THE ANDROID AND OUR CYBORG SELVES:

WHAT ANDROIDS WILL TEACH US

ABOUT BEING (POST)HUMAN

By

ANTONIE MARIE BODLEY

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHHY

WASHINGTON STATE UNIVERSITY
The Graduate School

MAY 2015

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of ANTONIE MARIE BODLEY find it satisfactory and recommend that it be accepted.

_____
Joseph K. Campbell, Ph.D., Chair

_____
Jon Hegglund, Ph.D.

_____
Pamela Thoma, Ph.D.

_____
Gene Apperson

ACKNOWLEDGEMENT

My deepest thanks go out to my committee for their continuing support and encouragement. Dr. Joe Campbell, Dr. Jon Hegglund, Dr. Pamela Thoma and Gene Apperson, all deserve my sincerest gratitude. Dr. Campbell, my committee chair, and Dr. Jon Hegglund were major supporters of my ideas from the very beginning. Without their confidence in my work beginning as early as ten years ago, this dissertation never would have been possible. Dr. Thoma helped show me that strong voices should be encouraged, no matter how quiet they may seem. From Mr. Apperson I was reminded that the fantastic imagination in science fiction generates real science and dreams for the future of robotics.

I would like to thank other faculty and staff in the Graduate School who helped make this possible. In particular, Dr. Pat Sturko: without her friendship and professional guidance, I never would have completed this.

Also foundational in the completion of this dissertation are my friends and family. Without their feedback and continuing interest in my work, I would not have been able to generate the ideas, the perspective, and the creative telling of my work. I could never have asked for a more interesting and intellectually curious group of friends to ask questions and quiz my knowledge about science and science fiction.

Lastly, to my husband, Garrett: my deepest thanks for your patience and enduring support. It is only with you that I could enjoy date night watching *RoboCop*.

THE ANDROID AND OUR CYBORG SELVES:

WHAT ANDROIDS WILL TEACH US

ABOUT BEING (POST)HUMAN

Abstract

by Antonie Marie Bodley, Ph.D.
Washington State University
May 2015

Chair: Joseph K. Campbell

In the search for understanding a future for our selves with the potential merging of strong Artificial Intelligence and humanoid robotics, this dissertation uses the figure of the android in science fiction and science fact as an evocative object. Here, I propose android theory to consider the philosophical, social, and personal impacts humanoid robotics and AI will have on our understanding of the human subject. From the perspective of critical posthumanism and cyborg feminism, I consider popular culture understandings of AI and humanoid robotics as a way to explore the potential effect of androids by examining their embodiment and disembodiment. After an introduction to associated theories of humanism, posthumanism, and transhumanism, followed by a brief history of the figure of the android in fiction, I turn to popular culture examples. First, using two icons of contemporary AI, Deep Blue, a chess playing program and Watson, a linguistic artificially intelligent program, I explore how their public performances in games evoke rich discussion for understanding a philosophy of mind in a non-species specific way. Next, I turn to the *Terminator* film series (1984-2009) to discuss how the humanoid embodiment of artificial intelligence exists in an uncanny position for our emotional attachments to nonhuman entities. Lastly, I ask where these relationships will take us in our intimate lives; I explore personhood and human-nonhuman relationships in what I call the

nonhuman dilemma. Using the human-Cylon relationships in the reimagined *Battlestar Galactica* television series (2003-2009), the posthuman family make-over in the film *Fido* (2006), as well as a real-life story of men with their life-sized doll companions, as seen in the TLC reality television series *My Strange Addiction* (2010), I explore the coming dilemma of life with nonhuman humanoids.

TABLE OF CONTENTS

LIST OF FIGURES

Dedication

This work is dedicated to all human(oid)s of the future.

May you continue to be inspired by science and science fiction.

The Android and our Cyborg Selves:
What Androids Will Teach Us about Being (Post)Human
By Antonie Marie Bodley

CHAPTER ONE: INTRODUCTION TO ANDROID THEORY

> "We ask of the computer not just about where we stand in nature, but about where we stand in the world of the artifact. We search for a link between who we are and what we have made, between who we are and what we might create, between who we are and what, through our intimacy with our own creations, we might become."
>
> -- Sherry Turkle, *The Second Self: Computers and the Human Spirit* (1984).

INTRODUCTION

In 1984 Sherry Turkle began her search for a link between ourselves and our creations; between ourselves and the artifact in *The Second Self: Computers and the Human Spirit*. She used the computer as her point of interest, as her "evocative object." As a child of the eighties, someone that could have easily been part of her studies of children and technology, I want to bring that reflection back into focus again, but with the figure of the android as my evocative object. While her work took her to explore those connections between and among people through technology, in particular computers and later mobile devices, this work seeks that link through fiction in the singular image of the android housed with strong Artificial Intelligence.

In the search for understanding what the future will be like for our *selves* and our species, it is important to find some solid ground, a stable perspective upon a subject/object that we can use for investigation and extrapolation. Here the figure of the android functions as such a focus. Currently under development in many scientific fields, from humanoid robotics to coding the artificial intelligence, the android is a rich subject for discussion because it has been in our

cultural imagination for decades. Here, the android becomes my evocative object, an object poised upon the edge of previously stable boundaries. For Turkle, "Theory enables us, for example, to explore how everyday objects become part of our inner life: how we use them to extend the reach of our sympathies by bringing the world within" (*Evocative Objects* 307). The goal of examining evocative objects is to "defamiliarize" ourselves from the objects to help bring them back into focus in a way that matters to our inner self.

Soon, androids will be part of our everyday lives and they will have a profound effect on our inner self and our homes. I propose here to explore those potential effects by examining the embodiment and disembodiment of the android through contemporary popular culture examples. After an introduction to associated theories of humanism, posthumanism, and transhumanism, followed by a brief history of the figure of the android in fiction, I will turn to examples. First, using Watson, a linguistic artificially intelligent program, I explore how his performance on the television gameshow *Jeopardy!* evokes rich discussion for understanding a philosophy of mind in a non-species specific way. Next, I turn to the *Terminator* film series (1984-2009) to discuss how the humanoid embodiment of artificial intelligence exists in an uncanny position for our emotional attachments to nonhuman entities. Lastly, I ask where these relationships will take us in our intimate lives; I explore personhood and human-nonhuman relationships using the human-Cylon relationships in the reimagined *Battlestar Galactica* television series (2003-2009) as well as a real-life story of men with their life-sized doll companions, seen in the TLC reality television series *My Strange Addiction* (2010).

In fiction, stories of androids, from the functional to the social, are at first examples of queer bodies and I consider these evocative and queer bodies. For Donna Haraway, "Queering has the job of undoing 'normal' categories, and none is more critical than the human/nonhuman

sorting operation" (xxiv). The body of the android undoes the category of human/nonhuman simply by being neither human in construction nor in acceptance and therefore "queer" may be an appropriate choice. Androids are also not fully *non*human as they are designed to "fit" in with humanity in a way that is more comfortable than other machines. In fact, this attempt to "pass" within the populations of humanity, suggests another aspect of androids' queerness which reaches into theories of mind, body and society.

ANDROID THEORY

Drawing from the fields of American Studies, Film Studies, Philosophy, and Literary Studies, I use a multidisciplinary approach to explore the android from two fronts – both the theory from fiction surrounding the android and the actual, literal development. I seek answers in theories of posthumanity and explorations of the cyborg. Along with these theoretical perspectives, I turn to contemporary currents in transhumanist philosophies. With Future Studies blossoming as an academic discipline in think-tanks like the Singularity University and the Future of Humanity Institute, it would be detrimental to ignore the actual development of androids and AI in this discussion. I will explore links among these fields with three primary foci: the mind, the body and society, each with a respective chapter. I choose this discussion now, not simply because of my love and fascination for science fiction, but also because we are on the threshold of an entirely new way of living which includes artificial entities in very human-like form.

In cultural studies, the posthuman is explored exclusively with the human figure as the focus, but I propose a shift from the human to the android – both the fictional creature and the actual creation. Using what I call "Android Theory," this research seeks to explore the figure of the android to form a vocabulary that can extrapolate to a future allowing for a (post)human self,

able to live and interact within communities of actual androids and other potential entities that we cannot even imagine at this time. The figure of the android will be addressed as both a literal entity existing in the form of humanoid robotics, and as a figurative entity found in fiction. In this exploration I hope to find that the human is not in crisis by the boundary blending generally proposed by concepts like cyborg theory. Rather, we are opening ourselves up to the new articulations that Judith Butler describes in *Undoing Gender*: "[it is necessary to keep] our very notion of the 'human' open to a future articulation" because that openness is "essential to the project of a critical international human rights discourse and politics" (222).

CRITICAL POSTHUMANISM

Some of the keywords introduced so far for this project include "transhumanism," "posthumanism," and "cyborg." Each of these requires some explanation before fitting within a discussion of science fiction and androids. Transhumanism generally refers to a philosophy, a world-view regarding the direction of technology development and the nature of the human condition, often associated with the Extropian movement (More, Relke, Wolfe). Posthumanism can be described as both a literal entity as part of the future of the transhumanist philosophy but also a theoretical framework within cultural theory. While transhumanism and posthumanism can be "cousins" of sorts, they both "[invite] a critical riposte from a position distinct from speculative posthumanism or transhumanism: critical posthumanism" (Roden 29). Bart Simon describes this confusion by positioning one as "popular posthumanism, or transhumanism" and the other as "critical posthumanism," with the phrasing attributed to Jill Didur (2). Both the popular post/transhumanisms of Extropian thought and the critical response to such thinking offer rich collections of work surrounding who and what we will potentially become in the future. Representing the transhumanists are writers such as Max More, Nick Bostrom, and Ray

Kurzweil, to name a few,[1] although their views on how to approach the future are very different. Some, like Bostrom, are sounding warnings while others, like Kurzweil, promise the coming panacea from the bounties of science. Responding to fiction and philosophy, the critical posthumanists are often represented by scholars like Donna Haraway, Katherine Hayles, and Cary Wolfe.

A person who buys in to the "transhumanist philosophy" is a person who believes in improving the human condition – including the body, self and society – through technological enhancement. There are many views surrounding exactly how most ideally to reach such goals is. From biological and genomic enhancement to cybernetic appendages and exploration into Artificial Life (biologically based, computationally-based, or some other base that we have not yet conceived of), the Transhumanist believes that what defines "human" is always in a state of flux. In general, the transhumanist philosophy endorses the belief that what makes us "human" is in a developmental state toward what could be called "posthuman." In this transhumanist philosophy, to *become* posthuman is to be effectively enhanced through technological means so that they have surpassed the "limitations that define the less desirable aspects of the 'human condition'" (More 4).

Posthuamanism as part of a transhumanist philosophy has a complex history, some of which grew out of the celebration of humanism and a return to a romanticized vision of technology from the Renaissance era. Humanism, as described by Brian Cooney, is "One of the Great Ideas western culture inherited from the classical Greeks" (xx-xxi). This "religious humanism" idealized traits that were distinctively human, one trait of which was the ability to be

[1] Relke asks, "is it any wonder [that] Extropianism, with its relentlessly optimistic focus on the future, is increasingly popular among techno-savvy young men?" (81). She reminds us that the Extropian movement, and other transhumanists, are typically white, privileged men… Suggesting that the future will only be populated by more of the same, but with cybernetics.

"tool makers." Humanism thrived as the species spread to the West and celebrated great feats of technological inventions, as evidenced at gatherings like the World's Fair.

With the postmodern era, quite the backlash arose against the imperialist attitudes of traditional Western humanist thinking. Of course, technological developments continued and thinking about improving the human condition found a voice again, but this time with the "post" – inferring the *beyond* or the *after* humanism. Arthur Kroker believes that "technology functions to deliver us to a future that is distinctly posthuman in its radical undermining of all the previous markers of the 'human' – unitary species-logic, private subjectivity, hierarchical knowledge – with human beings as the universal value-standard of all events" (5). This temporal concept of being *after* human literally includes technological developments that are not far off, including but not limited to, sAI (strong artificial intelligence), human-like robotics, cloning and gene therapy, space travel and many more possibilities. As both part of and instigators of these changes, the human will be caught up in the changes as well – some believe this will be for the betterment of humanity and earth as a whole.

Max More, in the *Transhumanist Reader*, argues that "becoming posthuman means exceeding the limitations that define the less desirable aspects of the 'human condition.'" For More and others, like James Hughes and members of the Extropian Institute or Humanity+,[2] these "less desirable aspects" include: disease, aging and inevitable death (More and Vita-More 4, "Mission"). Through developments in computer science and engineering, cognitive science,

[2] With the many different organizations and think tanks dedicated to the movement toward a "future society," there are also multiplying definitions of transhumanism. For example, *Humanity Plus* (+) is an organization which is, according to their website, "The world's leading nonprofit dedicated to the ethical use of technology to extend human capabilities." *Humanity* + defines transhumanism as "The intellectual and cultural movement that affirms the possibility and desirability of fundamentally improving the human condition through applied reason, especially by developing and making widely available technologies to eliminate aging and to greatly enhance human intellectual, physical, and psychological capacities." In other words, they believe that humans have the ability to better themselves by extending lives, improving memory, strength and other biological traits through, what they claim to be, ethical use of technoscience. ("Transhumanist FAQ")

AI and others, More believes that "posthuman beings would no longer suffer from disease, aging, and inevitable death… They would have a vastly greater physical capability and freedom of form… [They] would also have much greater cognitive capabilities and more refined emotions" (4).

A person who subscribes to the "transhumanist philosophy" is a person who believes in improving the human condition – including the body, self and society – through technological enhancement. There are many views surrounding how to most ideally reach such goals. From biological and genomic enhancement to cybernetic appendages and exploration into Artificial Life (biologically based, computationally-based, or some other base that we have not yet conceived of), the Transhumanist believes that what defines "human" is always in a state of flux. In general, the transhumanist philosophy endorses the belief that what makes us "human" is in a developmental state toward what could be called "posthuman." In this transhumanist philosophy, to *become* posthuman is to be effectively enhanced through technological means so that they have surpassed the "limitations that define the less desirable aspects of the 'human condition'" (More 4).

Critical posthumanists want to resist the appeal of the utopian promises of transhumanist thinking. The "qualitative shift in our thinking about what exactly is the basic unit of common reference for our species," Rosi Braidotti suggests, raises serious questions as to the very structures of our shared identity – as humans" (2). And when asking those questions, Diana Relke wants to remind us that "is it any wonder [that] Extropianism, with its relentlessly optimistic focus on the future, is increasingly popular among techno-savvy young men?" (81). The Extropian movement, and other transhumanists, Relke points out, are typically white, privileged men… Suggesting that the future will only be populated by more of the same, but

with cybernetics. And if this is the case, Braidotti is correct in reminding us that these "Discourses and representations of the non-human, the inhumane and the posthuman proliferate and overlap in our globalized, technologically mediated societies" (2).

Clearly, the question of who determines what the "less desirable aspects of the human condition" is perfectly reasonable question but is often side-stepped by transhumanists who simply say technology will solve the problem. In a future with the great bounties of technology realized, there would be no need to worry about poverty or hunger or the supposed "digital divide" because all would be made equal. The *Star Trek* series is often criticized on a point like this. Critics argue that the humanist vision of the shows' creator is an impossible "wet dream" of the future (Relke). Yet others argue that there is enough "techno-skepticism" in *Star Trek* to argue for its continuing relevance in a critical posthumanist discussion (Relke). As Jason Eberl and Kevin Decker explain, "Rather than mere escapism, all the incarnations of *Star Trek* ought to be seen as an entertaining, edifying preparation for thinking through the problems that the future will undoubtedly throw at us" (xvi).

Part of this posthuman future will apparently include living side by side with clones and humanoid robotics, potentially housed with AI. I follow those who assert that such a future requires that we radically rethink laws and social structures so that we may flourish together with these entities. As part of our tool making history, futurist Jim Dator is "increasingly convinced that we humans are inevitably in the process of creating entities that mimic, extend, and in many ways exceed our own mental, behavioral, and emotional capabilities" (51). And in that future, "humanity is about to be surrounded by all kinds of novel intelligent beings that will demand, and may or may not receive, our respect and admiration" (52). For Dator and others like Bostrom it will be crucial to move ahead with developments in a thoughtful way that takes agent-

status of others not for granted in a species-sense, but rather in a way that can allow for entities that assume and or excel our own cognitive and emotional capacities.

The belief that artificial entities, in particular AI, will outpace our own abilities is often framed in the discussion of what has come to be called "The Singularity." Despite popular belief, Kurzweil was not the inventor of the phrase "The Singularity." Rather, it grew from a number of conferences surrounding philosophers and programmers considering the exponential growth of computer programming speed. While Kurzweil is one of the most well-known for promoting knowledge about the singularity, other influential figures include SF writer Vernor Vinge, John von Neumann, Eliezer Yudkowsky, and Ray Solomonoff. All worked to increase knowledge surrounding the idea of "accelerating change" or the "intelligence explosion" (Chalmers).

The concept of the Singularity was first proposed by mathematician I.J. Good in 1965, who envisioned a world in which "smart machines would design even more intelligent machines" ("Scientists Worry" Markoff). This notion has gained growing attention as designers at Silicon Valley and beyond are working harder to unveil smart cars, smart phones, and disembodied AI. Dubbed "The Singularity" by computer scientist and science fiction writer Vernor Vinge this "intelligence boom" is most commonly associated with the work of Kurzweil due to his popular science celebrity status. Kurzweil, expanding on "Moore's Law" (a description of exponential growth in computer processing power),[3] famously predicted in 2005 that the "arrival" of posthuman evolution would occur in 2045 ("Coming Superbrain" Markoff). Since then there has been growing interest in the concept of the Singularity. Dr. Horvitz explains, "Technologists are providing almost religious visions, and their ideas are resonating in

[3] Dr. Gordon Moore, co-founder of Intel and conceiver of "Moore's Law," should not be confused with computer scientist and futurist Max More, co-editor and contributor to the *Transhumanist Reader*.

some ways with the same idea of the Rapture" ("Scientists Worry" Markoff). Others, like Vinge maintain a sort of agnostic view of the supposedly coming singularity, believing that we cannot even begin to imagine such a future (Roden).

Despite the disparate views within the futurist camps, there are some things most agree upon. For example, many futurists who believe in the coming Singularity, would agree with More, and while often conflated with posthumanism, would adopt the label of transhumanism in which the posthuman is a future iteration of the human. For More, this posthuman is something transhumanists are always striving for – it is a philosophy of improvement without end. This brand of posthumanism, also referred to as *trans*humanism or in its most extreme, Extropianism,[4] adopts utopian beliefs about the future. Transhumanism, or at least More's version of it, according to Mervyn Bendle, "leaves little or no room for doubt": "Disbelief is suspended and centuries of hard-won experience and intense critical thinking about science, technology and the social formation within which they flourish are swept aside by an uncritical 'will-to-believe' propositions about the possibilities of science and technology that are often preposterous, and even undesirable" (50). For Bendle, the transhumanist future, full of posthuman entities, is something to be dubious of and he, among others, wonder what that will mean for the *human*.

Similarly, Eugene Thacker fears that the extropian vision of the future will be a significant step backward for the liberal humanist subject: "Like the Enlightenment's view of science and technology, extropians also take technological development as inevitable progress for the human. The technologies of robotics, nanotech, cryonics, and neural nets all offer modes of enhancing, augmenting, and improving the human condition" (74). The humanist vision places certain aspects of the human as special or part of an "essential humanness": "… like the

[4] Extropianism, according to Bendle "sees its (rather daunting) mission as combating the entropic (i.e. disorderly) tendencies of the universe, especially where these impact on human well-being and potential."

types of humanisms associated with the Enlightenment, the humanism of extropianism places at its center certain unique qualities of the human – self-awareness, consciousness and reflection, self direction and development, the capacity for scientific and technological progress, and the valuation of rational thought" (74). Even instances of extreme posthuman visions in fiction or advertising, some argue, are still embedded with the humanist ideology – like never being able to separate the "humanism" from the "post" (Badmington; N. Campbell; Hird and Roberts; Pordzik). Roberts describes this as finding "a humanist text in a posthuman guise whose work is to affirm the immutable, essential nature of the human" (n.p.).

This return to a humanist celebration of the "human" may seem promising, but to some, this return also raises questions for a future with nonhuman entities. Braidotti, for example, argues that "the posthuman condition introduces a qualitative shift in our thinking about what exactly is the basic unit of common reference for our species, our polity and our relationship to the other inhabitants of this planet" (1-2). Not only will our species come in to questions, but for Braidotti, "This issue raises serious questions as to the very structures of our shared identity – as humans – amidst the complexity of contemporary science, politics and international relations. Discourses and representations of the non-human, the inhumane and the posthuman proliferate and overlap in our globalized, technologically mediated societies" (2).

While some feel this return to exalting the humanist vision to be hopeful, the technological developments enacted to achieve this utopian future seem to be simultaneously dismantling, sometimes literally, the human body. Hayles' book *How We Became Posthuman*, for example, opens with her recounting a nightmare-like epiphany while reading Hans Moravec's thought provoking quasi-fictional philosophical text *Mind Children* (1990). In Moravec's work, he illustrates the possibility of "uploading" the human consciousness into a

computer. Hayles describes this process as achieved by a "robot surgeon [who] purees the human brain in a kind of cranial liposuction, reading the information in each molecular layer as it is stripped away" (1). Not just the human body, but our sense of self could be stripped away as well. In fact, for some the posthuman future of the transhumanists will cede power to the technology we create. For Bendle, Kurzweil's particular vision is "an ideological misrecognition of humanity's relationship to technology. In a manner that fundamentally inverts this relationship, posthumanism cedes to technology a determinism over human affairs that it does not, cannot, and should not enjoy" (61)

For some, these relationships with technology could lead to an ideological loss of the sense of self. Michelle Chilcoat explains that "the projected obsolescence of the body also implied the loss of biological matter, traditionally viewed as the immovable or fixed material upon which to construct gender differences and inscribe male privilege" (156). For Chilcoat, this boundary breach goes right to the heart of humanism, including a threat to male privilege. For others, this threatened boundary is explored in terms of "suffering" that can be inflicted upon the human body as technology is not just embraced by the human but rather ruptures the human (Miccoli).

At first Hayles' vision of "information losing its body" seems terrifying, as the human is sucked into a blender and disembodied, she returns to an argument that expands the possibilities of what it means to be human. "When Moravec imagines 'you' choosing to upload yourself into a computer, thereby obtaining through technological mastery the ultimate privilege of immortality," Hayles writes, "he is not abandoning the autonomous liberal subject but is expanding its prerogatives into the realm of the posthuman. (287). For Hayles, "the posthuman offers resources for rethinking the articulation of humans with intelligent machines" (287).

Similarly, for Neil Badmington, even this most posthumanist vision of uploading consciousness is not as much a threat to humanism as at first seems. For Badmington, this imagery comes "from the distinctly humanist matrix of Cartesian dualism. Humanism survives the apparent apocalypse and, more worryingly, fools many into thinking that it has perished. Rumors of its death are greatly exaggerated" (11).

Indeed, it seems Hayles and others, when considering a critical posthuman thought, agree that humanism will remain even in a future that abandons species-specific definitions of the human. "Posthumanism," for Wolfe, "isn't posthuman at all – in the sense of being 'after' our embodiment has been transcended – but it is only posthuman*ist,* in the sense that it opposes the fantasies of disembodiment and autonomy, inherited from humanism itself" (xv). But at the same time, Braidotti reminds us that "Not all of us can say, with any degree of certainty, that we have always been human, or that we are only that. Some of us are not even considered fully human now, let alone at previous moments of Western social, political and scientific history" (1).

Not simply a utopian vision or philosophy for the future (as the extropians and transhumanists would have it), thinking of the posthuman is also used as a tool for critical literary analysis. For Hayles, this means "serious consideration needs to be given to how certain characteristics associated with the liberal subject, especially agency and choice, can be articulated within a posthuman context" (5). It is a discussion of the nonhuman versus the human as a way to better understand where the Self is located. In a way, a posthuman reading is a way to "uncover those uncanny moments at which things start to drift, of reading humanism in a certain way, against itself and the grain" (Badmington 19). For many, this means that "the "post-" of posthumanism does not (and, moreover, cannot) mark or make an absolute break from the legacy of humanism" (Badmington 21). And while humanism will continue to be alive and