan, emerged because “[t]he study of the senses and of the ways in which media extend and modify sensibility needed systemic attention” (ix). For McLuhan and McLuhan, the study of media technologies *is* the study of the senses.

According to McLuhan, each media technology creates a new environment that reorganizes our senses: “A medium creates an environment. An environment is a process; it is not a wrapper. It’s an action, and it goes to work on our nervous systems and on our sensory lives, completely altering them” (“The Medium Is the Message,”

²¹ “[O]ur human senses, of which all media are extensions [...] configure the awareness and experience of each one of us” (McLuhan, *Understanding Media*, 21). In an interview, Stelarc articulates: “I think that one of [McLuhan’s] most beautiful statements is the one I mentioned yesterday that technology is the external organs of the body. I think that is just a wonderful statement about our relationship with our technologies. Technology is an essential but external organ” (Abrahamsson, 301).

91).²² These environments are not passively inhabited spaces, but rather “active processes” that reconfigure our nervous systems and our senses, and thus posit new organizations of our bodies (“Introduction to the Second Edition,” *Understanding Media*, viii). These environments, McLuhan theorizes, are invisible while they are currently “active,” while they are operating on us. Indeed, the more potent and immersive the environment, the more invisible it is to us.²³ It is only when these environments are supplanted by newer media technologies that these no-longer current environments become visible. “But in the case of environments that are created by new technologies, while they are quite invisible in themselves, they do tend to make visible the old environments. We can always see the Emperor’s old clothes, but not his new ones” (“The Invisible Environment,” 164). In other words, we can only see older environments when they are no longer active processes, and thus no longer working on and reconfiguring our senses. Like the senses, which are invisible to us in their activity, in our perceptual encounters with and unfoldings of the world, current, active technological environments are invisible to us.

²² We cannot underestimate the effects of our environments on us. For example, McLuhan cites Erwin Strauss’ reading of Pavlov’s behavior-conditioning experiment as effected not through stimuli, but rather through the environment: “[Pavlov] didn’t get his conditioning effects by means of stimuli or signals to his experimental subjects. Rather he did it by environmental controls. He put his subjects in environments in which there was no sound, in which the heat and other sensory controls were very carefully adjusted and maintained steadily. Pavlov discovered that if you tried to condition animals in an ordinary environment, it did not work. The environment is the real conditioner, not the stimulus or the content” (“The Invisible Environment,” 165).

²³ “This is another mysterious feature about the new and potent electronic environment we now live in. *The really total and saturating environments are invisible. The ones we notice are quite fragmentary and insignificant compared to the ones we don’t see.* The English language, for example, as it shapes our perceptions and all our habits of thought and feeling, is little perceived by the users of the English language. It becomes much more perceptible if we switch suddenly to French” (McLuhan, “The Invisible Environment,” 164 [my emphasis]).

While the invisible new technological environment makes visible the older media technology, this older media technology can also be said to suggest that which is invisible. In other words, we might understand the visibility of the older technology as pointing to that which this visibility obscures. This move from invisibility to visibility, from the current to the obsolete, is what McLuhan invokes in his famous phrase, “the medium is the message.”²⁴ With every new mediating technology, the previous medium ceases to be invisible medium and becomes, in its visibility, content or message. In other words, media technologies and the environments they create are invisible to us when they are currently mediating, and only become visible when they are no longer the means by which the world is brought to us, but rather the world itself that is brought to us in technological encounter.

McLuhan’s discussion of media technologies relies on an understanding of the senses as always unified. While the ratio of the senses shifts in relation to each media technology, the senses never work in isolation: “Each of our senses makes its own space, but no sense can function in isolation. Only as sight relates to touch, or kinesthesia, or sound, can the eye see” (McLuhan and Nevitt, 13). Each new environment, each new technology affects this unity of the senses, changing the ratios by which the senses co-

²⁴ “If the new environment is invisible, it does serve to make very visible the preceding environment” (McLuhan, “The Invisible Environment,” 164). Media theorists Jay David Bolter and Richard Grusin’s theory of remediation is an account of this mutual reliance and tension between this quality of media visibility (“hypermediacy”) and media invisibility (“immediacy”). Bolter and Grusin describe remediation as “the formal logic by which new media refashion prior media forms” (273), whereby newer media “refashion” existing media. Remediation involves the dual logics of immediacy and hypermediacy: “Although each medium promises to reform its predecessors by offering a more immediate or authentic experience, the promise of reform inevitably leads us to become aware of the new medium as a medium. Thus, immediacy leads to hypermediacy.” (*Remediation*, 19).

operate.²⁵ In changing the ratios of the senses, each technological encounter reconstitutes the body itself, rendering the previous organization of the body visible only when it is supplanted by a new corporeal organization.

4.6 Stelarc, McLuhan, and the Role of the Artist

The artist, according to McLuhan, is concerned primarily with creating “counter-environments”: “[...] the role of art in the past has been not so much the making of environments as making of counter-environments or anti-environments” (“The Invisible Environment,” 165). In this creation of counter- or anti-environments, artists create new perceptual organizations that in fact allow us to perceive our previously invisible current environments. The counter-environment renders the current visible in its presence, rather than in its obsolescence.

By way of their ability to perceive and make perceptible our current technological environments, artists play a crucial role in our society. In a lecture given in 1964, McLuhan suggests that the artist has a unique capacity to perceive the current environment, and thus models these new environments for the non-artist: “The artist tends to be a man who is fully aware of the environmental. The artist has been called the ‘antennae’ of the race. The artistic conscience is focused on the psychic and social implications of technology. The artist builds models of the new environments and new social lives that are the hidden potential of new technology” (“Cybernetics,” 49). Art

²⁵ Indeed, this technological reorganization of the senses, both in their functions and locations, is precisely what Stelarc’s *Extra Ear* effects.

frames new, current environments, inverting the standard relief by which new environments make visible older environments. In other words, art, as counter-environment, is introduced as a new environment, thus making visible and perceptible our current environment. In art, our current environment is momentarily suspended and converted into content, and made visible.

Stelarc's *Extra Ear* directly speaks to the reorganization of the senses that the work, as art, effects. "Displacing percepts is the role of the artist" (McLuhan and Nevitt, 17). According to McLuhan, art's counter-environments, like environments themselves, work on and through the senses, operating on the "changing sensory ratios and sensory patterns"²⁶ of the body. Stelarc's extra ear literalizes and makes visible precisely this perceptual displacement, this recalibration of the senses. This third ear is excessive according to previous sensory ratios, disrupting pre-existing sensory patterns; and yet, if we reorganize the senses and their collaborative ratios, this ear as excess ceases to make sense, and is instead integrated into the organization. Indeed, notions of excess – and conversely, of lack – cease to apply within this reading of Stelarc.

The concept of the prosthetic, which is frequently used to characterize Stelarc's work, falls short of the reorganization of the senses and the body that Stelarc's works,

²⁶ Contemporary philosophers Jacques Rancière and Davide Panagia identify this reorganization of the senses as political. Rancière identifies art's introduction of new sensory organizations as a means to affect new political subjectivities (see *The Politics of Aesthetics*). Similarly, Panagia describes the "political life of sensation" as the interruption and reorganization of the senses. According to Panagia, these moments of sensory interruption are political in their disarticulation of old associations and subsequent engendering of new ones (see *The Political Life of Sensation*).

particularly his *Extra Ear*, perform.²⁷ Elizabeth Grosz, during a discussion session at the June 1996 *Anybody* conference in Buenos Aires, critiques the prosthesis as relying on an idea of a previously unified body, and posits the interface as a more useful way to understand the relationship between technology and the body:

Prostheses imply an intact body, or at least the potential of an intact body. Interface implies that there is no organic unity of the body, that the body is a multiplicity of sites that link in a multiplicity of ways with a whole series of things. I should point out that my parents actually make prostheses, so I'm not entirely against prostheses. I like them in certain contexts. But prostheses imply an identity of the body augmented by an external supplement. The notion of the interface, on the other hand, is interesting because it has an internal requirement, an internal disconnection that allows a reconnection elsewhere. (*Anybody*, 145)

Grosz, in the above passage, describes the logic of the prosthetic as the logic of excess.

And this logic is precisely what becomes moot in Stelarc's production of the obsolete body. What can be understood as prostheses, as technological externalities that are ontologically excessive for the "intact body" – the robotic hand of *Third Hand*, the "extra" ear of *Extra Ear*, the six-legged pneumatic walking machine of *Exoskeleton* – might be better understood outside the conceptual parameters of excess, lack, and prosthesis, outside a notion of the "organic unity of the body." In the counter-environment produced by Stelarc's *Extra Ear*, the extra, technological, anthropomorphic non-object of abjection – the ear – points to technological engagement as perpetual sensory and corporeal reorganization outside of the intact body. Stelarc's obsolete body

²⁷ For discussions of Stelarc through the prosthetic, see Julie Clarke, who amends Stelarc's own discussion of his *Third Hand* work as a "prosthetic attachment [that] represents excess," by suggesting that this excess "is always accompanied by loss" ("The Human/Not Human," 37); Arthur and Marilouise Kroker, who call Stelarc "the artist par excellence of prosthetic culture" ("We Are All Stelarc's Now," 72); and Zylinska, who, in her discussion of Stelarc, posits the prosthetic as a way of thinking about ethics, identity, and difference ("The Future... Is Monstrous," 216).

is not this intact body, but the body that is perpetually reorganized in its unavoidable openness to and engagement with the world.

5 Anthropomorphic Transduction and Speculation: Spielberg's *A.I.* and Microsoft's Milo

Steven Spielberg's *A.I. Artificial Intelligence* (2001) and Microsoft's currently unreleased virtual and interactive boy, Milo, draw on the various technologies, theories, and imaginaries discussed in the previous chapters: conversing artificial intelligences, embodied and conscious machines, emotional and empathetic robots, and anthropomorphic incorporations and expulsions. In *A.I.* and Milo's fusions of the cultural imaginary and the technological, anthropomorphization is turned outward, operating in and through the screens on which *A.I.* is projected and which Milo inhabits. This chapter grapples with this anthropomorphization in *A.I.* and Milo alongside recent philosophical theories that, in very different ways, seek to understand the world outside of anthropocentrism. These theories, Adrian Mackenzie's theory of transduction, Graham Harman's object-oriented philosophy, and Quentin Meillassoux's speculative materialism, privilege the human only to the extent that they equally privilege objects and, more generally, the not-human.¹

Virtual boy Milo and *A.I.*, which features an android boy with the capacity to love, are modeled on the human and interactive with the human. In conversation with the deanthropocentric theories of Mackenzie, Harman, and Meillassoux, these humanoid

¹ Though their work differs significantly, Harman and Meillassoux are often brought together under the contemporary philosophical movement, speculative realism. Edited by Robin Mackay, the "Speculative Realism" issue of the journal *Collapse* includes works by Quentin Meillassoux, Graham Harman, Paul Churchland, and Reza Negarastani, all of which propose new ways of understanding the relationship between human and world.

forms amplify, and at times disrupt, these theorizations of the human, the not-human, the world with humans, and the world independent of humans. Conversely, thinking the not-human through these theoretical approaches also highlights the processes by which Milo and *A.I.* evoke and direct anthropomorphization. The conversation between these humanoid forms and de-anthropocentric theories does not so much highlight and illuminate aspects of the forms and the theories, but rather opens up interventions across these two conceptual registers.

5.1 The Humanoid Transduced

For Mackenzie culture, lived experience, and the human are indissociable from technology, to the extent that to even name “the technical” and “the social” is to obscure how deeply enmeshed these two spheres are:

The volatile essence, the mutable, divergent and eventful character of technologies within our collectives eludes classification as merely “technical” or “social.” The interplay between what counts as social/cultural and what counts as technical is far more convoluted than most existing accounts admit. (2)

Humanoid technologies and representations inhabit and demonstrate this constitutive interplay in a singular way; the humanoid, in its technologies and its cultural representations, are created and imagined precisely to embody this interplay. As technologies, they are just as much imagined as they are engineered. And as imaginaries, they are just as much technical as they are fictional. While culture and technology enjoy a mutually constitutive relation, they are not synchronous with each other. “When we think about originary technicity, we can expect to find a complicated interlacing of

anticipation and delays. There is instability and movement at the point between technology and culture, but this is not because either is an autonomous agent” (10).

Technology and culture are not autonomous, but they also do not unfold at the same pace.

Mackenzie’s concept of “originary technicity” pairs deconstruction’s supplement with Gilbert Simondon’s “technicity.” “More unnerving and unlocatable” than prosthetic addition, Mackenzie describes “originary” as constitutive supplement, in which the addition is embedded in the prior constitution of that which is being added on to.

[O]ne way to describe something more unnerving and unlocatable than merely strapping on, implanting or even injecting gadgets into living bodies. By now, “originary” has become familiar shorthand for the deconstructive logic of the supplement. The logic of the supplement describes all those situations in which what was thought to be merely added on to something more primary turns out to be irreversibly and inextricably presupposed in the constitution of what it is said to be added on to. (7)

This logic of the supplement, Mackenzie points out, is thus inextricably temporal, fundamentally binding the question of technology to bodies and time.²

In drawing on Gilbert Simondon’s concept of “technicity,” Mackenzie moves his discussion beyond the purview of “the human.” Technicity “can be brought forward to show how a margin of indeterminacy is associated with technology that neither belongs solely to human life nor belongs to some intrinsic dynamism of technology” (10). This “margin of indeterminacy” expands both the human and technology beyond themselves and into relation with each other.

² “[T]o speak of the inextricability of bodies with technology is also and always to speak of time.” (Mackenzie, 8).

Alongside originary technicity, Mackenzie turns to the concept of transduction, also from the work of Simondon. “*The hallmark of a transductive process is the intersection and knotting together of diverse realities*” (13, my emphasis). The transductive process creates a nexus of multiple realities. In his discussion of transduction, Mackenzie points specifically to computer game technologies and interfaces, arguing that we understand them not as presenting closed worlds that erase difference, but rather as productively generating connections between different worlds, bodies, and images.³ Transductive processes require that we think transductively about them. Mackenzie describes this transductive thinking as “suspending any prior, separate substantial unity in either technology or the collectives (societies, cultures, civilizations, etc.), and attending to the processes that separate and bind them” (19). To think transductively about the humanoid is to think first about the *relation* between human and not-human. In Mori’s Uncanny Valley, the terms human and not-human come first, and condition the relation between them as opposition. Mori’s approach not only reifies the terms, but the relation as well. In a transductive approach, the terms emerge from the relation, which is nuanced in its dynamism and complexity.

³ “The stereotype of computer games regards them as closed worlds, purged of differences, and mostly involving narratives that emphasize extermination of differences rather than affirmative engagement with them. Those circuits, however, also engender linkages between bodies and images which compose the mediatized whatever body” (Mackenzie, 147).

5.2 Milo, Molyneux, and the Creation of the Human

Milo, created by Lionhead Studios game company, runs on Microsoft's Project Natal system; this game system tracks and senses the human body in its entirety, rather than employing an external controller. "You Are the Controller. Introducing Project Natal, a revolutionary new way to play: no controller required. See a ball? Kick it, hit it, trap it or catch it. If you know how to move your hands, shake your hips or speak you and your friends can jump into the fun -- the only experience needed is life experience" (Project Natal). As neither Milo nor Project Natal has been released as of this writing, my discussion of Milo will draw primarily from Lionhead's official video demonstration. Thus, my concern will be less with the technical components of Milo and Natal than with the narrative by which he is introduced.⁴

Peter Molyneux, the director of Lionhead Studios, debuted both Project Natal and Milo at the 2009 Electronic Entertainment Expo (E3), an annual showcase for video and computer games. Steven Spielberg introduced Project Natal to the crowd: "The gamer in me went out of my mind,' [Spielberg] said. 'It's not about reinventing the wheel, it's having no wheel at all'" (qtd in Greenop).⁵ Molyneux, by way of introducing Milo to the public, then debuted the video demonstration. The video opens with Molyneux explicitly invoking the cultural imaginary in humanoid technology. "Science fiction writers,

⁴ The video has since been widely circulated over the Internet. On YouTube, for example, the demo has been watched over one million times.

⁵ Elsewhere, Spielberg described Project Natal as a historical technological development, saying "I felt like I was present for a historic moment, a moment as significant as the transformation from the square-shaped movie screen to CinemaScope and then to Imax" (qtd in Lee).

filmmakers, they haven't imagined what we've been able to do today. We've been experimenting with something here. I'd like you to meet a boy called Milo" (Milo Demo). Molyneux posits a tenuous opposition between the cultural and the technological, as Milo emerges just as much from Lionhead as it does from a larger cultural imaginary comprised of science fiction, film, science, and technology, among others. While Molyneux may be right that writers and filmmakers have not imagined Milo *exactly*, Lionhead may not have been able to imagine and create Milo without the imaginings of science fiction writers and filmmakers, not insignificantly those of Spielberg.⁶

5.2.1 What Claire Knew

In the demo, Molyneux hands off Milo's introduction to Claire, with whom Milo will interact. "[T]his is Claire. She's going to introduce you to Milo."

⁶ Journalists Johnson and Lee both mention Milo's resemblance to the projected videos of the lead character's deceased son in Spielberg's film, *Minority Report*. Further pointing to the already richly intertwined spheres of technology and the cultural imaginary, as well as to the rich vectors of influence *within* these respective spheres, this film is based on a short story by Philip K. Dick, entitled "The Minority Report."



Figure 9: Milo and Claire greet each other.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

The video opens with Milo on a tree swing. When he spots Claire, he jumps off the swing and walks toward her. Claire greets him:

Hiya Milo, how're doing?
Hi Claire. You okay?
Actually, I'm a bit nervous.

Molyneux, who narrates over Claire and Milo's encounter, explains to the viewer that Milo "recognizes" Claire. "Here we're seeing Claire being recognized. And the emotion in Claire's voice being recognized. And that emotion reflecting in Milo's face." Milo identifies Claire's emotional state from her voice. And, in a sympathetic gesture, reflects Claire's emotion on his own face. Molyneux's explanatory narrative does not simply translate what the viewer is seeing in relation to the new technology, but is in fact filling

in the imaginative – indeed, anthropomorphic – gaps when what is happening is not so clearly legible to the viewer (see figure 10).

Molyneux’s narration, working in conjunction with the interactive demonstration, intervenes on the level of the anthropomorphic imaginary. His narration amplifies, if not prompts the viewer altogether, to project intentionality and interiority onto and into Milo in his interaction with Claire. For example, in the first exchange between Claire and Milo, during which Milo recognizes and reflects Claire’s nervousness, I do not see Claire’s nervousness reflected on Milo’s face. It is not until Molyneux fills in the narrative gaps that I project, however reluctantly and dubiously, and even if only as discarded possibility, this intentionality, emotion recognition, and reflection onto Milo. If this process of emotion recognition and reflection is not legible to me as I am watching the video, at the very least I understand that this is what Milo is *supposed* to do, that he is *supposed* to be, among other things, sympathetic.

This video introduces Milo into the cultural imaginary, presenting him as a sympathetic boy with whom one can speak and play football, who has school projects and classmates. Primarily through Molyneux’s narrative overlays, the video also introduces the human with which Milo interacts. There are several levels of interaction in this video: between Claire and Milo; Claire-Milo as interactive dyad and the viewer; Molyneux and the viewer; and Molyneux-Claire-Milo and the viewer. All of these interactions aim to introduce Milo as always more than himself – as always interactive with. Thus, the video is not introducing only Milo, but also the human who will interact

with Milo. This human emerges both through modeling Claire’s behavior with Milo, as well as through Molyneux’s narrative interpellations.



Figure 10: Molyneux fills in the gaps.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

The demo continues with Claire referencing a previous game of football.⁷ “So listen, I was thinking today you should let me beat you in football again. That is, if you finished your homework.” At this point, Milo gives Claire a sideways glance and starts walking, leading Claire to ask, “You have finished your school project...” The video does not show Claire and Milo playing soccer together; rather, Claire and Milo’s conversation cues the viewer to imagine this scenario, prompting him or her to both anthropomorphize,

⁷ As Molyneux, Claire, and Milo are British, it is safe to assume that Claire is asking Milo to play soccer.

to envision Milo playing soccer with Claire, and to imagine this scene in relation to Milo and Project Natal's technological capacity, to speculate that a human, for example the viewer him or herself, could have such an interactive, multiple-reality engagement with Milo.⁸

Molyneux deciphers Claire's side of the conversation for the viewer: "What happened there is that Claire knew Milo so well she knew when he was worried about something. The head was down, he wasn't looking at the camera so much, and this is about you meeting a character, a person." Molyneux is not deciphering the technological mechanisms of Milo and Natal; rather, he is giving Milo a mental interiority, narratively interjecting intentionality into Milo's actions and appearance. Just as humans are induced to do with ELIZA, as well as with Kismet and Leonardo, Molyneux anthropomorphically projects intentionality into Milo. In so doing, Molyneux cues the viewer, as future player and future consumer, to similarly interact with and imagine Milo as possessing this interiority.⁹ Moreover, Molyneux similarly anthropomorphizes Claire, ascribing particular intentions and motives to her actions and speech. That Molyneux might be able to speak to what Milo is "feeling" is one thing; that Molyneux has access to what "Claire knew" is another. This discussion of what Claire knew points to the complex logic of anthropomorphization in humanoid-human interactivity:

⁸ This is not to say that Project Natal does not have the capacity to create such an engagement. In fact, Ricochet, a game run on Project Natal, demonstrates precisely this capacity. In Ricochet, the player, using his or her own body to control an avatar, kicks and hits a ball against a wall in a virtual room.

⁹ Upon Project Natal's release, expected later this year, I look forward to expanding my discussion of Milo beyond the viewing of the demo video, to the experience of interacting game players.

anthropomorphization constitutes and informs the humanoid, but it does so only by also constituting and informing the human.

5.2.2 What Claire Felt

From inside the screen, Milo throws a pair of goggles at Claire (see figure 2), which she then catches. Milo then instructs her on how to put on the goggles (see figures 3 and 4).



**Figure 11: Milo throws Claire a pair of goggles.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.**



Figure 12: Milo shows Claire how to put on the goggles.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.



Figure 13: Claire puts on the goggles as instructed.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”

Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

Molyneux describes this exchange as demonstrating Claire's connection to Milo's world. "Claire has been thrown a pair of goggles. Notice what she did. This wasn't acted. She felt the need to reach down for those goggles. Now, everybody, every single person that has experienced this reaches down, because they feel so connected to Milo's world." At this point, Molyneux expands his projection of intentionality beyond Claire, and to "every single person that has experienced this." These other people, like Claire, reach down for the goggles "because they feel so connected to Milo's world." Molyneux expands the possible connection to Milo's virtual world beyond the interaction between Milo and Claire, and to other people outside of the world of the demo. Molyneux thus provides multiple points of possible identification in the video: the viewer can identify with Claire, in her interactions with Milo; with Molyneux, as the one who knows, who possesses privileged knowledge about Milo, Natal, and Claire; and with these invisible others who are not shown interacting with Milo, but who have interacted with him and experienced the same deep connection to Milo and his world. Future interacting-humans who, having viewed this video, will know exactly what to do when Milo throws something at them; they might even ascribe their own catches to "feel[ing] so connected to Milo's world." With just a catch, these future players can thus become part of this community of connectedness to Milo's world.

5.2.3 Where Claire Is

Milo leads Claire to the pond. The screen then moves Claire past Milo, to the edge of the pond. “Go on,” Milo says, “put your hand in. It’s not too cold” (Milo demo). Claire, whose image is reflected in the water (figure 14), moves her hand in front of the screen, as if to move it through the water. In response, the water ripples (see figure 15).¹⁰ Milo tells Claire to “swish the water about a bit. See if you can touch a fish” (see figure 16).



Figure 14: Claire sees her reflection in the water.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

¹⁰ Project Natal operates through a sensor bar, which is connected to the Xbox 360 game console. This sensor bar is equipped with the capacity to track the player’s three-dimensional movement and action. (Johnson)



Figure 15: Claire puts her hand in the water.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

“There Claire is,” Molyneux narrates. “In Milo’s world. She’s in that pond. Every hand movement is being recognized. Being able to touch fish, being able to swish the water with her hand. Everyone who’s experienced it, their hairs are standing up on the back of their hand” (Milo demo). For the viewer, Claire may appear to be in Milo’s world; indeed, she is reflected in Milo’s pond. However, Claire does not feel the wetness that is associated with touching water, just as she does not feel the heft of the goggles that Milo throws to her. Claire is partially situated in Milo’s world, though Milo is not so in Claire’s world. And while Claire affects Milo’s world, Milo does not similarly affect Claire’s world. Milo throws goggles that disappear, weightless, once they cross the

threshold of the screen; and the water in Milo’s world responds to Claire’s touch, but it does not touch her back.

5.2.4 What Claire Drew

Molyneux continues, “Now, what’s about to happen is some real magic.” Milo, and therefore also Claire, walk away from the edge of the pond. Claire picks up a piece of paper and an orange marker and draws a fish. “What do you think?” Claire asks Milo, as she moves the piece of paper towards the sensing box (see figure 8). Milo, in response, reaches his hand up to grasp the paper; he pulls it down into his world (see figure x). “Orange,” he says, recognizing something of what Claire has drawn.



Figure 16: Claire passes the drawing to Milo.
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.



Figure 17: “Orange.”
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

“This is true technology that science fiction has not even written about,” Molyneux states. Here, Molyneux distinguishes between the imaginary of science fiction and the “real magic” of true technology. And yet Molyneux, in the demo, relies significantly on narrative to give Milo an internal life of emotions, sympathy, intentionality, and materiality. In other words, Molyneux’s creation is just as much in the realm of the imaginary, the fictional, as it is in the technological. Within these intermeshed spheres of the imaginary and the technological, it is not only Milo who is multiply co-constituted, but Claire, as well as the viewer as future player and customer. Through narrative, Molyneux appeals to the imaginary, persuading the viewer to anthropomorphize both

Milo and Claire, the latter of which also serves as a point of identification for the viewer. Through this identification, Claire serves as a kind of double for the demo-viewer; thus, what Claire knew, felt, and drew is by proxy what the viewer knows, feels, and draws. And while the viewer can only watch Claire draw, he or she can, through Molyneux, “know” and “feel” what Claire knew and felt, placing the viewer in even greater proximity to Claire, and thus to Milo’s world.

5.3 Worlds Buffered, Worlds Collided

At E3, Molyneux introduced the Milo video demonstration using the concept of interactivity, suggesting that a direct connection between virtual game worlds and our world has until now been fettered by game controllers.

Surely we’ve been making interactive games for 20 years haven’t we? or 30 years? Well, no I don’t think we have. Because that thing in our hands, that thing that’s evolved in our hands, that got more and more complex, it’s got more and more buttons, actually has been the biggest barrier to what we want to create. Cause what we want to create is a connection to our world, and that’s what I believe Natal does. (Milo demo)

Project Natal, according to Molyneux, removes the controller in order to create a more direct connection between worlds; even absent this controller, Milo’s world and the world outside of Milo remain heavily insulated from each other. Claire does not leave her world and enter his, nor does Milo leave his world and enter hers. Rather, they encounter each other, each in their own worlds, occasionally coming into contact with each other and each other’s worlds through Project Natal’s sensory apparatus and the screen.

Graham Harman's object-oriented philosophy understands all encounters between objects as buffered. All objects in the world only encounter each other indirectly and in the sensual realm, the realm in which things appear to us and in which we perceive things. As the human is merely one object among many objects, human knowledge is also simply one of many objects' perspectives on and means of accessing the world.¹¹ Harman suggests, then, that we "rejec[t] any privilege of human access to the world, and pu[t] the affairs of human consciousness on exactly the same footing as the duel between canaries, microbes, earthquakes, atoms, and tar [...]" ("Vicarious," 188). Humans, canaries, earthquakes, and tar are all merely objects existing in the same world, none privileged in any absolute sense, and none encountering another directly. Harman describes this de-anthropocentric world of buffered encounters as comprised of "objects [that] exist in utter isolation from all others, packed into secluded private vacuums" (*Guerilla*, 1). If objects exist as such, the question for Harman is how do such objects encounter each other? How do they come into contact? In other words, what is the nature of their relation?

Harman's object-oriented philosophy moves away from two anthropocentric conceptions of the world. Firstly, Kant's distinction between noumena and phenomena:

Although I will claim that real objects do exist beyond human sensual access to them, this should not be confused with Kant's distinction between phenomena and noumena. Whereas Kant's distinction is something endured by humans alone, I

¹¹ Harman speaks of this anthropocentric approach as ignoring how objects, independent of humans, interact with each other: "No one sees any way to speak about the interaction of fire and cotton, since philosophy remains preoccupied with the sole relational gap between humans and the world – even if only to deny such a gap" ("On Vicarious Causation," 188).

hold that one billiard ball hides from another no less than the ball-in-itself hides from humans. (“Vicarious,” 138)

Harman does not dispute that there are aspects of objects that cannot be accessed or known by humans; rather, he argues that this inaccessibility of objects is not a quality reserved for humans alone. Objects are indifferent to other objects, human or not-human alike. In other words, just as there is nothing privileged about the way the human accesses and knows the world, there is also nothing privileged about the way access to the world is *withheld* from the human.

Harman’s second deanthropocentric intervention takes place around the correlation between thought and being: “this equation of being and thought must be rejected, since it leaves us stranded in a human-world coupling that merely reenacts the breakthroughs of yesteryear (“Vicarious,” 189). Both of Harman’s interventions abolish the human from the epistemological center of the world, thus radically recasting the nature of the world itself. As the objects that comprise this world are “autonomous realities of any kind,” Harman’s deanthropocentric world of objects is a world of multiple realities, of multiple and discrete worlds (“Vicarious,” 189).

To return to the question of how these isolated and secluded objects encounter each other, according to Harman they do so only vicariously, “through some vicar or intermediary.”¹² Harman’s object-oriented philosophy is a useful way to think through

¹² “Vicarious,” 190. In *Guerilla Metaphysics*, Harman takes up Merleau-Ponty’s concept of the flesh (53-5). For Harman, flesh speaks to the encounter of objects, but by way of buffer, of separation, unlike in Hardt and Negri, who take up Merleau-Ponty’s flesh as a means to expand connectivity among bodies, both individual and social (*Multitude*, 192).

the vicarious and buffered encounters between Milo's and Claire's worlds; additionally, Harman's de-anthropocentric intervention is certainly in line with my own study of anthropomorphization and the human. However, Harman's philosophy cannot fully account for the anthropomorphic forces from which Milo emerges and within which he operates. Milo cannot be thought outside Claire, outside the human, outside Molyneux, outside the demo viewer who participates in the imaginary construction of Milo.

As technological humanoid object, Milo cannot be understood outside of these various spheres of the human, unlike Harman's billiard ball. Milo emerges from a deep anthropomorphization, and cannot then be understood outside of it. What Harman's object-oriented philosophy does bring to humanoid technologies is precisely this understanding – that we can think neither the human nor the humanoid, nor their respective worlds, outside each other. The refusal of the human and the humanoid to stay in their “secluded private vacuums” points to the limits of Harman's philosophy. In relation to Milo as anthropomorphized humanoid, Harman's object-oriented philosophy breaks down precisely around the question of constitution. Harman's object-oriented philosophy is thus not the most useful way to think about anthropomorphized objects, human and non-human alike, to understand anthropomorphization not as originating in the human and working on Milo, but rather as a constitutive relationality from which both human and Milo emerge.

5.4 Post-Human World, Post-Human Time

Steven Spielberg's *A.I. Artificial Intelligence* (2001) takes place in a future Earth ravaged by greenhouse gases. Entire cities, from Amsterdam to New York are now underwater. Because of the scarcity of resources, governments have restricted procreation. Thus, children have also become scarce resources. And robots, who do not draw on the already scarce resources, have become "so essential an economic link in the chain mail of society." Dr. Hobby, of the robotics manufacturing company Cybertronics of New Jersey, fuses the essential with scarcity, proposing the creation of a Mecha¹³ that can love, "Love like the love of a child for its parents." Interrupting the building excitement in the room, a female Cybertronics employee asks about the moral dimensions of this undertaking, "You know, it occurs to me with all this animus existing against Mechas today, it isn't simply a question of creating a robot who can love. But isn't the real conundrum, can you get a human to love them back?" Unlike Dr. Hobby and virtually every other human in the film, this Cybertronics employee understands that to design a robot to interact with humans is in fact to design the human.¹⁴

Created by Dr. Hobby, David is this Mecha child, built to love, "a perfect child caught in a freeze-frame... always loving, never ill, never changing." Humans, who have

¹³ Mecha is shorthand for mechanical, as opposed to organic humans.

¹⁴ The female executive is also the only African-American character in the film. In that she is the character to point to the moral responsibility of the human in their treatment of Mechas, the film gestures to the history of slavery and race relations in the U.S. as a way of understanding human ethics and responsibility in their treatment of Mechas. My thanks to Priscilla Wald for drawing my attention to the connection between robots and the history of slavery in the U.S.

been destroying the world, similarly mistreat, abuse, and destroy their Mecha creations. Humans' treatment of David is largely no exception. At the beginning of the film, David is placed with Monica and Henry, a couple devastated by the illness of their son, Martin. Kept alive but unconscious by medical technology, Martin is, at the beginning of the film, incurable. Dr. Hobby searches the company's employee files for a family to test out David. He selects Henry precisely for, as a fellow Cybertronics employee points out, this "family tragedy that may qualify him above the rest." To the extent that Dr. Hobby "designs" humans for David to love, Dr. Hobby is not concerned with whether they can love David in return, but with what they need from David and what David can give to humans. As William Beard points out, "David is a product inspired by and directed towards the dreadful experience of loss, a consolation and a substitute for suffering people" (Beard, 7). Mecha David is less emotional robot than he is emotional prosthetic for Monica and Henry, as well as for Dr. Hobby, who designed Mecha David in the image of his own dead son, also named David.

Numerous reviews and critical readings of the film point to the cruelty and horror of humanity.¹⁵ Beard, who understands the human in *A.I.* as "straightforward degeneration," asks, "why would you want to be *human*?" (10). Humans create Mechas, only to mistreat, abandon, and destroy them. David, who has been programmed to love "Mommy" and pursue her love in a singularly focused and intense way, is driven to the woods and abandoned by Monica. Devastated, David is caught by Mecha hunters and

¹⁵ For examples, see Morrissey's "Growing Nowhere: Pinocchio Subverted in Spielberg's *A.I.: Artificial Intelligence*."

taken to the Flesh Fair, a circus-like event in which humans gleefully watch Mechas being torn apart, disintegrated with acid, and, we can only imagine, destroyed in other horrifying ways. One by one, Mechas are taken to the center ring and killed. When David's turn comes, his pleas for his life silence the crowd: ““Don't burn me! Don't burn me! I'm not Pinocchio! Don't make me die! I'm David! I'm David! I'm David!”” Amidst the silence, one woman stands up, disturbed: “Mecha don't plead for their lives. What is that? He looks like a boy.” Understanding David to be a real boy, the crowd turns against the emcee, “You're a monster!”, “He's just a boy!”, “Let that boy go!” David, because he is thought to be human, escapes. While humans do not love or care for Mechas, they do in fact care for and love each other, as we see in both Monica caring for a sick Martin, Monica and Henry's love for each other, and even the crowd's protection of David.¹⁶ It is not that humans do not love, or are incapable of love; rather, they do not or cannot love Mechas, even while creating Mechas with the express purpose to love humans.

Often overlooked in these discussions of human cruelty in the film are the exceptions to this cruelty. These exceptions are rare, but inflect “humanity” just as much as the moments of cruelty and brutality. The first example is the Cybertronics employee's question about human responsibility in relation to a Mecha child that can love. The second example, which takes place before the Flesh Fair, leads to the expulsion of David from his human family. At a birthday party for a miraculously

¹⁶ This is not to conflate Monica's love for Martin and Henry for what might be more aptly understood as species-loyalty in the Flesh-Fair scene.

recovered Martin, a group of boys begin picking on David, “We’re organic, you’re mechanical. Orga, Mecha. Orga, Mecha. Orga, Mecha.” “Can you pee?” a boy asks David. “I cannot,” he responds. “Then let’s see what you *can*’t pee with,” the boy says, reaching for David’s shorts amidst the other boys’ laughter. Martin, who had been competitive with and cruel to David since returning from the hospital, intervenes on David’s behalf. Twice Martin tells the boys to stop, placing his arm between David and the boys both times, as if to place himself between the boys and David. Martin’s protection indicates David’s inclusion in this human family. However, Martin’s well-intentioned protectiveness quickly becomes intensely distorted, almost killing Martin and subsequently casting David out of his family. One of the boys jabs David’s arm with a knife, deliberately activating David’s Damage Avoidance System (DAS). David, in DAS mode, runs to Martin and clings to him for protection, repeating, “Keep me safe, Martin. Keep me safe.” Not letting go of Martin, David pulls him into the pool, nearly drowning him. The next day, Monica drives David out to the woods and abandons him.

The third exception takes place around David’s escape from the Flesh Fair. To suggest that David avoids demolition at the Flesh Fair solely because he passes for human overlooks an important exception to the generally bleak depiction of humanity and humans in relation to Mechas. David’s escape is just as much owed to his passing as it is to the help of a man working at the Flesh Fair. Amanda, this man’s young daughter, sees David and tells her father, “There’s a boy in a cage.” “A real boy. He’s stuck in the cage.” Amanda’s father X-ray’s David, discovering that he is Mecha. However, when the crowd turns on the emcee, Amanda’s father seizes the opportunity to help David

escape. Unlike the crowd at the Flesh Fair, which protests David's destruction because they mistake him for a human, Amanda's father helps David while knowing him to be Mecha. David, with love-Mecha Gigolo Joe in tow, runs out of the Flesh Fair. The next scene opens with a series of family photos of Dr. Hobby and his son, David. This juxtaposition of fathers, one who saves David despite knowing he is Mecha, and the other who created David as a response to his own dead son, suggests that a mutually ethical and compassionate human-Mecha relationship is possible, but only if Mechas do not exist solely in relation to a pre-existing loss. When Mechas exist as artificial substitutions, they are not properly loved humans, much less treated ethically or responsibly, accorded the autonomy and agency that Mechas already exhibit. Because Amanda's father asks nothing of David, this is, for David, a singular interaction with a human. Recalling the Cybertronics employee who earlier challenged Dr. Hobby on the issue of human responsibility, David's encounter with Amanda's father hints at how possibility humans and Mechas might co-exist outside of this cruelty, as how things could be otherwise.

After the Flesh Fair, David's quest to become a real boy leads him, accompanied by Joe, to Dr. Hobby's home in Manhattan. There David is confronted with another David Mecha. Reframing Mori's uncanny valley as the reverberation between robotic entities, Mecha David encounters his Mecha double. This encounter completely upends David's sense of his own uniqueness. David, disturbed and believing his claim to Monica threatened, picks up a lamp and attacks his double, lopping off his double's mechanical head in a violent rage. "She's mine. And I'm the only one. I'm David! I'm David! You can't have her! I'm the only one! [...] I'm special! I'm unique! You can't

have her!”¹⁷ David, exhausted by his fit and this new information, tells Dr. Hobby, “I thought I was one of a kind.” Dr. Hobby explains, in what must have been to David the cruelest terms, that David was never, nor will ever be unique, “My son was one of a kind. You are the first of a kind.” Dr. Hobby, as promised during his opening presentation, is mass manufacturing Davids, as well as their female counterparts, Darlenes. “At Last – A Love of Your Own,” their packaging reads. David, upon having stumbled onto these countless Davids and Darlenes, moves from despair to despondence and throws himself into the water that has submerged most of Manhattan, with one last call, “Mommy.” David has experienced and recognized the full weight of human cruelty, but is still beholden to the human agents of this cruelty.

Tim Kreider describes *A.I.* as “a new story about the death of humanity itself” (34). If *A.I.* is indeed about “the death of humanity,” its mourning is of a deeply ambiguous nature. The final stage of the film takes place two thousand years later, after the world has frozen over. Buildings and cities still stand, but there are no more humans to inhabit them. This world is populated solely by super-Mechas. The extinction of the human species in this now glacial world highlights the different temporalities of Mechas and humans. Earlier in the film, David asks Monica, “How long will you live?” “For ages. For fifty years,” she says, thinking she is reassuring David. “Is fifty years a long

¹⁷ When David first sees his doppelganger, doppelganger David is reading a book entitled *What Is a Fish?*. As a counterpoint to David’s confusion about and anger towards his double, David₂ is happy to meet David. “Let’s be friends,” David₂ offers. If there is a genealogy of the imaginary to be drawn for Milo, it does not include the obsessively devoted David, but rather the genial, friendly, fish-interested David₂.

time?,” David later asks Teddy, his faithful robot bear companion.¹⁸ “I don’t think so.” What is “ages” for Monica is a moment for David who, trapped under the ocean for two thousand years, prays to a submerged statue – his wish-granting “Blue Fairy”; “Please, please make me into a real live boy. Please. Please. Make me real,” he repeats, for centuries. When David wakes up, he is in a world after humans. The film’s frozen, snowy, non-human world aptly illustrates what Quentin Meillassoux describes as the non-human “world of Cartesian extension.” Meillassoux describes this world, which he seeks to recuperate in his de-anthropocentric speculative realism, as “indifferent” to the humans’ relation to it.

[A] world that acquires the independence of substance, a world that we can henceforth conceive of as indifferent to everything in it that corresponds to the concrete, organic connection that we force with it – it is this *glacial* world that is revealed to the moderns, a world in which there is no longer any up or down, centre or periphery, nor anything else that might make of it a world designed for humans. For the first time, the world manifests itself as capable of subsisting without any of those aspects that constitute its concreteness for us. (115)

¹⁸ The animatronic Teddy was built by Stan Winston Studios, the same special effects company that collaborated with Breazeal on Leonardo (*Animating A.I., Special Visual Effects and Animation: ILM, A.I.: Artificial Intelligence Special Features* disc).

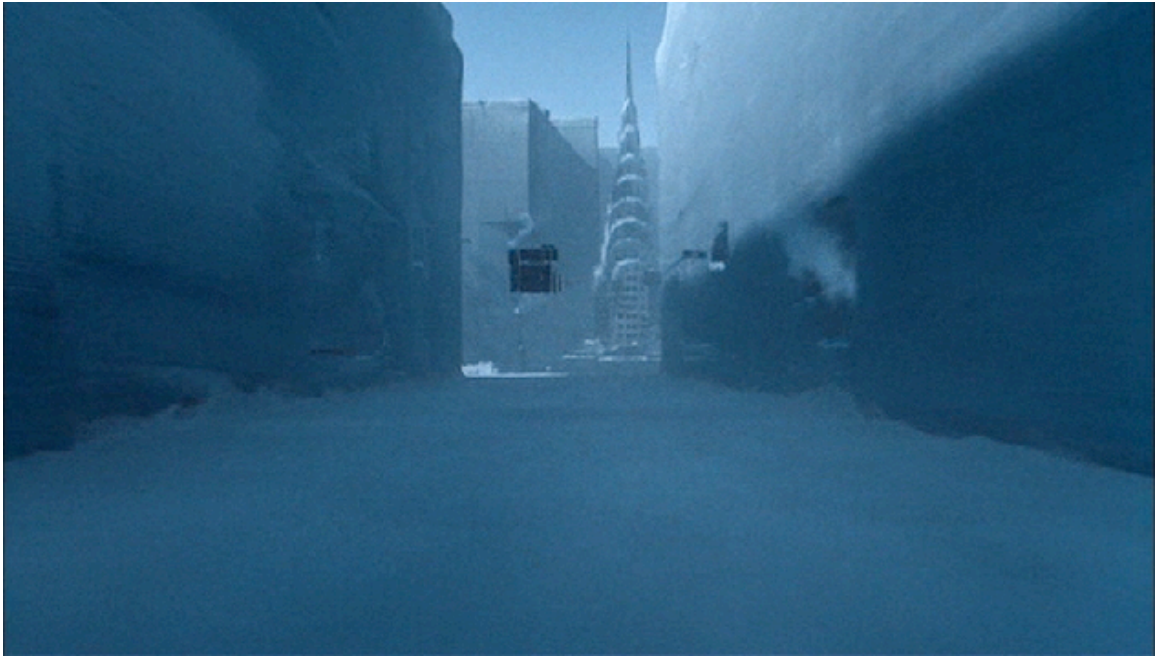


Figure 18: The glacial world after human extinction.
Spielberg, Steven dir. *A.I. Artificial Intelligence*. Warner Bros. Pictures. 2001.

To say that this non-human world is indifferent to humans is not to say that humans do not have a relation to it. Humans have a relation to this world, and this relation is constitutive. This relation is not one of mastery or absolute knowledge, but of attending to the limits by which the human can know this world. In other words, the relation between human and the non-human world is speculative.

5.4.1 Contingency and Chaos

Meillassoux's speculative materialism argues for a return to the absolute as a way to account for existence independent of and completely indifferent to the human. Like Harman, Meillassoux critiques human-centered philosophies that rely on correlationism,

the impulse to always understand thought and being in relation to each other, “the idea according to which we only ever have access to the correlation between thinking and being, and never to either term considered apart from the other. We will henceforth call *correlationism* any current of thought which maintains the unsurpassable character of the correlation so defined” (*After Finitude*, 5). Correlationist thought privileges the human in such a way that does not allow for anything outside of the human who thinks and experiences. Thus, correlationism posits an inability to think the subject and the object, the human and the world, independent of each other.¹⁹

The stakes of this critique of correlationism lie in the ability to think a world independent of humans, a world outside of anthropocentrism.²⁰ Meillassoux’s particular critique turns on two concepts, ancestry and the arche-fossil, which dismantle correlationism by indicating a time and a world before humans. The ancestral is a time that precedes humans: “I will call ‘ancestral’ any reality anterior to the emergence of the human species [...]” (10). The arche-fossil is the manifestation of this ancestral past in our present.

I will call “arche-fossil” or “fossil-matter” not just materials indicating the traces of past life, according to the familiar sense of the term “fossil,” but materials indicating the existence of an ancestral reality or event; one that is anterior to terrestrial life. An *arche-fossil* thus designates the material support on the basis

¹⁹ “Correlationism consists in disqualifying the claim that it is possible to consider the realms of subjectivity and objectivity independently of one another. Not only does it become necessary to insist that we never grasp an object “in itself”, in isolation from its relation to the subject, but it also becomes necessary to maintain that we can never grasp a subject that would not always-already be related to an object” (Meillassoux, 5).

²⁰ Within the scope of this chapter, I do not engage explicitly with the rich debate about the nature of Meillassoux’s intervention within philosophy. Here, I take up Meillassoux’s speculative materialism as it speaks to questions of anthropomorphization and de-anthropomorphization.

of which the experiments that yield estimates of ancestral phenomena proceed – for example, an isotope whose rate of radioactive decay we know, or the luminous emission of a star that informs us as to the date of its formation. (10).

The arche-fossil is not just the indication of ancestrality, but also the means by which we can understand something of the ancestral. For Meillassoux, the question that upends correlationism is, “How do we make sense of ancestrality?” Or, phrased differently, how, given the arche-fossil, do we think the time of ancestrality in the present? How do we think about a time that completely precedes human experience? Making sense of the ancestral necessitates the pulling apart of thinking and being, precisely in the act of thinking that which precedes our being. Meillassoux’s speculative materialism proposes that the only necessity of the world is radical contingency. Not thought, nor being, nor the correlation between the two – none of these exists as necessary components of the world. Only contingency.

Having cleared the philosophical ground with his temporally-organized critique of correlationism, Meillassoux turns his attention to the restoration of the absolute. This absolute is not totalizing or absolutist. Thinking the absolute is not *knowing* the absolute, but rather acknowledging the absolute as that which we can never know.²¹ In this way, the absolute is decidedly “de-absolutizing.” Meillassoux describes the absolute as not only autonomous of humans, but in fact completely “indifferent” to us and how we might encounter, sense, and think it. In fact, the absolute exists “regardless of whether we are

²¹ Meillassoux distinguishes the absolute from an absolutizing force on p. 34.

thinking of it or not” (7). This absolute is the not-human, whose conceptual genealogy Meillassoux draws to the ancestral, the before-human.

Meillassoux, like Harman, aims to move humans outside of the center of the world. For Meillassoux, this decentering is largely effected through human thought. Thus, to say that the world outside of the human is indifferent to humans does not mean that we do not have a relation to it, or that this relation is not constitutive. Meillassoux proposes instead a recalibration of the human-world relationality, such that this relation is organized by the de-absolutizing absolute, what we cannot know of the world. It is through the faculties of the human mind that we can attend to, without knowing, this absolute.

A student of Alain Badiou, Meillassoux considers mathematics the means by which one can think the absolute in its de-totalizing unknowability, its indifference to the human. “From its inception, the mathematization of the world bore within it the possibility of uncovering knowledge of a world more indifferent than ever to human existence, and hence indifferent to whatever knowledge humanity might have of it” (116). Through a discussion of mathematics, specifically Georg Cantor’s set theory, Meillassoux argues for radical contingency as the only necessity of the world, with no exception being granting to natural laws. That something is does not negate all the possibilities that are not; the difference between what is and all of the possibilities that did not occur is simply and only contingency. Peter Hallward describes Meillassoux’s radical contingency as evacuating causality: “[...]there is no such thing as reason or cause. The truth is not just that a given cause might give rise to a hundred different

effects, but that an infinite variety of ‘effects’ might emerge on the basis of no cause at all, in a pure eruption of novelty *ex nihilo*” (52). If everything can always have been something else, to think what is, is to always think beyond it to all of its possibilities. For Meillassoux, the factual, or the what is, is the grounds for speculative thought itself (*After Finitude*, 128).²²

In this world of radical contingency, the concepts of chance, probability, and predictability are no longer operable, as they all function within a pre-ordered, totalized system. “The contingent, in a word, is *something that finally happens* – something other, something which, in its irreducibility to all pre-registered possibilities, puts an end to the vanity of a game wherein everything, even the improbable, is predictable” (108). On the level of Meillassoux’s radical contingency, the system is never totalized and the set, in mathematical terms, is never calculable or thinkable in its entirety. This mathematized world of radical contingency is a world that is not only anterior to the human, by way of ancestry, but indifferent to the human even in the present:

From its inception, the mathematization of the world bore within it the possibility of uncovering knowledge of a world more indifferent than ever to human existence, and hence indifferent to whatever knowledge humanity might have of it. In this way, science carried within it the possibility of transforming every datum of our experience into a dia-chronic object – into a component of a world that gives itself to us as indifferent, in being what it is, to whether it is given or not. (116)

²² Gottfried Leibniz’s theory of compossibility, as does Hallward’s critique on the basis of “actual change” and Harman’s critique of stability amidst chaos (both of which I will discuss), argues that not all possibilities are always possible in a given moment. “Not all possibles are compossible. Thus, the universe is only a certain collection of compossibles, and the actual universe is the collection of all existing possibles, that is to say, those which form the richest composite. And since there are different combinations of possibilities, some of them better than others, there are many possible universes, each collection of compossibles making up one of them” (Leibniz, qtd in Rutherford, 182).

Hallward critiques Meillassoux's radical contingency as an abstraction that does not account for "an actual process of transformation or development." Hallward's critique continues, "Once Meillassoux has purged his speculative materialism of any sort of causality he deprives it of any worldly historical purchase as well. The abstract logical possibility of change (given the absence of any ultimately sufficient reason) has strictly nothing to do with any concrete process of actual change" (55). Harman, in a somewhat related critique, suggests that Meillassoux's "hyper-chaos," the world of radical contingency, has not yet adequately dealt with stability, namely "how stability emerges from out of chaos" ("Quentin Meillassoux," 111; 113). Almost as if anticipating these critiques, Badiou's Preface to *After Finitude*, points to how Meillassoux's thought might speak to both of these questions of transformation, materiality, and stability. He describes Meillassoux's philosophy, and the world described therein, as reorienting the ways that we understand and exist within the world, moving us from "what can I know?" to "what must I do?" and "what can I hope?" (vii). For Badiou, Meillassoux's speculative thinking is about acting and effecting, as well as hoping. And, while Meillassoux himself does not address these questions explicitly, he provides an interesting and productive prism by which to think through these questions of transformation and political action *precisely* by destabilizing stability itself, by asserting the "what is" as always "how it could be otherwise." To my mind, this is a powerful tool for transformation that does not abstract materiality, but rather provides the possibility for intervention at every moment, in every substance.

While cultural forms such as art seem like rich avenues to both effect and think through the transformative potential of Meillassoux's speculation, he rejects art in favor of mathematics:

But what is most fundamental in all this – and this was already one of the guiding intuitions of *Being and Event* – is the idea that the most powerful conception of the incalculable and unpredictable event is provided by a thinking that *continues to be* mathematical – rather than one which is artistic, poetic, or religious. It is *by way of* mathematics that we will finally succeed in thinking that which, through its power and beauty, vanquishes quantities and sounds the end of play. (*After Finitude*, 108).

Why Meillassoux creates such a rigid opposition between mathematics and the artistic, the poetic, and the religious is unclear.²³ In fact, the creative, as one way of responding to Hallward's critique, seems a deeply appropriate forum for Meillassoux's mathematical speculation, as well as a very important and effective force for speculative intervention and transformation.

Meillassoux's de-absolutizing absolute and radical contingency, which insist that things could always be otherwise, provide a way to understand the very brief moments of exception in the film. These moments are moments of speculative engagement, suggesting how things could be otherwise, how humans could be something other than cruel brutalizers of Mechas.²⁴ Within Meillassoux's philosophical framework, it is important to understand the critique of humans and humanity within the film's proposal of why and how things could be otherwise. A final moment of human exception takes

²³ The work of many Oulipian poets, for example,

²⁴ These moments, you will remember, include the Cybertronics executive who suggests that creating a Mecha that loves humans necessitates creating a human who can love Mechas, and Amanda's father helping David and Joe escape from the Flesh Fair.

place in the time after humans. For the super-Mechas, ancestry can be understood as the time of humans, that certain yet unknowable time anterior to the super-Mechas. And for the super-Mechas, David is a relic of this time, a present manifestation of this Mecha-anterior time. “This machine was trapped under the wreckage before the freezing. Therefore, these robots are originals. They knew *living people*,” a super-Mecha says, in a language that in no way resembles human speech. David, as arche-fossil, references the time of humans, and in this way he is at last unique.

According to Friedman, the super-Mechas value David for himself, unlike the humans in the film, who value David in relation to their own loss (30). However, the super-Mechas, while more compassionate and emotionally evolved, also value David for their lost connection to their human creators, for what David can tell them about humans. They value him precisely as arche-fossil. David, through his memories, provides the super-Mechas with knowledge about humans. And as a gift to David, they give him one perfect day with Monica, a day with Monica in which “there was no Henry, there was no Martin, there was no grief. There was only David.” This world after humans is a world of compassion. The super-Mechas, Beard notes, are “more important for their *emotional* perfection than their technical superiority. [...] Like David, but unlike human beings, these are creatures who can love properly, and it is they who extend to David the tenderness and care that no humans ever did, including his mother”²⁵ (10). At the

²⁵ Lester D. Friedman articulates a very similar reading of humans as “morally inferior to the humanoids they create.” (30).

beginning of the film, Dr. Hobby introduces the idea of Mechas that can love. By the end of the film, it is the super-Mechas that create a human who can love in return.

Krieder astutely points out that this one day of pure happiness rewrites all of the previously ugly or sinister moments of their previous life together (37). For example, the scene in which David, tricked by Martin, unwittingly menaces a sleeping Monica by cutting a lock of her hair, is transposed into a playful, joyful scene in which Monica washes and combs David's hair. Similarly, Martin's birthday party, which leads to David's expulsion from his family, is rewritten as a beautiful celebration of David's birthday. This perfect day, imperfect only in its finitude, ends with Monica finally telling David, "I love you, David. I do love you. I have always loved you." The narrator continues, "That was the everlasting moment [David] had been waiting for. And the moment had passed, for Monica was sound asleep. More than merely asleep. Should he shake her she would never rouse. So David went to sleep too. And for the first time in his life, he went to that place where dreams are born." This day with Monica, given to David by the super-Mechas, is David's ideal Monica (who loves him completely), in David's ideal world (absent everyone else). The actuality of Monica and his life with her was, of course, far from this ideal.

For Beard, this artificially created ideal is at the center of the film itself: "The notion that the ideal must be *imagined* and *artificially manufactured* if it is to exist at all is at the core of the meta-Spielberg nature of *A.I.*" (10). We see this with David, whom Dr. Hobby created to be an "ideal" child who loves completely and never ages, as well as with Monica, whom the super-Mechas create to return David's love. As the super-

Mechas can only bring Monica back for one day, it is Monica who is now “caught in a freeze-frame... always loving, never ill, never changing.”²⁶ The film’s “imagined and artificially manufactured” ideals speak to Meillassoux’s speculative thought and how it might transform and intervene into the actual. The super-Mechas do not give David a factual day with Monica, but rather a day that collected all the horrible moments of his life with Monica and replayed them not as they were, but rather as they were not, as they could have been otherwise.

At the end of David’s singularly perfect day, Monica goes to sleep forever. In this perfect day, David, who was not built with the ability to sleep, falls asleep next to Monica. The ambiguity of this ending has produced numerous interpretations. Does David mercifully shut down, effectively dying (Beard and Dunn)? Alternatively, does he lie awake forever, next to a never-waking Monica (Krieder, 39)? Is he finally alive, a real boy, in this time after life (Sterritt, 60)? Or perhaps in this ideal day with his ideal Monica, David is also his ideal self, the real boy who is loved by his mother, who sleeps, and who dreams.

Moving back from this future world without humans, and forward nine years after the theatrical release of *A.I.*, we find ourselves in the immediate present with Milo. Molyneux learned the lesson from *A.I.*; he not only creates a virtual boy to interact with humans, he also creates the human who will be able to interact with the virtual boy. In other words, the interacting-human emerges from the Milo-human interactivity that

²⁶ Pointing again to the different temporalities of Mechas and humans, David’s freeze-frame lasts for millennia, while Monica’s lasts for only one day.

Molyneux, like Breazeal with Kismet and Leonardo, designs. Molyneux's video introduction of Milo is effectively a primer on how to anthropomorphize.



Figure 19: David
Spielberg, Steven dir. *A.I. Artificial Intelligence*. Warner Bros. Pictures. 2001. DVD.



Figure 20: Milo
“E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD.”
Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.

Timothy Lenoir identifies the cultural imaginary as driving a desire for computer-generated imagery (CGI) in films: “Through films such as *Jurassic Park* and *Toy Story*, media industries have created a desire for computer-generated imagery. [...] We have come to desire these effects even when the film could be made without them. The appetite for ‘realism’ in visual effects forms a feedback loop with whatever technologies are currently available” (“All but War,” 305). Milo draws on a desire for special effects

realism that combines with and operates through anthropomorphization, a desire that is earlier activated by David in *A.I.*. In other words, the video does not so much introduce the technology, so much as the anthropomorphization and the cultural imaginary from which Milo emerges. How else might we understand the introduction of Molyneux, “Britain’s most celebrated video game designer” (Kendall) and Project Natal by Spielberg, whose films have significantly shaped the cultural imaginary from which Molyneux draws? In drawing on the cultural imaginary as such, Molyneux, Project Natal, and Milo are also affecting, influencing, and participating in this imaginary. Neither solely technology, though drawing significantly from AI and robotics, nor solely cultural imaginary, Milo is a hyperarticulation of the inseparability of the technological and the cultural imaginary.

Conclusion: The Human as Speculative and Singular

The human can be understood as speculation, and anthropomorphization as itself a speculative operation. When thought speculatively, anthropomorphization does not reify certain conceptions of the human (for example, the human as intelligent in the Turing test, as emotional in Breazeal's sociable robots, as engaging in specific kinds of interactivities with Milo), but can in fact generate new possibilities for the human (as seen in Brooks' embodied robots, Dick's novels, Hanson's robots, and Stelarc). In other words, in light of Meillassoux's speculative materialism, the fuzziness and elasticity of "the human" becomes a generative and productive allowance. Within this framework, anthropomorphization of different humans across different humanoid technologies and imaginaries does not reify or foreclose other forms and conceptions of the human. Instead, speculative thought demands that in thinking these only-provisionally reified humans, we also think the humans that, in that moment, exist only as possibilities, as imaginaries.

However, there is a limit to thinking possibility. My project, in line with Hallward and Harman's critiques of Meillassoux, accounts for the multiple possibilities and contingencies of the humans that do and do not exist, but is also deeply concerned with how certain conceptions of the human, both in technology and in literary forms, nonetheless emerge amidst this chaos and contingency. Kenneth Surin describes the singular in Deleuze as what is or can be as "the outcome of an always specific convergence of forces" (*Freedom*, 192). My dissertation understands the human by

tending to the “specific convergence of forces” from which specific conceptions of the human emerge. Twinning the speculative with the philosophical singular as such attends to possibilities and contingency, but not at the expense of what Hallward calls “[the] actual” and Harman calls “stability.” According to Surin, this is a significant purchase of Deleuze’s concept of the singular: “A powerful empiricism is at work in this conception of a singularity, an empiricism which derives from the insistence that singularities result from an always particular constellation of material forces” (Freedom, 193).

My project, having looked at anthropomorphization across different theoretical registers, imaginaries, and technologies, and attending to both the various humans articulated therein, as well as the operations by which these humans emerge, identifies the humanoid as uniquely inflecting, articulating, and distorting theories and philosophies of the human. Anthropomorphization moves through the cultural – literature, film, art – and the technological – artificial intelligence, robotics, motion sensing and visual effects – creating the humanoid in the image of the human, and in turn creating the human. Seaming the technological and the cultural imaginary, anthropomorphization in the humanoid articulates only in that it disarticulates; the human emerges only through the originary conflation between human and humanoid. The human, not the humanoid, is the object of anthropomorphization par excellence.

Bibliography

- a_kep. "o9_ars electronica festival." 2 Sept. 2009. Online image. Flickr. Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic. <http://www.flickr.com/photos/judo10/3881104948/>. 9 Mar. 2010.
- Abrahamsson, Christian and Sebastian Abrahamsson. "In Conversation with the Body Conveniently Known as Stelarc." *cultural geographies* 14 (2007): 293-308. Print.
- Ahn, Hyung-il, Alea Teeters, Andrew Wang, Cynthia Breazeal, and Rosalind Picard. "Stoop to Conquer: Posture and Affect Interact to Influence Computer Users' Persistence." *ACII 2007 LNCS* 4738: 582-593. Print.
- Allen, Frederick. "Unreasonable Facsimile." *Atlantic Monthly*. Aug.1994: 20-23. Print.
- Anderson, James A. "Turing's Test and the Perils of Psychohistory." *Social Epistemology* 8.4 (1994): 327-332. Print.
- "Anthropomorphization." *Oxford English Dictionary*. Web. 21 Mar. 2010.
- Appleby, John. "Planned Obsolescence: Flying into the Future with Stelarc." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 101-113. Print.
- Aristotle, *Aristotle's Physics*. Trans. Hippocrates G. Apostle. Grinnell, IA: The Peripatetic Press. 1969. Print.
- Aristotle. *Poetics*. Trans. Richard Janko. Indianapolis: Hackett Publishing Co., 1987. Print.
- Arning, Bill. "prosthetics." *sensorium: embodied experience, technology, and contemporary art*. Ed. Caroline A. Jones. Cambridge, MA: The MIT Press, 2006. 194-199. Print.
- Baldwin, Thomas. "Introduction." *Maurice Merleau-Ponty: The World of Perception*. Trans. Oliver Davis. London: Routledge 2004. 1-33. Print.
- Baraibar, Aitor. "Stelarc's Post-Evolutionary Performance Art: Exposing Collisions between the Body and Technology." *Women & Performance: A Journal of Feminist Theory* 11.21 (1999): 157-168. Print.

- barryhensey. "Babbage's Brain." Flickr. 2 Jan. 2009. Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic. 3 Mar. 2010.
http://www.flickr.com/photos/barry_hensey/3246183544/.
- Bartneck, Christoph, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. "Is the Uncanny Valley an Uncanny Cliff?" *16th IEEE International Conference on Robot & Human Interactive Communication* August 26-29, 2007. Jeju, Korea. 368-373. Print.
- Beard, William. "A.I. or, the Agony of Steven Spielberg." *CineAction* 66 (2005): 2-11. Print.
- Bieri, Peter. "Thinking Machines: Some Reflections on the Turing Test." *Poetics Today* 9.1 (1988): 163-186. Print.
- Boden, Margaret, ed. *Philosophy of Artificial Intelligence*. Oxford: Oxford University Press, 1990. Print.
- Bolter, Jay David. "Artificial Intelligence." *Daedalus* 113. 3 (1984): 1-18. Print.
- . and Richard Grusin. *Remediation: Understanding New Media*. Cambridge, MA: The MIT Press, 2000. Print.
- Booker, M. Keith and Anne-Marie Thomas. *The Science Fiction Handbook*. West Sussex, UK: Wiley-Blackwell, 2009. Print.
- Botting, Fred and Scott Wilson. *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 149-167. Print.
- Botvinick, Matthew and Jonathan Cohen. "Rubber hands 'feel' touch that eyes see." *Nature* 391 (1998): 756. Print.
- Bould, Mark and Sherryl Vint. "Of Neural Nets and Brain in Vats: Model Subjects in *Galatea 2.2* and *Plus*." *Biography* 30.1 (2007): 84-105. Print.
- Brady, Michael and Huosheng Hu. "The Mind of a Robot." *Philosophical Transactions: Physical Sciences and Engineering, the Royal Society* 349.1691A (1994): 15-28. Print.
- Brassier, Ray and Iain Hamilton Grant, Graham Harman, Quentin Meillassoux. "Speculative Realism." *Collapse: Philosophical Research and Development* III (2007). 307-449. Print.

- Breazeal, Cynthia. "Affective Interaction between Humans and Robots." *ECAL 2001 LNAI 2159*: 582-591. Print.
- . *Designing Sociable Robots*. Cambridge, MA: The MIT Press, 2001. Print.
- . "Function Meets Style: Insights from Emotion Theory Applied to HRI." *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews* 34.2 (2004): 187-194. Print.
- . and Lijin Aryananda. "Recognition of Affective Communicative Intent in Robot-Directed Speech." *Autonomous Robots* 12 (2002): 83-104. Print.
- . and J. Gray, G. Hoffman, M. Berlin. "Social Robots: Beyond Tools to Partners." *Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication*. Kurashiki, Okayama Japan. September 20-22, 2004: 551-556. Print.
- . and Rosalind Picard. "The Role of Emotion-Inspired Abilities in Relational Robots." *Neuroergonomics: The Brain at Work*. Eds. Raja Parasuraman and Matthew Rizzo. New York: Oxford University Press, 2007. 275-292. Print.
- Brooks, Rodney. "Elephants Don't Play Chess." *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press, 1999. 111-132. Print.
- . *Flesh and Machines: How Robots Will Change Us*. New York: Vintage. 2003.
- . "I, Rodney Brooks, Am a Robot." *IEEE Spectrum*, 45.6 (2008): 68-71. Print.
- . "Intelligence without Reason." *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press, 1999. 133-186. Print.
- . "Intelligence without Representation." *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press, 1999. 79-102. Print.
- . "New Approaches to Robotics." *Science* 253.5035 (1991): 1227-32. Print.
- . "Preface." *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press, 1999. vii-xii. Print.
- . "The Robots Are Here." *Technology Review*. (2008): 30. Print.
- . and Lynn Andrea Stein, "Building Brains for Bodies," *Autonomous Robots* 1 (1994): 7-25. Print.

- Brown, Bill. "Thing Theory." *Critical Inquiry* 28.1 (2001): 1-22. Print.
- Buckland, Warren. *Direct by Steven Spielberg: Poetics of the Contemporary Hollywood Blockbuster*. New York; London: Continuum, 2006. Print.
- Bull, Michael. "mediation." *sensorium: embodied experience, technology, and contemporary art*. Ed. Caroline A. Jones. Cambridge, MA: The MIT Press, 2006. 167-168. Print.
- Cañamero, Lola. "Did Garbo care about the uncanny valley?" *Interaction Studies* 7.3 (2006): 355-359. Print.
- Čapek, Karel. *R.U.R. Rossum's Universal Robots*. New York: Penguin. 2001. Print.
- Carnegie Mellon University. "Robot Hall of Fame." Web. 10 Jan. 2010.
- Carpenter, Julie, Matthew Eliot, and Daniel Schultheis. "The Uncanny Valley: Making Human-Nonhuman Distinctions." *Proceedings of the 5th International Conference on Cognitive Science*. Vancouver, B.C., Canada. 2006. 81-82. Print.
- Catts, Oron and Ionat Zurr. "The Art of the Semi-Living and Partial Life: Extra Ear – ¼ Scale." *Art in the Biotech Era*. Ed. Melentie Pandilovski. Adelaide, S. Australia: Experimental Art Foundation Inc., 2008. Print.
- Cavallaro, Dani. "The brain in a vat in cyberpunk: the persistence of the flesh." *Studies in History and Philosophy of Biological and Biomedical Sciences* 35 (2004): 287-305. Print.
- . *Cyberpunk and Cyberculture: Science Fiction and the Work of William Gibson*. London: The Athlone Press, 2000. Print.
- Clarke, Julie. "Corporeal Mélange: Aesthetics and Ethics of Biomaterials in Stelarc and Nina Sellars's *Blender*." *Leonardo* 39. 5 (2006): 410-416. Print.
- . "The Human/Not Human in the Work of Orlan and Stelarc." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 33-72. Print.
- Clynes, Manfred E. "Cyborg II: Sentic Space Travel." *The Cyborg Handbook*. Ed. Chris Hables Gray. New York; London: Routledge, 1995. 35-42. Print.
- . "An Interview with Manfred Clynes." *The Cyborg Handbook*. Ed. Chris Hables Gray. New York; London: Routledge, 1995. 43-53. Print.

- . and Nathan S. Kline. "Cyborgs and Space." *The Cyborg Handbook*. Ed. Chris Hables Gray. New York; London: Routledge, 1995. 29-33. Print.
- cognitive.evolution. "Leonardo." Flickr. Creative Commons Attribution-Noncommercial-Share Alike 2.0 Generic. <http://www.flickr.com/photos/mrgordon/519250977/>. 9 Mar. 2010.
- Copeland, Jack. *Artificial Intelligence: A Philosophical Introduction*. Oxford, UK: Blackwell Publishers Ltd, 1993. Print.
- . "The Turing Test." *Minds and Machines* 10 (2000): 510-539. Print.
- . and Proudfoot's "Turing's Test: A Philosophical and Historical Guide," *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts, and Grace Beber. New Yorker: Springer, 2008: 119-138. Print.
- Cochran, Michael. "Introduction." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 18. Print.
- Crevier, Daniel. *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York: BasicBooks, 1993. Print.
- Csicsery-Ronay, Istvan. "Cyberpunk and Neuromanticism." *Mississippi Review* 16.2/3 (1988): 266-78. Print.
- Curtain, Tyler. "The 'Sinister Fruitiness' of Machines: *Neuromancer*, Internet Sexuality, and the Turing Test." *Novel Gazing: Queer Readings in Fiction*. Ed. Eve Kosofsky Sedgwick. Durham: Duke University Press, 2007. 128-148. Print.
- Danius, Sara. *The Senses of Modernism: Technology, Perception, and Aesthetics*. Ithaca: Cornell University Press, 2002. Print.
- Darwin, Charles. *The Expression of the Emotions in Man and Animals*. 3rd Edition. New York: Oxford University. 1998. Print.
- Davidson, Cynthia C. "Introduction." *Anybody*. Ed. Cynthia C. Davidson. Cambridge, MA: The MIT Press, 1997. Print.
- Davidson, Donald. "Turing's Test." *Modelling the Mind*. Ed. K.A. Mohyeldin Said, W. H. Newton-Smith, R. Viale, K. V. Wilkes. Oxford: Clarendon Press, 1990. 1-11. Print.
- Deese, James. "Mind and Metaphor: A Commentary." *New Literary History*. 6.1 (1974):

211-217. Print.

Deleuze, Gilles. *Expressionism in Philosophy: Spinoza*. Trans. Martin Joughin. New York: Zone Books, 1990. Print.

---. *Francis Bacon: The Logic of Sensation*. Trans. Daniel W. Smith. London; New York: Continuum, 2003. Print.

---. and Félix Guattari. *Anti-Oedipus: Capitalism and Schizophrenia*. Trans. Robert Hurley, Mark Seem, Helen R. Lane. Minneapolis: University of Minnesota, 1977. Print.

---. and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Trans. Brian Massumi. Minneapolis: University of Minnesota Press, 1987. Print.

Demechenko, Eugene and Vladimir Veselov. "Who Fools Whom?: The Great Mystification, of Methodological Issues on Making Fools of Human Beings." *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts, and Grace Beber. New York: Springer, 2008. 447-459. Print.

Demos, E. Virginia. "An Affect revolution: Silvan Tomkins's affect theory." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 17-23. Print.

Déniz, O. et al. "Natural Interaction with a Robotic Head." *IWINAC*. Eds. J. Mira and J.R. Álvarez. 2007: 71-80. Print.

Dennett, Daniel C. *Brainchildren: Essays on Designing Minds*. Cambridge, MA: The MIT Press, 1998. Print.

Derrida, Jacques. "The Animal That Therefore I Am (More to Follow)." *Critical Inquiry* 28.2 (2002): 369-418. Print.

---. *The Animal That Therefore I Am*. Ed. Marie-Louis Mallet. Trans. David Wills. New York: Fordham University Press, 2008. Print.

---. "White Mythology: Metaphor in the Text of Philosophy." *Margins of Philosophy*. Trans. Alan Bass. Chicago: The University of Chicago Press, 1982. Print.

Dery, Mark. *Escape Velocity: Cyberculture at the End of the Century*. New York: Grove Press, 1995. Print.

- Descartes, René. *Meditations on First Philosophy*. Fourth Edition. Trans. Donald A. Cress. Indianapolis: Hackett Publishing Company. 1998. Print.
- Dick, Philip K. "The Android and the Human." *The Dark Haired Girl*. Willimantic, CT: Ziesing. 1988. 121-62. Print.
- . *Do Androids Dream of Electric Sheep?* New York: Random House. 1968. Print.
- . "Man, Android and Machine." *The Dark Haired Girl*. Willimantic, CT: Ziesing. 1988: 201-231. Print.
- . *We Can Build You*. New York: Vintage Books, Random House, Inc. 1994. Print.
- Dixon, Steve. "Metal Performance: Humanizing Robots, Returning to Nature, and Camping About." *Critical Digital Studies: A Reader*. Eds. Arthur Kroker and Marilouise Kroker. Toronto: University of Toronto Press, 2008. 442-475. Print.
- Dolan, David. "A New Image of the Artist." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 67. Print.
- Dolar, Mladen. *A Voice and Nothing More*. Cambridge, MA: The MIT Press, 2006. Print.
- Drozdek, Adam. "Human Intelligence and Turing Test." *AI & Society* 12 (1999): 315-321. Print.
- Druckery, Timothy. "An Itinerary and Five Excursions." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. 33-61. Print.
- Dumit, Joseph. "neuroexistentialism." *sensorium: embodied experience, technology, and contemporary art*. Ed. Caroline A. Jones. Cambridge, MA: The MIT Press, 2006. 182-189. Print.
- Dunn, Timothy. "A.I.: Artificial Intelligence and the Tragic Sense of Life." *Steven Spielberg and Philosophy: We're Gonna Need a Bigger Book*. Ed. Dean A. Kowalski. Lexington, KY: The University Press of Kentucky, 2008. 82-94. Print.
- "E3 2009 - Project Natal - Milo Demo with Peter Molyneux 720p HD." Lionhead Studio. 2 June 2009. Online video clip. YouTube.com. 18 Feb 2010.
- Edmonds, Bruce. "The Social Embedding of Intelligence: Towards Producing a Machine that Could Pass the Turing Test." *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts, and Grace Beber. New York: Springer, 2008. 211-235.

Print.

Ekman, Paul. "Silvan Tomkins and facial expression." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 209-214. Print.

---. and Wallace V. Friesen. *Unmasking the Face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975. Print.

---. Wallace V. Friesen, and Phoebe Ellsworth. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. New York: Pergamon Press Inc., 1972. Print.

Ellis, Paula M. and Joanna J. Bryson. "The Significant of Textures for Affective Interfaces." *IVA 2005*, Eds. T. Panayiotopoulos etc al.: 394-404. Print.

"Empathy." *Oxford English Dictionary*. Web. 10 Jan. 2010.

Fitzpatrick, Kathleen. "The Exhaustion of Literature: Novels, Computers, and the Threat of Obsolescence." *Contemporary Literature* 43.3 (2002): 518-559. Print.

French Robert M. "Subcognition and the Limits of the Turing Test." *Mind* 99.393 (1990): 53-65. Print.

Freud, Sigmund. "The Uncanny." *Collected Papers, Vol. IV*. Trans. Joan Rivière. London: The Hogarth Press. 1949. 368-407.

Friedman, Lester D. *Citizen Spielberg*. Urbana: University of Illinois Press, 2006. Print.

Fuss. "Introduction: Human, All Too Human." *Human, All Too Human*. Ed. Diana Fuss. New York: Routledge. 1996. Print.

Galloway, Alex. *Protocol: How Control Exists after Decentralization*. Cambridge, MA: The MIT Press, 2004. Print.

Galvan, Jill. "Entering the Posthuman Collective in Philip K. Dick's 'Do Androids Dream of Electric Sheep?'" *Science Fiction Studies* 24.3 (1997): 413-29. Print.

Gardner, Howard. *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books, Inc., 1985. Print.

Gee, F.C., W.N. Browne, and K. Kawamura. "Uncanny Valley Revisited." *2005 IEEE International Workshop on Robots and Human Interactive Communication*. 151-157.

- Geller, Tom. "Overcoming the Uncanny Valley." *IEEE Computer Graphics and Applications* (2008): 11-17.
- Genova, Judith. "Response to Anderson and Keith." *Social Epistemology* 8.4 (1994): 341-343. Print.
- . "Turing's sexual guessing game." *Social Epistemology* 8.4 (1994): 313-326. Print.
- Gerlenter, David Hillel. *The Muse in the Machine: Computerizing the Poetry of Human Thought*. New York: Free Press, 1994. Print.
- Gibson, William. "Foreword: 'The Body.'" *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. Print.
- Goodall, Jane. "The Will to Evolve." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. 1-31. Print.
- Gordon, Andrew M. *Empire of Dreams: The Science Fiction and Fantasy Films of Steven Spielberg*. Lanham, MD: Rowman & Littlefield Publishers, Inc., 2008. Print.
- Graham, Neill. *Artificial Intelligence*. Blue Ridge Summit, PA: Tab Books, 1979. Print.
- Gray, Chris Hables. "In Defence of Prefigurative Art: The Aesthetics and Ethics of Orlan and Stelarc." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 181-192. Print.
- Gray, Jesse, Cynthia Breazeal, Matt Berlin, Andrew Brooks, and Jeff Lieberman. "Action parson and goal inference using self as simulation." *2005 IEEE International Workshop on Robots and Human Interactive Communication*. 202-209. Print.
- Green, André. *The Fabric of Affect in the Psychoanalytic Discourse*. Trans. Alan Sheridan. London: Routledge, 1999. Print.
- Greenop, Matt. "E3: Project Natal – Motion Slickness." *New Zealand Herald*. Jun 2, 2009. Print.
- Grossman, Wendy M. "Do Robots Dream of Copyright?" *Guardian* September 14, 2006. Print.
- Grosz, Elizabeth. *Volatile Bodies: Toward a Corporeal Feminism*. Bloomington: Indiana University Press, 1994. Print.

- Gržinić, Marina. "Stelarc – Politics of the Body." *Stelarc: Political Prosthesis & Knowledge of the Body*. Ed. Marina Gržinić. Ljubljana, Maska 2002. 13-21. Print.
- Hacking, Ian. "The Cartesian Vision Fulfilled: Analogue Bodies and Digital Minds." *Interdisciplinary Science Reviews* 30.2 (2005): 153-166. Print.
- Halberstam, Judith. "Automating Gender: Postmodern Feminism in the Age of the Intelligent Machine." *Feminist Studies* 17.3 (1991): 439-460. Print.
- Hall, Gary. "Para-Site." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 131-146. Print.
- Hallward, Peter. "Anything Is Possible" [review of *After Finitude*]. *Radical Philosophy* 152 (2008): 51-57. Print.
- Hansen, Mark B. N. *New Philosophy for New Media*. Cambridge, MA: The MIT Press, 2004. Print.
- . "The Time of Affect, or Bearing Witness to Life." *Critical Inquiry* 30 (2004): 584-626. Print.
- Hanson, David. "Exploring the Aesthetic Range for Humanoid Robots." *Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*. 2006. 39-42. Print.
- . "Humanizing Interfaces – An Integrative Analysis of the Aesthetics of Humanlike Robots." Diss. U of Texas. Dallas. 2007.
- . "Re: a research query?" Message to the author. 29 Dec. 2009. E-mail.
- . and Andrew Olney, Steve Prilliman, Eric Mathews, Marge Zielke, Derek Hammons, Raul Fernandez, Harry Stephanou. "Upending the Uncanny Valley." *Proceedings of the 20th Natural Conference on Artificial Intelligence Vol. 4*. 2005: 1728-1729. Print.
- . and G. Pioggia, S. Dinelli, F. Di Francesco, R. Francesconi, and D. De Rossi. "Identity Emulation (IE): Bio-inspired Facial Expression Interfaces for Emotive Robots," in *AAAI Technical Report WS-02-18*: 72-81. Print.
- Haraway, Donna J. "Cyborgs and Simbionts: Living Together in the New World Order." *The Cyborg Handbook*. Ed. Chris Hables Gray. New York: Routledge, 1995. xi-xx. Print.

- . "A Game of Cat's Cradle: Science Studies, Feminist Theory, Cultural Studies." *Critical Digital Studies: A Reader*. Ed. Arthur Kroker and Marilouise Kroker. Toronto: University of Toronto Press, 2008. 45-55. Print.
- . *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge, 1991. Print.
- Hardt, Michael. "Foreword: What Affects Are Good For." *The Affective Turn: Theorizing the Social*. Ed. Patricia Ticineto Clough. Durham: Duke University Press, 2007. ix-xiii. Print.
- . and Antonio Negri. *Multitude: War and Democracy in the Age of Empire*. New York: Penguin Press. 2004. Print.
- Harman, Graham. *Guerrilla Metaphysics: Phenomenology and the Carpentry of Things*. Chicago: Open Court, 2005. Print.
- . "On Vicarious Causation." *Collapse II* (2007): 187-221. Print.
- . "Quentin Meillassoux: An New French Philosopher." *Philosophy Today* 51.1 (2007): 104-17. Print.
- Hauser, Larry. "Look Who's Moving the Goal Posts Now." *Minds and Machines* 11 (2001): 41-51. Print.
- Hayes, Patrick and Kenneth Ford. "Turing Test Considered Harmful." *Proceedings of IJCAI-95*. Montréal, Canada. (1995): 972-977. Print.
- Hayles, N. Katherine. *Chaos Bound: Orderly Disorder in Contemporary Literature and Science*. Ithaca: Cornell University Press. 1990. Print.
- . *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: The University of Chicago Press, 1999. Print.
- . "Hyper and Deep Attention: The Generational Divide in Cognitive Models." *Profession* (2007): 187-199. Print.
- Heuser, Sabine. *Virtual Geographies: Cyberpunk at the Intersection of the Postmodern and Science Fiction*. Amsterdam; New York: Rodopi. 2003. Print.
- Hodges, Andrew. *Alan Turing: The Enigma*. New York: Simon and Schuster, 1984. Print.

- Hollinger, Veronica. "Cybernetic Deconstructions: Cyberpunk and Postmodernism." in *Storming the Reality Studio: A Casebook of Cyberpunk and Postmodern Science Fiction*. Ed. Larry McCaffrey. Durham: Duke University Press. 1991. 203-218. Print.
- Horsley, Victor. "Description of the Brain of Mr. Charles Babbage, F.R.S.," *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character* 200 (1909): 117-131. Print.
- Ichbiah. *Robots: From Science Fiction to Technological Revolution*. New York: Harry N. Abrams. 2005. Print.
- Ishiguro, Hiroshi, Minoru Asada, Stuart C. Shapiro, Michael Thielscher, Cynthia Breazeal, Maja J. Mataric, and Hiroshi Ishida. "Trends and Controversies: Human-Inspired Robots." *IEEE Intelligent Systems* 21.4 (2006) 74-85. Print.
- Jameson, Fredric. *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions*. London; New York: Verso. 2007. Print.
- jeanbaptisteparis. "Leonardo" 4 Mar. 2009. Online image. Flickr. Creative Commons Attribution-Share Alike 2.0 Generic. <http://www.flickr.com/photos/jeanbaptisteparis/3328683281/>. 9 Mar. 2010.
- Johnson, Bobbie. "Technology: Newly asked questions: How exactly does Microsoft's Project Natal work?" *Guardian Technology Pages*. June 4, 2009. p. 2. Print.
- Jones, Amelia. "Stelarc's Technological 'Transcendence'/Stelarc's Wet Body." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. 87-123. Print.
- Jones, Caroline A. "The Mediated Sensorium." *sensorium: embodied experience, technology, and contemporary art*. Ed. Caroline A. Jones. Cambridge, MA: The MIT Press, 2006. 5-49. Print.
- Jones, Meredith and Zoë Sofia. "Stelarc and Orlan in the Middle Ages." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 56-72. Print.
- Keith, William. "Artificial Intelligences, Feminist and Otherwise." *Social Epistemology* 8.4 (1994): 333-340. Print.
- Kendall, Nigel. "Peter Molyneux Interview." *Times Online*. 19 March 2010. Web. 24 Mar 2010.

- Khan, A. F. Umar. "The Ethics of Autonomous Learning Systems." *Android Epistemology*. Eds. Kenneth M. Ford, Clark N. Glymour, and Patrick J. Hayes. Cambridge, MA: The MIT Press, 1995. 253-365.
- Kim, Do Hyoung, Kwang Ho An, and Myung Jin Chung. "An Artificial Emotion Imitator." *IAPR Conference on Machine Vision Applications*. MVA2005. 273-276. Print.
- Kim, Do Hyoung, Kwang Ho An, Yeon Geol Ryu, and Myung Jin Chung. "A Facial Expression Imitation System for the Primitive of Intuitive Human-Robot Interaction." *Human-Robot Interaction*. Ed. Nilanjan Sarkar. Vienna, Austria: Itech Education and Publishing, 2007. 213-234. Print.
- Kim, Do Hyoung and Hui Sung Lee and Myung Jin Chung. "Biologically Inspired Models and Hardware for Emotive Facial Expressions." *2005 IEEE International Workshop on Robots and Human Interactive Communication*. 679-685. Print.
- Kittler, Friedrich. *Gramophone, Film, Typewriter*. Trans. Geoffrey Winthrop-Young and Michael Wutz. Stanford: Stanford University Press, 1986. Print.
- Koch, Christof and Giulio Tononi. "Can Machines Be Conscious?: Yes – and a New Turing Test Might Prove It." *IEEE Spectrum* 45.6 (2008): 55-59. Print.
- Koistinen, Olli and John Biro. "Introduction." *Spinoza: Metaphysical Themes*. Eds. Olli Koistinen and John Biro. New York: Oxford University Press, 2002. Print.
- Kreider, Tim. "Review: A.I.: Artificial Intelligence." *Film Quarterly* 56.2 (2002-3): 32-39. Print.
- Kristeva, Julia. *Powers of Horror: An Essay of Abjection*. Trans. Leon S. Roudiez. New York: Columbia University Press, 1982. Print.
- Kroker, Arthur and Marilouise. "We Are All Stelarcs Now." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. 63-85. Print.
- Kurzweil, Ray and Mitchell Kapur. "A Wager on the Turing Test." *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts and Grace Beber. New York: Springer, 2008. 463-476. Print.
- Langer, R. and Vacanti, JP. "Tissue Engineering." *Science* 260. 5110 (1993): 920-926. Print.

- Lapham, Lewis H. "Introduction to the MIT Press Edition: The Eternal Now." *Understanding Media: The Extensions of Man*. Cambridge, MA: The MIT Press. 1994. Print.
- Lawlor, Leonard. "'Animals Have No Hand': An Essay on Animality in Derrida." *CR: The New Centennial Review* 7.2 (2007): 43-69. Print.
- . "Following the Rats: Becoming-Animal in Deleuze and Guattari." *SubStance* #117 37.3 (2008):169-187. Print.
- Leder, Drew. "Flesh and Blood: A Proposed Supplement to Merleau-Ponty." *Human Studies*13.3 (1990). 209-219. Print.
- Lee, Hui Sung, Do Hyoung Kim, Kwang Ho An, Jeong Woo Park, Yeon Geol Ryu, and Myung Jin Chung. "Two Types of Facial Robots for Visual Attention and Emotional Expressions." *International Conference Control, Automation, Robotics and Vision* 2004. 1553-1558. Print.
- Lee, Oo Gin. "Look ma, no controller!" *The Straits Times*. June 10, 2009. Print.
- Lenat, Douglas B. "Building a Machine Smart Enough to Pass the Turing Test: Could We, Should We, Will We?" *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts, and Grace Beber. New York: Springer, 2008. 261-282. Print.
- Lenoir, Tim. "All but War Is Simulation: The Military-Entertainment Complex." *Configurations* 8 (2000): 289-335. Print.
- . "Makeover: Writing the Body into the Posthuman Technoscape: Part One: Embracing the Posthuman." *Configurations* 10.5 (2002): 203-220. Print.
- Lévinas, Emmanuel. *Alterity & Transcendence*. Trans. Michael B. Smith. New York: Columbia University Press. 2000. Print.
- . *Humanism of the Other*. Trans. Nidra Poller. Urbana: University of Illinois Press, 2003. Print.
- . *Totality and Infinity: An Essay on Exteriority*. Trans. Alfonso Lingis. Pittsburgh: Duquesne University Press, 1969. Print.
- Lingis, Alfonso. "Translator's Preface." *The Visible and the Invisible*. Ed. Claude Lefort. Trans. Alphonso Lingis. Evanston: Northwestern University Press, 1968. Print.

The Loebner Prize. Web. 27 May 2008.

---. "1995 Loebner Prize Information." Web. 5 May 2008.

---. "2001 Loebner Prize Contest Results." Web. 27 May 2008.

---. "Alice connected to Judges." Web. 27 May 2008.

---. "Jabberwacky entry to the Loebner Prize 2005." Web. 15 Dec. 2008.

---. "Loebner Prize 2004 Rules." Web. 5 May 2008.

MacDorman, Karl F. "Introduction to the Special Issue on Android Science." *Connection Science* 18.4 (2006): 313-317. Print.

---. "Mortality Salience and the Uncanny Valley." *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots*. December 5-7, 2005. Tsukuba, Japan. 399-405. Print.

---. "Subjective Ratings of Robot Video Clips for Human Likeness, Familiarity, and eeriness: An Exploration of the Uncanny Valley." *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*. July 26, 2006. Vancouver, Canada. 48-51. Print.

---. and Robert D. Green, Chin-Chang Ho, Clinton T. Koch. "Too Real for Comfort? Uncanny responses to Computer Generated Faces." *Computers in Human Behavior* 25 (2009): 695-710. Print.

---. and Hiroshi Ishiguro. "Opening Pandora's uncanny box: Reply to commentaries on 'The uncanny advantage of using androids in social and cognitive science research.'" *Interaction Studies* 7.3 (2006): 361-368. Print.

---. and Hiroshi Ishiguro. "The Uncanny Advantage of Using Androids in Cognitive and Social Science Research." *Interaction Studies* 7.3 (2006): 297-337. Print.

---. and Hiroshi Ishiguro. "The Uncanny Advantage of Using Androids in Cognitive and Social Science Research." *Interaction Studies* 7.3 (2006): 297-337. Print.

Mackay, Robin. "Editorial Introduction." *Collapse II* (2007): 3-13. Print.

Mackenzie, Adrian. *Transductions: Bodies and Machines at Speed*. London; New York: Continuum, 2002. Print.

Manning, Erin. *Politics of Touch: Sense, Movement, Sovereignty*. Minneapolis:

University of Minnesota Press, 2007. Print.

Marcus, Aaron et al. "Sci-Fi at CHI: Cyberpunk Novelists Predict Future of Interfaces." *CHI'92*, May 3-7: 435-7. Print.

Marino, Mark C. "Review: *The Electronic Literature Collection Volume 1: A New Media Primer*." *DHQ: Digital Humanities Quarterly* 2.1 (2008): n.pag. Web. 11 Nov. 2009.

Massumi, Brian. "Notes on the Translation and Acknowledgments." *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minneapolis Press, 1987. Print.

---. *Parables for the Virtual: Movement, Affect, Sensation*. Durham: Duke University Press, 2002. Print.

McCaffery, Larry et al. "Cyberpunk Forum/Symposium." *Mississippi Review* 16.2/3 (1988): 16-65. Print.

McLuhan, Eric. "Preface." *Laws of Media: The New Science*. Toronto: University of Toronto Press, 1988, vii-xi . Print.

McLuhan, Marshall. "Art as Survival in the Electric Age (1973)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 206-224. Print.

---. "Cybernetics and Human Culture (1964)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 44-55. Print.

---. "Electronic Revolution: Revolutionary Effects of New Media (1959)." *Understanding Me: Lectures and Interviews*. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 1-11. Print.

---. "Fordham University: First Lecture (1967)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 139-146. Print.

---. "The Future of Man in the Electric Age (1965)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 56-75. Print.

---. "The Future of the Book (1972)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003.

- 173-186. Print.
- . *The Gutenberg Galaxy: The Making of Typographic Man*. Toronto: University of Toronto Press, 1962. Print.
- . "Introduction to the Second Edition." *Understanding Media: The Extensions of Man*. New York: New American Library, 1964. Print.
- . "The Invisible Environment: The Future of an Erosion." *Perspecta* 11 (1967): 163-167. Print.
- . "Man and Media (1979)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 277-298. Print.
- . "The Marfleet Lectures (1967)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 103-138. Print.
- . "McLuhan's Laws of the Media." *Technology and Culture* 16.1 (1975): 74-78. Print.
- . "The Medium Is the Massage (1966)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 76-97. Print.
- . "Open-Mind Surgery (1967)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 147-157. Print.
- . "Predicting Communication via the Internet (1966)." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 98-102. Print.
- . "Popular/Mass Culture: American Perspectives (1960)" in *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. 12- 43. Print.
- . "Roles, Masks, and Performances." *New Literary History* 2.3 (1971): 517-531. Print.
- . "Speed of Cultural Change." *College Composition and Communication* 9.1 (1958): 16-20. Print.
- . *Understanding Media: The Extensions of Man*. Cambridge, MA: The MIT Press, 1994. Print.

- . *Understanding Media: The Extensions of Man*. Second Edition. New York: McGraw-Hill Book Company, 1964. Print.
- . and Quentin Fiore. *The Medium is the Massage: An Inventory of Effects*. San Francisco: HardWired, 1967. Print.
- . and Eric McLuhan. *Laws of Media: The New Science*. Toronto: University of Toronto Press, 1988. Print.
- . and Barrington Nevitt. "The Argument: Causality in the Electric World." *Technology and Culture* 14.1 (1973): 1-18. Print.
- McCorduck, Pamela. *Machines Who Think*. San Francisco: W. H. Freeman and Company, 1979. Print.
- McHale, Brian. "Elements of a Poetics of Cyberpunk." *Critique* 33.3 (1992): 149-175. Print.
- Means, Loren. "Interview with Rudy Rucker." *YLEM Journal* 25.10 and 12: 4-8. Print.
- "Science Fiction and Its Discontents." *YLEM Journal* 25.10 and 12: 2, 17, 30. Print.
- Meillassoux, Quentin. *After Finitude: An Essay on the Necessity of Contingency*. Trans. Ray Brassier. London; New York: Continuum. 2008. Print.
- . "Spectral Dilemma." *Collapse IV* (2008): 261-75. Print.
- Menzel, Peter and Faith D'Aluisio. *Robo sapiens: Evolution of a New Species*. Cambridge, MA: The MIT Press, 2001. Print.
- Merleau-Ponty, Maurice. *Phenomenology of Perception*. Trans. Colin Smith. London: Routledge, 1999. Print.
- . "The Primacy of Perception and Its Philosophical Consequences." Trans. James M. Edie. *The Primacy of Perception and Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics*. Ed. James M. Edie. Evanston: Northwestern University Press, 1964. 12- 42. Print.
- . *The Visible and the Invisible*. Ed. Claude Lefort. Trans. Alphonso Lingis. Evanston: Northwestern University Press, 1968. Print.
- "Metaphor." *The Oxford English Dictionary*. Web. 5 Dec. 2008.

- Millar, P.H. "On the Point of the Imitation Game." *Mind* 82.328 (1973): 595-597. Print.
- Minsky, Marvin. "Our Robotized Future." *Robotics*. Ed. Marvin Minsky. Garden City, N.Y.: Anchor Press/Doubleday, 1985. 304. Print.
- "Philip K. Dick robot." Mirka23. 25 July 2006. Online image. Flickr. Creative Commons Attribution-Noncommercial 2.0 Generic.
<http://www.flickr.com/photos/mirka23/198619829/>. 10 Mar. 2010.
- Monfort, Nick. *Twisty Little Passages: An Approach to Interactive Fiction*. Cambridge, MA: The MIT Press, 2005. Print.
- Moor, James H. "The Status and Future of the Turing Test." *Minds and Machines* 11 (2001): 77-93. Print.
- Moravec, Hans P. *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press, 1988. Print.
- . *Robot: Mere Machine to Transcendent Mind*. New York: Oxford University Press, 1999. Print.
- Mori, Masahiro. "The Uncanny Valley." Trans. Karl F. MacDorman and Takashi Minato. *Energy* 7.4 (1970): 33-35. Print.
- . *The Buddha in the Robot*. Tokyo: Kosei Publishing Company, 1989. Print.
- Morris, Nigel. *The Cinema of Steven Spielberg: Empire of Light*. London; New York, Wallflower Press, 2007. Print.
- Morrissey, Thomas. "Growing Nowhere: Pinocchio Subverted in Spielberg's *A.I.: Artificial Intelligence*." *Extrapolation* 45.3 (2004): 249-262. Print.
- Murray, Timothy. "Artistic Simulacra in the Age of Recombinant Bodies." *Literature and Medicine* 26.1 (2007): 159-179. Print.
- Naur, Peter. "Thinking and Turing's Test." *BIT* 26 (1986): 175-187. Print.
- Negri, Antonio and Michael Hardt. "Value and Affect." *boundary 2* 26.2 (1999): 77-88. Print.
- Newman, M. H. A., Alan M. Turing, Sir Geoffrey Jefferson, and R. B. Braithwaite. "Can

- Automatic Calculating Machines Be Said to Think?" *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*. Ed. Stuart Shieber. Cambridge, MA: The MIT Press, 2004. 117-132. Print.
- Nilsson, Nils J. *Shakey the Robot*. Technical Note 323. AI Center, SRI International, Menlo Park, CA. Apr 1984. Print.
- North, Ian. "Link Exhibition: Catalogue Statement." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 66. Print.
- Olivier, Bert. "When Robots Would Really Be Human Simulacra: Love and the Ethical in Spielberg's *AI* and Proyas's *I, Robot*." *Film-Philosophy* 12.2 (2008): 30-44. Print.
- Paffrath, James D. "An Interview with Stelarc." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 16-17. Print.
- Palmer, Christopher. *Philip K. Dick: Exhilaration and Terror of the Postmodern*. Liverpool: Liverpool University Press. 2003. Print.
- Panagia, Davide. *The Political Life of Sensation*. Durham: Duke University Press, 2009. Print.
- Peek, Nadya. "Kismet is momentarily lifeless." 11 Mar. 2008. Online image. Flickr. Creative Commons Attribution 2.0 Generic. <http://www.flickr.com/photos/nadya/2325676697/>. 9 Mar. 2010.
- Pentland, Alex. "Socially Aware Computation and Communication." *IEEE*, March (2005): 33-40. Print.
- Picard, Rosalind W. *Affective Computing*. Cambridge, MA: The MIT Press, 1997. Print.
- . "Affective computing: challenges," in *International Journal of Human-Computer Studies* 59 (2003): 55-64. Print.
- . S. Papert, W. Bender, B. Blumberg, C. Breazeal, D. Cavallo, T. Machover, M. Resnick, D. Roy, and C. Strohecker. "Affective learning – a manifesto." *BT Technology Journal* 22.4 (2004): 253-269. Print.
- Pickering, Andrew. "Cybernetics and the Mangle: Ashby, Beer and Pask." *Social Studies of Science* 32.3 (2003): 413-437. Print.

- . "Emergence and Synthesis: Science Studies, Cybernetics and Antidisciplinarity." *Technoetic Arts: A Journal of Speculative Research* 6.2 (2008): 127-133. Print.
- Poster, Mark. "High-Tech Frankenstein, or Heidegger Meets Stelarc." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 15-32. Print.
- Powers, Richard. *Galatea 2.2: A Novel*. New York: Picador, 1995. Print.
- Project Natal. *Xbox.com*. 22 Mar. 2010. Web.
<http://www.xbox.com/en-US/live/projectnatal/>.
- Putrill, Richard L. "Beating the Imitation Game." *Mind* 80.318 (1971): 290-294. Print.
- Rajan, Tilottama. "Trans-Positions of Diference: Kristeva and Post-structuralism." *Ethics, Politics, and Difference in Julia Kristeva's Writing*. Ed. Kelly Oliver. New York: Routledge. 1993. 215-237. Print.
- Rancière, Jacques. *The Politics of Aesthetics: The Distribution of the Sensible*. Trans. Gabriel Rockhill. London; New York: Continuum, 2004. Print.
- Ricoeur, Paul. "The Metaphor Process as Cognition, Imagination, and Feeling." *Critical Inquiry* 5.1 (1978): 143-159. Print.
- Rizzolatti, Giacomo and Laila Craighero. "The Mirror-Neuron System." *Annual Review in Neuroscience* 27 (2004): 169-192. Print.
- Robnik, Drehli. "Saving One Life: Spielberg's *Artificial Intelligence* as Redemptive Memory of Things." *Jump Cut: A Review of Contemporary Cinema* 45 (2002). Web. 12 Feb. 2010.
- Rogers, Carl R. *Client-Centered Therapy: Its Current Practice, Implications, and Theory*. Boston: Houghton Mifflin Company, 1951.
- . *On Becoming a Person: A Therapist's View of Psychotherapy*. Boston: Houghton Mifflin Company, 1961.
- Roland, Alex and Philip Shiman. *Strategic Computing: DARPA and the Quest for Machine Intelligence, 1983-1993*. Cambridge, MA: The MIT Press, 2002. Print.
- Rucker, Rudy. "37 Questions." *Seek!*. New York: Four Walls Eight Windows, 1999. 8-30. Print.
- . *Artificial Life Lab*. Corte Madera, CA: The Waite Group. 1993. Print.

- . *Software*. New York: Avon Books, 1982. Print.
- . "What Is Cyberpunk?" in *Seek!*. New York: Four Walls Eight Windows, 1999. 315-322. Print.
- Rutherford, Donald. *Leibniz and the Rational Order of Nature*. New York. Cambridge University Press. 1995. Print.
- Sack, Warren. "Artificial Intelligence and Turing's Imitation Game." April 2003.
- Saltzman, Marc. "Recharged E3 draws plenty of star power." *The Toronto Star*. June 6, 2009. E01. Print.
- Sampson, Geoffrey. "In Defence of Turing." *Mind* 82.328 (1973): 592-594. Print.
- Scheer, Edward. "What Does an Avatar Want? Stelarc's E-motions." *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylińska. London; New York: Continuum, 2002. 81-100.
- Schultz, Alan C., Cynthia Breazeal, Terry Fong, and Sara Kiesler. "The Second International Conference on Human-Robot Interaction." *AI Magazine* (2007): 92-93. Print.
- Searle, John. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3 (1980): 417-457. Print.
- Sedgwick, Eve Kosofsky. *Touching Feeling: Affect, Pedagogy, Performativity*. Durham: Duke University Press, 2003. Print.
- . and Adam Frank. *Shame and Its Sisters: A Silvan Tomkins Reader*. Durham: Duke University Press, 1995. Print.
- Serres, Michel. *The Five Senses: A Philosophy of Mingled Bodies*. Trans. Margaret Sankey and Peter Cowley. London: Continuum International Publishing Group, 2008. Print.
- Shanahan, Murray. "Reinventing Shakey." *AAAI Technical Report FS-98-02* (1998): 125-135. Print.
- Shanon, Benny. "Metaphor: From Fixedness and Selection to Differentiation and Creation." *Poetics Today* 13.4 (1992): 659-685. Print.
- Shieber, Stuart. "Early Harbingers of Later Issues." *The Turing Test: Verbal Behavior as*

- the Hallmark of Intelligence*. Ed. Stuart Shieber. Cambridge, MA: The MIT Press, 2004. Print.
- . "Immediate Responses." *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*. Ed. Stuart Shieber. Cambridge, MA: The MIT Press, 2004. 135-139. Print.
- . "The Spark of Intentionality." *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*. Ed. Stuart Shieber. Cambridge, MA: The MIT Press, 2004. 199-200. Print.
- Short, Emily. *Galatea*, in *Electronic Literature Collection Volume One*. Ed. N. Katherine Hayles, Nick Monfort, Scott Rettberg, Stephanie Strickland. Creative Commons 2.5 License 2006. CD-ROM.
- Shouse, Eric. "Feeling, Emotion, Affect." *M/C Journal* 8.6 (2005). Web. 12 Jan. 2009.
- Smith, M. Brewster. "Introduction." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 1-12. Print.
- Smith, Marquard. "Preface: The Conception of *Stelarc: The Monograph*." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. Print.
- Sobchack, Vivian. "Love machines: boy toys, toy boys and oxymorons of *A.I.: Artificial Intelligence*." *Science Fiction Film and Television* 1.1 (2008): 1-13. Print.
- Sokal, Alan and Jean Bricmont's *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: Picador, 1988. Print.
- Sosnowski, Stefan and Ansgar Bittermann, Kolja Kühnlenz, and Martin Buss. "Design and Evaluation of Emotion-Display EDDIE." *Proceedings of the 2006 IEEE/R SJ International Conference on Intelligent Robots and Systems*. October 9-15, 2006, Beijing, China. 3113-3118. Print.
- Spexard, Thorsten P., Marc Hanheide, Gerhard Sagerer, "Human-Oriented Interaction with an Anthropomorphic Robot," in *IEEE Transactions on Robotics*, vol . 23, no. 5, October 2007, pp. 852-862. Print.
- Spielberg, Steven dir. *A.I. Artificial Intelligence*. Warner Bros. Pictures. 2001. DVD.
- Spinoza, Benedict. *Ethics*. Trans. R. H. M. Elwes. Amherst: Prometheus Books, 1989. Print.

- Spivey, Michael J. "Turning the Tables on the Turing Test: The Spivey Test." *Connection Science* 12.1 (2000): 91-94. Print.
- Stein, Lynn Andrea. "Imagination and Situated Cognition." A.I. Memo No. 1277 (February 1991). Print.
- Stelarc, "Extra Ear: Alternate Anatomical Architectures 2007," Brunel University School of Arts. Print.
- . "The Myth of Information." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 24. Print.
- . "Parasite Visions: Alternate, Intimate and Involuntary Experiences." *Stelarc: Political Prosthesis & Knowledge of the Body*. Ed. Marina Gržinić. Ljubljana, Maska 2002. 13-21. Print.
- . "The Splitting of the Species." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 134. Print.
- . "Strategies and Trajectories." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 76. Print.
- . "Triggering an Evolutionary Dialectic." *Obsolete Body/Suspensions/Stelarc*. Ed. James D. Paffrath with Stelarc. Davis, CA: JP Publications, 1984. 52. Print.
- . and Marquard Smith. "Animating Bodies, Mobilizing Technologies: Stelarc in Conversation." *Stelarc: The Monograph*. Ed. Marquard Smith. Cambridge, MA: The MIT Press, 2005. 215-241. Print.
- Sterrett, Susan. "Nested Algorithms and 'The Original Imitation Game Test': A Reply to James Moor." *Minds and Machines* 12 (2002): 131-136. Print.
- . "Too Many Instincts: Contrasting Philosophical Views on Intelligence in Humans and Non-Humans." *Journal of Experimental & Theoretical Artificial Intelligence* 14 (2002): 39-60. Print.
- . "Turing's Two Tests for Intelligence," *Minds and Machines* 10 (2000): 541-559. Print.
- Sterritt, David. "Spielberg, iconophobia, and the mimetic uncanny." *New Review of Film and Television Studies* 7.1 (2009): 51-65. Print.
- Stiehl, Walter Dan, Levi Lalla, and Cynthia Breazeal. "A 'Somatic Alphabet' Approach

- to ‘Sensitive Skin.’” *Proceedings of the 2004 IEEE International Conference on Robotics & Automation*. New Orleans, LA. April 2004. 2865-2870. Print.
- Suchman, Lucy. *Human-Machine Reconfigurations: Plans and Situated Actions*, 2nd Edition. New York: Cambridge University Press, 2007.
- Sullins, J. P. “Friends by Design.” *Philosophy and Design: From Engineering to Architecture*. Eds. Pieter E. Vermaas, Peter Kroes, Andrew Light, Steven A. Moore. New York: Springer, 2008. 143-157. Print.
- Surin, Kenneth. “Force.” *Gilles Deleuze: Key Concepts*. Ed. Charles J. Stivale. Montréal & Kingston: McGill-Queen’s University Press. 2005. 19-30. Print.
- Sutin, Lawrence. *Divine Invasions: A Life of Philip K. Dick*. New York: Harmony Books, 1989. Print.
- Swanson, Don R. “Toward a Psychology of Metaphor.” *Critical Inquiry* 5.1 (1978): 163-166. Print.
- “Sympathy.” *Oxford English Dictionary*. Web. 10 Jan 2010.
- Thomaz, Andrea Lockerd, Matt Berlin, and Cynthia Breazeal. “An Embodied Computational Model of Social Referencing.” *2005 IEEE International Workshop on Robots and Human Interactive Communication*. 591-598. Print.
- . “Robot Science Meets Social Science: An Embodied Computational Model of Social Referencing.” *Cognitive Science Society*. July 25-26, 2005. Stresa, Italy. 7-17. Print.
- Tinwell, A. and M. Grimshaw. “Bridging the Uncanny: An Impossible Traverse?” *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*. Tampere, Finland (2009): 66-73. Print.
- Tomkins, Silvan. *Affect Imagery Consciousness, Volume I – The Positive Affects*. New York: Springer Publishing Company, Inc., 1962. Print.
- . *Affect Imagery Consciousness, Volume III – The Negative Affects: Anger and Fear*. New York: Springer Publishing Company, 1991. Print.
- . “Evolution of the Affect System,” in *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 64-67. Print.
- . “Ideology and Affect.” *Exploring Affect: The Selected Writings of Silvan S. Tomkins*.

- Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 109-167. Print.
- . "Modifications in the theory – 1978." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 86-96. Print.
- . "The Phantasy behind the Face." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 263-278. Print.
- . "The Quest for Primary Motives." *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 27-63. Print.
- . "What and Where Are the Primary Affects?" *Exploring Affect: The Selected Writings of Silvan S. Tomkins*. Ed. E. Virginia Demos. Cambridge: Cambridge University Press, 1995. 217-262. Print.
- Tsakiris, Manos. "Looking for Myself: Current Multisensory Input Alters Self-Face Recognition." *PLoS ONE* 3.12, (2008): 1-6. Print.
- . "Re: a question about self-face recognition." Message to the author. 18 May 2009. E-mail.
- . and Patrick Haggard. "Experimenting with the Acting Self." *Cognitive Neuropsychology* 22.3/4 (2005): 387-407. Print.
- . and Patrick Haggard. "The Rubber Hand Illusion Revisited: Visuotactile Integration and Self-Attribution." *Journal of Experimental Psychology: Human Perception and Performance* 31.1 (2005): 80-91. Print.
- . et al. "Having a body *versus* moving your body: How agency structure body-ownership." *Consciousness and Cognition* 15 (2006): 423-432. Print.
- . et al. "Neural Signatures of Body Ownership: A Sensory Network for Bodily Self-Consciousness." *Cerebral Cortex* 17 (2007): 2235-2244. Print.
- Turing, Alan. "Computing Machinery and Intelligence." *Mind* 59.236 (1950): 433-460. Print.
- Turkle, Sherry. *The Second Self: Computers and the Human Spirit*. New York: Simon and Schuster, 1984. Print.

- Wallace, Richard S. "The Anatomy of A.L.I.C.E." *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Eds. Robert Epstein, Gary Roberts, and Grace Beber. New York: Springer, 2008. 181-210. Print.
- Wardrip-Fruin, Noah. *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. Cambridge, MA: The MIT Press, 2009. Print.
- Waxman, Sharon. "A Strange Loss of Face, More Than Embarrassing," *New York Times*, June 24, 2006. Print.
- we-make-money-not-art. "IMG_6363." 1 Feb. 2007. Online image. Flickr. Creative Commons Attribution-Share Alike 2.0 Generic. <http://www.flickr.com/photos/nearfuture/376891679/19>. Mar. 2010.
- Wegenstein, Bernadette. "Getting Under the Skin, or How Faces Have Become Obsolete." *Configurations* 10 (2002): 221-259. Print.
- Weizenbaum, Joseph. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W. H. Freeman, 1976. Print.
- . "ELIZA – A Computer Program from the Study of Natural Language Communication between Man and Machine." *Communications of the ACM* 9.1 (1966): 36-45. Print.
- White, V. Alan. "A.I.: Artificial Intelligence: Artistic Indulgence or Advanced Inquiry?" in *Steven Spielberg and Philosophy: We're Gonna Need a Bigger Book*. Ed. Dean A. Kowalski. Lexington, KY: The University Press of Kentucky, 2008. 210-226. Print.
- Wiener, Norbert. *Cybernetics: or Control and Communication in the Animal and the Machine*. 2nd edition. Cambridge, MA: The MIT Press, 1961. Print.
- . *The Human Use of Human Beings: Cybernetics and Society*. Garden City, NY: Doubleday & Company, Inc., 1954. Print.
- Wolfe, Tom. "Foreword." *Understanding Me: Lectures and Interviews*. Ed. Stephanie McLuhan and David Staines. Toronto: McClelland & Stewart, 2003. xxv-xxvi. Print.
- Zdenek, Sean. "Passing Loebner's Turing Test: A Case of Conflicting Discourse Functions." *Minds and Machines* 11 (2001): 53-76. Print.
- Zylinska, Joanna. "Extending McLuhan into the New Media Age: An Introduction." *The*

Cyborg Experiments: The Extensions of the Body in the Media Age. Ed. Joanna Zylynska. London; New York: Continuum, 2002. 1-12. Print.

---. “‘The Future... Is Monstrous’: Prosthetics as Ethics.” in *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Ed. Joanna Zylynska. London; New York: Continuum, 2002. 214-236. Print.

---. “Of Swan and Ugly Ducklings: Bioethics between Humans, Animals, and Machines.” *Configurations* 15 (2007): 125-150. Print.

---. and Gary Hall. “Probing: An Interview with Stelarc.” *The Cyborg Experiments: The Extensions of the Body in the Media Age*. Joanna Zylynska. London; New York: Continuum, 2002. 114-130. Print.

Biography

Jennifer S. Rhee, born November 1, 1975 in Barberton, Ohio and received her Bachelor's of Arts in English from Princeton University in 1998. She is co-editor of *Electronic Techtonics: Thinking at the Interface, the Proceedings of the First International HASTAC Conference* (Lulu.com, 2008) and the author of "Time Embodied: The Lived Body in On Kawara's *Date Paintings*" in *Thresholds* 31: Ephemera (2006). She has an article on Richard Powers' *Galatea 2.2* and Joseph Weizenbaum's ELIZA under review, and is currently working on an article, provisionally entitled "Double Singularity," that looks at Rudy Rucker's *Postsingular* through two concepts of singularity – the philosophical and the technological posthuman. She has been the recipient of generous financial support from the Program in Literature, as well as a graduate fellowship from INTERFACE, the John Hope Franklin Humanities Institute's Franklin Seminar in the Humanities. She has also been awarded the Duke Graduate School summer research fellowship, the Graduate School's Conference Travel Fellowship, a tuition scholarship for the University of California Humanities Research Institute's Seminar in Experimental Critical Theory, technoSpheres:FutureS of Thinking. She is a member of the American Comparative Literature Association; Humanities, Arts, Science, and Technology Advanced Collaboratory; the Modern Language Association; the Society for Literature, Science, and the Arts.