The University of Chicago

Can Robots go to Heaven?

Artificial Intelligence & Moral Standing in Christian Ethics

By Taylor Matalon

August 2021

A Paper Submitted in partial fulfillment of the requirements for the Master of Arts Degree in the Master of Arts Program in the Social Sciences

Faculty Advisor: William Schweiker

Preceptor: Victoria Gross

Introduction

In 2007, Russel Moore, a pastor and professor of theological ethics, published a blog post titled, "Robotic Church Membership? Christian Ethics in the Year 2088." In it, he asks whether or not we should baptize robots, human clones, or other forms of artificial life. He poses the following hypothetical scenario:

> *"It is the distant future. You are 106 years-old, and in good health with a sound mind. Your great-grandson, Joshua, is a young pastor in the Southern Baptist Convention (now called the Galactic Immersionist Federation). He is seeking your counsel because, as he puts it, "I've got a question and there's nothing about this in the Bible."*
>
> *Modern technology has enabled infertile couples to engineer what the press of the day calls "robo-frankenbabies." These babies' bodies are constructed partially with, as in the Frankenstein novel of old, body parts from human corpses and partially with body parts produced via human cloning. These children are real flesh and blood in every way, except with a robotic brain. This cyber-brain is programmed with advanced artificial intelligence so that the child is able to truly think on his own. He is able to express joy and sorrow, grief and gladness, the full range of human emotions.*
>
> *At Vacation Bible School (now called "Reverb"), Aidan, age eleven, came to see your great-grandson, the pastor. Aidan's parents are unbelievers, but he has been moved by the Gospel presentation given at the end of Reverb week. He cries in Pastor Joshua's office. In fact, he is convulsing in tears.*
>
> *"I know I'm a sinner," he said to Pastor Joshua. "And I know that I deserve to go to hell." He continued through his sobbing, "I love Jesus, and I want to know Him. What must I do to be saved?"Pastor Joshua stepped out of his office to confer with his Associate Pastor, Caleb.Josh's first instinct is to lead the boy to Christ.*

“Josh, I don’t know,” Caleb said. “Yes, he has a human body, but his mind and heart are artificial. His brain is that of a robot.” Joshua replied, “Yes, but he’s human, too-at least it sure seems so. His blood pumps, his heart beats, he sweats-and he thinks, makes choices, and feels. He even feels guilty!”

Caleb said, “Yes, but he’s programmed to feel and to make choices-even though those choices are random.”

Joshua responded, “But isn’t it the randomness of those choices, and the ability to long for communion with God, the ability to know the conviction of sin-doesn’t that mean something?” Joshua said, “I’ve been told my whole life to offer the Gospel to every repentant sinner. In my office there is a repentant sinner, and I don’t know what to do.” Pastor Joshua walked back into his study to see the trembling boy in his study. Aidan was reading Pastor Joshua’s Bible, an old King James Version that belonged to Josh’s great-great grandmother. Aidan looked up from the Bible and said, “Does Jesus love me, Pastor? Did Jesus die for me? Can I be saved from this guilt and, like you said in your message, from sin and death and hell?”

Aidan looked up and asked, “Am I a real boy? And can I be a real Christian?” Pastor Joshua said to Aidan, “I want you to stay here with Pastor Caleb for a few minutes while I teleport over to the nursing care facility so that I can talk with some one I think might know the answer to this.”

Joshua sits by your bed, asking for help. “Should I lead him to Christ?” he asks. “Should I baptize him?” What do you tell him?While waiting for you to respond, Pastor Joshua organizes the stuff he brought with him-hastily grabbed as he walked out of his office. In his hand is the King James Bible that little Aidan was holding. The Bible falls open to John chapter three. Joshua notices that the red letters in the middle of the page are smudged by something. In verse sixteen, the word “whosoever” is almost illegible. The page buckles underneath where it is printed.

> *"Something's on my old Bible," Joshua says to you, as you think of how to answer his question. "Some kind of salty water stain, I don't…." Joshua stops and looks up at you, with realization: The stain is from the tears of a crying seeker. And they sure look human to him."*[1]

The rapid development of artificial intelligence seems to indicate that the horizon of true 'thinking machines,' may be just around the proverbial corner. Forgiving the perhaps over-dramatic prose, Moore's anecdote elucidates several key questions regarding the future of Christian ethics and the role that such thinking machines ought to play therein. The question, "Do robots go to heaven?" is a relatively elementary one, yet I find it to be worth philosophical inquiry beyond mere public discourse. This thesis seeks to examine, on a more theoretical level, the various arguments and philosophical principles that inform scholars' differing perspectives apropos of artificial intelligence. In so doing, questions about human ontology, human telos, the existence of a soul, the origin of intelligence, the properties of intelligence, and human-technology relationships will most certainly arise.

First, I will explore the concept of human intelligence–and its potential replicability (or lack thereof)–from the perspective of leading Christian theologians and ethicists. I wish to, taking a more critical eye, offer a third-party perspective on which theological principles and arguments are most consistent, or most in conflict with a potential superintelligence explosion. Given the historical precedent for conflict between theological and scientific understandings of human ontology, will we see a similar reckoning with certain beliefs in Christian dogma? Or, are those theological frameworks that conflict with science relics of the past? Some theological perspectives are undoubtedly more resilient than others, and this thesis will attempt to ascertain which those might be. Second, I will look to secular–namely Neo-Darwinian (based on

[1] Moore, Russel. "Robotic Church Membership? Christian Ethics in the Year 2088." Russell Moore, October 5, 2015.

evolutionary biology) and physicalist (based on physics)–perspectives on the same topic to see what conflict and/or synthesis can be seen with those of the various Christian scholars. Thereafter, I will examine how these varying viewpoints on intelligence inform one's determination of the moral standing of artificial intelligence (or lack thereof) as it relates to that of human beings. That is to ask, is replicating human intelligence a necessary *and* sufficient condition for granting equal moral standing? In performing these analyses, I arrive at the following arguments, which will be explicated further later on:

> *(i)* Theological Anthropocentrism based on discrete unique human *properties* is weaker than an anthropocentrism positing a more holistic, transcendent, or supernatural justification of human superiority. In other words, if human superiority is predicated on the unique property X, the replication of property X by machines would falsify such anthropocentrism.
>
> *(ii)* A Neo-darwinian or Physicalist conception of human intelligence would suggest the possibility (not inevitability, necessarily) of superintelligent AI.
>
> *(iii)* Theological Anthropocentrists tend to reject the granting of equal moral status to artificial beings
>
> *(iv)* Neo-darwinian or physicalist criteria for moral standing (currently assessed in humans and animals) could theoretically be achieved by superintelligent AI.

The in-depth prescription of my own *personal* criteria for both intelligence and moral standing is beyond the scope of this thesis, and is something I wish to explore deeper in due time. I should say, for clarity's sake, that I do personally advocate for a radical egalitarian consequentialism that rids itself of species boundaries; it is un-anthropocentric in nature, opening the door for the moral standing of animals and theoretical artificial life. Rather than provide a formal prescription for

such a framework, this thesis instead should provide a third-party critical analysis of the interaction between different theological and secular perspectives on the topic.

A Brief Discussion of Caveats

I should first acknowledge the most noticeable elephant in the room, which is the near incomprehensible diversity of opinion that can fall under the umbrella of "Christian Theological Ethics." Per the "complexity thesis," it is impossible to ascertain a definitive "Christian" position on any given ethical or epistemological issue. As such, I will be examining those positions, and the justifications behind them, that have been most prevalent in recent *academic* discourse. I say this because there exists plenty of Christians, particularly in the U.S., who wield their faith as justification for conservative political positions. These individuals, however, tend to be quite far removed from contemporary academic theology. We will examine later, through some primary sources, the differences and similarities between the two groups.

Second, I want to make my use of key terms clear. Pedantic debate over any philosophical term could take up dozens of pages, so I should indicate that my use of "intelligence," refers, generally, to the cumulative cognitive capabilities of any agent. Though there is slight variation in definitions thereof, this paper is ambivalent thereto. I am not definitively asking what intelligence is, but rather examining how different conceptions of *human* intelligence compare to different conceptions of *artificial* intelligence. As such, this comparison is beholden to the definition of intelligence put forth by the specific scholar or framework being discussed. I should also be clear about my use of terms like "superintelligence," "strong AI," and the "intelligence explosion." Put most simply, strong or superintelligent AI is that which is beyond specific and discrete uses. A chess-playing AI can

only play chess, and is considered categorically "weak." The threshold for "superintelligence" is a strong AI that possesses human-or-greater cognitive capabilities across the board. The "intelligence explosion" is a chronological term, indicating the (hypothetical) point in time at which superintelligence is achieved.

Finally, I should state any potential biases and/or conflicts of interests. I am not Christian; I am a Jewish Atheist. Though I went to Catholic high school, this thesis is not serving as a mission of philosophical vengeance against my former school teachers. My interest in Christian Ethics is mostly academic in nature.

Intelligence from a Christian POV

According to Ted Peters, Christian scholars are faced with two potential paths when reckoning with the potential intelligence explosion. As he explains, Christian Scholars can "(1) abandon the view that the defining feature of humanity is rationality and propose an account of spirituality that dissociates it from reason; or, on the other horn, (2) find a way to invalidate the growing faith in a posthuman future shaped by the enhancements of Intelligence Amplification (IA) or the progress of Artificial Intelligence (AI)."[2] We will explore both of these options below, starting with the second.

Intelligence as the Soul

The scholars whose work I explore in this section see intelligence or the human mind as the definitive distinguishing quality of humanity. The mind, and the incredible functioning thereof, is itself necessary and sufficient evidence for the existence of a soul, usually explained

[2] Peters, Ted. "Artificial Intelligence versus Agape Love: Spirituality in a Posthuman Age." In *Forum Philosophicum*, vol. 24, no. 2, pp. 259-278. 2019.

as uniquely bestowed upon humanity by God. Thus, such scholars reject that an AI could ever be 'truly' intelligent, and subsequently reject any claim of equal moral standing.

Pope John Paul II, for example, in an official statement to the Pontifical Academy of Sciences in 1996, argues that "theories about evolution that consider the mind as emerging from the forces of living matter are incompatible with the truth about man." Instead, the mind emerges as a product of Creation in Imago Dei. This perspective is fundamentally anti-naturalist, as it rejects the possibility of a wholly rational and/or scientific explanation of human intelligence. Fearful, however, of appearing like a denier of evolutionary theory, his excellency clarifies: "If the human body takes its origin from pre-existent living matter, the spiritual soul is immediately created by God." Here, his excellency invokes classical mind/body dualism. In so doing, he has created a theoretical mechanism for both (i) accepting a naturalistic explanation of humanity's biological origin *and* (ii) holding that the mind is beyond natural explanation, as the mind is itself a-biological.

This dualism is a common thread of this style of argument, and its foundations can be traced back to Augustian thought. As Bruce Reichenbach summarizes:

> "The traditional (although currently not unanimous) Christian view of human beings is that we are dualistically composed of body and soul (Augustine 1948a, XIX, 3). Creating in his own image, God formed in the human dustly body "a soul endowed with reason and intelligence, so that he might excel all creatures of earth, air, and sea" (Augustine 1948a, XII, 23). Although the soul permeates the body, they are separable at death (Augustine 1948b, XVI, 25). Even when separate from the body, the soul can continue to

> exist until divinely reunited with a resurrected body in a perfected state, fleshly yet spiritual (Augustine 1948a, XXII, 19, 21)."[3]

Note that our *intelligence* is that which allows us to excel all other creatures. This soul–or that which gives us intelligent thought, breathed into us by God during creation–could not be artificially replicated, as it is an inseparable aspect of the principle of Imago Dei. As Reichenbach explains,

> "Many Christians have understood the image of God ontologically, expressed in human features such as intelligence, ability to reason, and the moral decision-making faculty of the soul that emulate, in some analogical way, those of the creator. Augustine held that the imago Dei refers to the *rational* soul, and specifically to its unique Trinitarian abilities of memory, understanding, and will."[4]

Here, we see rationality, and specific functions of the rational mind, as inextricably linked to the creation of humanity itself. To be made in God's image is to possess rational faculties. Any attempt at recreating or simulating human rationality, thus, is doomed to fail, according to this perspective. Thinking machines are infinitely limited by their mechanical composition. On the other hand, human intelligence, according to these scholars, could never be explained by mechanical means, or merely by its biological composition. As Swineburne argues, "science cannot provide an explanation of why particular brain events cause specific mental events, and vice versa, for it is unlikely that it can discover natural laws among or governing the phenomena."[5] To him, the phenomena of the human mind is supernatural in nature; naturalistic explanations are definitionally insufficient. There is a transcendent quality of the mind that goes

---

[3] Augustine , 1948a. " City of God." In *Basic Writings of Saint Augustine* II, edited by Whitney Oates. Grand Rapids, MI: Baker. Qtd. in Reichenbach, *Three Phases*

[4] Reichenbach, Bruce R."Christianity, Science, and Three Phases of Being Human." *Zygon*® 56, no. 1 (2021): 96-117.

[5] Swinburne, Richard. "The evolution of the soul." (1986).

beyond the material bounds of the brain's neurochemistry. As Ian Barbour explains, "The biblical view does indeed conflict with the determinist and materialist philosophical assumptions of many neuroscientists."[6] To these scholars, physicalist explanations of the mind fail to comprehensively explain its complexity. This anti-physicalism, or anti-reductionism, thus, necessarily precludes the possibility of a purely physicalist explanation of *artificial* intelligence; No matter how complex the infrastructure or programming, AI will always lack a soul, and thus the prerequisite conditions for true rationality.

Modern scholars have attempted to argue this point beyond mere theory, and have turned to contemporary artificial intelligence, noting how it fails to meet the unreachable standard of human intelligence. It is certainly clear, however, that the analytical and computational faculties of the human mind can indeed be replicated (i.e. mathematics, strategy, pattern recognition). Thus, these scholars tend to focus on more subjective, social, and emotional aspects of human intelligence in asserting its uniqueness. Winograd & Flores, for example, argue that language and relationships are necessary conditions for true intelligence. The emphasis on language has been at the forefront of AI development since Alan Turing's landmark paper "Computing Machinery and Intelligence." Where Turing, a secular scholar, predicts that machines could indeed replicate human linguistic capability,[7] Winograd & Flores assert that such linguistic mastery by an AI is "a pipe dream."[8] In their view, the social implications and subtext of language is crucial for the full understanding thereof. Where machines simply match symbols to meaning, "to be human is to be the kind of being that generates commitments, through speaking and listening. Without our ability to create and accept (or decline) commitments we are acting in a less than fully human

[6] Barbour, Ian G. "Neuroscience, artificial intelligence, and human nature: Theological and philosophical reflections." *Zygon*® 34, no. 3 (1999): 361-398.
[7] Turing, Alan M. "Computing machinery and intelligence." In *Parsing the turing test*, pp. 23-65. Springer, Dordrecht, 2009.
[8] Winograd, Terry, Fernando Flores, and Fernando F. Flores. *Understanding computers and cognition: A new foundation for design*. Intellect Books, 1986.

way, and we are not fully using language."[9] Herzfeld argues a similar point, explaining that true linguistic "understanding is both predicated on and produces social ties," neither of which artificial intelligence is capable.[10] Peters, additionally, argues that "the general human intelligence we have come to know is co-present with the self, the person-in-relationship, the soul. To date, machine intelligence lacks selfhood. To date, so-called machine intelligence does not emulate human intelligence in this regard. AI is selfless, soulless. The only human intelligence we know is found here: the embodied person-in-relationship."[11]

From language and social ties comes an emphasis on emotional capability. According to Herzfeld, "To navigate the world of relationships, one needs to express and perceive emotions, to manage one's own emotions, and to use emotions to facilitate thought. While we often think that our emotions obscure rational thought, they are actually necessary for cognition."[12] She continues, asserting that "Computers can be programmed to recognize emotions in facial expressions or tone of voice, and a robot or a video avatar can be programmed to express a variety of emotions, physically or verbally. However, feeling emotion requires a level of self-consciousness current machines lack."[13]

Self-consciousness and subjective experience are key here. As Reichenbach explains, the neo-darwinians' "reductionistic, rationalistic understanding of human beings omits much of what it is to be human: our interaction with the environment, our ability to converse, and our ability to navigate the world of relationships."[14] To these scholars, an artificial intelligence has no such subjective experience, and thus cannot accomplish the social or emotional feats that the human

[9] (Winograd and Flores 1991, 68).
[10] Herzfeld, Noreen L. "In Whose Image? Artificial Intelligence and the Imago Dei." in *Blackwell Companion to Science and Christianity* (2014).
[11] Peters
[12] Herzfeld
[13] Herzfeld
[14] Reichenbach

mind can. A computer can perform its directed functions, but it is unable to achieve self-awareness of such functioning and lacks the agency to order and choose which specific functions it ought to perform.

Tied to subjective experience and awareness is a sense of self-embodiment; that is, the capacity for a mind to move throughout the world and physically interface with objects and other agents. Herzfeld asserts that any intelligent agent must be embodied in some way. As such, "much of what we consider to be human intelligence is not an internal quality of the mind. Our designs and behaviors arise through and out of interaction with the environment."[15] That is to say, a priori thought is but one aspect of intelligence; another–empirical and sensual experience–is equally necessary, yet harder to reproduce with machines, if not impossible. It should be emphasized here that many of these critiques levied against the potentiality of computational intelligence are wielded as justifications for a non-physicalist account of human intelligence; that is to say, given that humans can indeed achieve advanced linguistic skills, emotional intelligence, and self-awareness and agency, yet computational systems cannot, there must be some non-mechanical, non-physical quality X that allows us to perform such functions. This argument is not inherently Christian, as there are a number of scholars from both secular schools and schools of other faiths that wield it.

The primary issue with this line of argumentation-by-example, however, is that the supposed superiority of the human mind over that of machines is *conditional*. In wielding specific instances in which artificial intelligence fails to meet human standards as evidence for the superiority of the human mind, one's conclusions depend wholly on the validity of such empirical premises. With language, for example, how would Winograd & Flores respond if indeed an artificial intelligence passed a Turing test with flying colors? What would Herzfeld say

[15] Herzfeld

if an artificial intelligence were able to fully control an external body and gather sensory data? Given the logic of this argumentative structure, their theological anthropocentrism would need serious revision if/when machines met these specific criteria for intelligence. As such, other scholars contend that intelligence is *not* the end-all be-all distinguishing factor of humanity. Our unique souls are more than our rational minds, and our unique relationship with God should be considered independent of our cognitive abilities.

Beyond Intelligence

This view of human ontology shifts the conversation away from one of essential human *properties*. In other words, the coming arguments hold that our unique position among creation is not due to our composition. These arguments, however, still tend to assert an inherent superiority or uniqueness of humanity, precluding the recognition of artificial intelligence as beings of equal value.

Ted Peters, for example, posits that our capacity for agape love is what sets us apart. In his words, "regardless of the level of intelligence residing in each of our brains and minds, it is love and not reason that becomes the measure of our godliness."[16] He suggests that "we should surrender the long-held belief that rationality is unique to the human person." In decentering human rationality, Peters opens the door for the possibility of strong artificial intelligence. In doing so, however, he forgoes intelligence as the primary indicator for the soul, or as sufficient evidence of imago Dei. Russell Bjork expands upon this view of creation. To Bjork:

> "If *intelligence* is made in Imago Dei, then any being capable of demonstrating intelligence is made in imago dei (including AI). However, if the divine image is to be understood in some other way, then there is no logical conflict between strong artificial

[16] Peters

> intelligence and the doctrine of the image of God. In particular, one can certainly hold that 'being created in the divine image requires (and hence implies) being rational' without holding the converse: 'being rational implies being created in the divine image.'"[17]

Bjork emphasizes the unique *relationship* that exists between God and humanity, rather than any set of divine qualities. As he explains, "when human worth is tied to human constitutional uniqueness, the possibility of strong AI seems to pose a serious threat to one of our most cherished concepts."[18] Moral worth should not be tied to the presence or absence of a rational soul; instead, human worth comes from our role as the shepherds of God's will, independent of our constitution.

Along the same vein, Hehn argues that imago Dei ought to be interpreted as a "job description" more than a description of compositional properties, shifting the discussion from one of ontology to one of teleology.[19] Von Rad, running with this teleological shift, posits that, "man is placed upon earth in God's image, as God's sovereign emblem. He is really only God's representative, summoned to maintain and enforce God's claim to dominion over the earth."[20] The theological position of man as merely a shepherd of God's will (hereafter shortened to 'man-as-shepherd) is, prima facie, one that is less anthropocentric than that which claims a superior human ontology. Still, in asserting oneself as a shepherd of God's will, one imbues a sense of dominion and moral authority that is reserved just for humans. Though this view tends to inform modern "ecotheology" and a deeper foray into animal welfare, such scholars tend to maintain that the fate of other beings remains in the hands of man-as-shepherd.

---

[17] Bjork, Russell C. "Artificial Intelligence and the Soul." *Perspectives on Science & Christian Faith* 60, no. 2 (2008).
[18] Bjork
[19] Hehn, Johannes. 1915. Zum Terminus "Bild Gottes." In Festschrift Eduard Sachau zum Siebzigsten Geburtstag. Berlin: G. Reimer, pp. 36–52, Qtd in Herzfeld
[20] Rad, Gerhard von. "Genesis: A Commentary, translated by John H. Marks." (1961).

Other scholars suggest that attempts at ascertaining a definitive set of unique human properties–intelligence or otherwise–are inherently futile, as the true nature of human ontology is beyond rational understanding. Derek Schurmaan, for example argues that our "ontological starting point" ought to be within the *spiritual* realm; scholars in the 'properties' camp fall into a reductionist trap of setting discrete limits on human ontology, likening us to "complex biochemical machines."[21] Similarly, Niebuhr argues that "man is more than a rational creature," and that our true nature should be understood as transcendent, rather than neurological.[22] Here, we see the same anti-naturalist tones as scholars like Swineburne and JPII; yet, rather than describing intelligence itself as the inexplicable, supernatural phenomenon, Schurmaan and Niebuhr view the "spiritual" or "transcendent" nature of humanity to be holistic. To ascribe certain human quality X as transcendent, but quality Y as naturally explicable would be an arbitrary and reductionist practice. As such, this holistic, transcendent view of human ontology conflicts with the stricter mind/body dualism endorsed by John Paul II and Augustine.

These holistic scholars recognize the epistemological danger that comes with anchoring humanity's divine image on a specific set of rational properties, such as language or self-awareness. By distancing imago Dei from mere rationality, they can maintain an axiological anthropocentrism while acknowledging the possibility of strong AI. In essence, no technological feat or scientific discovery could shake the foundation of their perspective; this anthropocentrism is *unconditional*, compared to the anthropocentrism of Herzfeld, Winograd & Flores, etc. In this sense, both the man-as-shepherd and holistic transcendent justifications of human superiority are better equipped to withstand an intelligence explosion than one based on unique human

[21] Schuurman, Derek C. "Artificial Intelligence: Discerning a Christian Response." *Perspectives on Science & Christian Faith* 71, no. 2 (2019).
[22] Niebuhr, Reinhold. *The Nature and Destiny of Man: A Christian Interpretation: Human Nature*. Westminster John Knox Press, 1996.

properties. This strong anthropocentrism also sets up a more ethical discussion we will return to later; if intelligence is *not* the primary requisite for normative value, there is no need or incentive to recognize artificial life as worthy of such value.

Secular Intelligence

In the above section, we explored the two major arguments put forward by Christian scholars apropos of artificial intelligence: *(i)* a denial of the possibility of strong AI, due to intelligence being uniquely human *(ii)* a shift away from intelligence as the primary indicator of human superiority. In the following section, I will briefly explore the leading secular arguments on the nature of human intelligence and the potentiality of strong AI, noting any strong conflicts or areas of congruency with the Christian perspectives.

On the origin and nature of human intelligence, leading secular scholars tend to follow a Neo-Darwinian approach, citing evolutionary principles as sufficient explanatory tools in describing the phenomenon of the human mind. This approach is fundamentally naturalistic, inherently supposing that all phenomena–even those beyond current understanding–can theoretically be explained by natural law and physical principles. E.O. Wilson, one of the fathers of Neo-Darwinism, argues that *any* quality of humankind can be explained evolutionarily, or by tracing back our biochemistry. Such as we are creatures of nature, all of our qualities necessarily come from nature as well. He offers the following argument in response to those who propose that certain qualities (e.g. intelligence) are instead transcendent:

1. If E (evolved), then not T (transcendent)
2. E
3. Therefore, not T

To Wilson, no explanation of a species' nature is necessary "beyond the imperatives of its genetic history."[23] Using his diagram, *nothing* cannot be said to be evolved. Here arises an immediate conflict with John Paul II's dualist supposition. It is inconsistent, in Wilson's eyes, to suppose that *some* aspect of homo sapien is evolved, yet others are bestowed upon us by God. If we possess a thorough and internally consistent genetic history of the human brain and all related functions, how can it be said that the intervention of God is what causes intelligence? If we can point to specific evolutionary developments, specific departures from other apes, or specific mutations in neurochemistry, and demonstrate how such events affect human psychological behavior, where in the picture can primary divine intervention fit? I should note, however, that Wilson's view is primarily critiquing those theological views that view creation as a direct act of intervention by some figure "God." It is less in conflict with those who take a more "secondary" interpretation of divine intervention (e.g. God working through natural processes like evolution).

Beyond Wilson, Reichenbach raises similar Darwinian concerns with a strict dualist view. As he explains,

> "Both [dualist] views—that the soul emerged at some point in the ancestral human lineage or that God implants it individually in each human—encounter the problem of the "first." At what point did the first ensouled *human* being arise? On the creationist account, did God implant souls into Neanderthals or Denisovans, or only into homo sapiens? Did the first ensouled human wonder whether his or her parents had self-consciousness, moral consciousness, or free choice?"[24]

Strict definitions of "humankind" run into all sorts of biological and historical trouble, as other members of the hominid tree had similar if not fully equal cognitive capabilities. Indeed, is the

[23] Wilson, Edward O. "What is sociobiology?." *Society* 15, no. 6 (1978): 10-14.
[24] Reichenbach

soul unique to the homo sapien subspecies of hominids? We can extend this issue into the future as well. In 5,000 years, if interplanetary colonization and radical genetic engineering have been achieved, surely the genetic makeup of different hominid groups would be vastly different, even more so than the difference between homo sapien and homo neanderthalis. Which of these would get endowed with a soul? Which of these would maintain the rational faculties imparted upon us by God?

Wilson and Reichenbach are not the only scholars to note the flaws in a strictly dualist approach to the human mind. Much of Neo-Darwinian thought, which works on the plane of evolutionary biology, can be traced back to a more pure physicalism, which seeks to provide explanations for natural phenomena with fundamental physical law. Evolutionary principles can be explained by molecular biology, which can be explained by chemistry, which can be explained by physics, as the framework goes. Richard Dawkins, for example, argues that "human thoughts and emotions emerge from exceedingly complex interconnections of *physical* entities within the brain."[25] Dawkins vehemently critiques (rather harshly) any attempt to work beyond this physicalist framework. Alan Turing stated that he is "unable to accept any part of" theological dualism. He argues that a view of rational thought as conferred by God is not only arbitrary, but also fails as a justification for human superiority, as God could choose to impart a soul into an elephant or into a machine on any day, if he wanted.[26] Stephen Wolfram's "Principle of Computational Equivalence" suggests that "all processes, whether they are produced by human effort or occur spontaneously in nature, can be viewed as computation."[27] As such, the functions of the human mind–unprecedented as they may be–are just as computational as other natural processes, according to this principle.

---

[25] Dawkins, Richard, and Lalla Ward. *The god delusion*. Boston: Houghton Mifflin Company, 2006.
[26] Turing
[27] Wolfram, Stephen. *A new kind of science*. Vol. 5. Champaign, IL: Wolfram media, 2002

An interesting wrench in the physicalist framework is the phenomenon of emergence, which we briefly discussed earlier. Emergent phenomena are, definitionally, those that function beyond the expectations or calculable limits of their physical composition. Many scholars, secular or religious, view human consciousness as one such instance of emergence. Though we can (mostly) map the human brain and all of its systems, neuroscience is far from a complete project, as there still exists a significant knowledge gap in the bridge between neurological anatomy and psychological behavior. Some, like Schurrman and Niehbor, take the emergent perspective of consciousness as evidence of a divine or supernatural intervention. Emergence does not, however, necessarily preclude a physicalist understanding of natural phenomena. Bjork, for example, argues the following:

> "It does seem theologically plausible, then, to hold that personhood emerges from the (physical) interaction of neurons in the brain. Such a view is consistent both with the holistic tenor of Scripture and with empirical evidence for continuity among living creatures and for mind-brain interdependence. If this is the case, then there would not seem to be—in principle—a theological reason why personhood could not emerge in similar fashion from the operation of a sufficiently complex technological artifact"[28]

Bjork, here, is working within a Christian framework while also arguing in favor of a physicalist account of emergent phenomena, showcasing that physicalism is not wholly incompatible with Christian thought.

If one accepts a purely physicalist account of human intelligence and its origin, it follows that such intelligence could be replicated and expanded upon by technological means. That is to say, if there is *not* some transcendent quality of intelligence, all of the necessary components exist naturally. Thus, given sufficient scientific and technological advancement, there is no

[28] Bjork

reason that such natural components couldn't be fully developed artificially. As such, most physicalists tend not to view AI as an axiomatic impossibility. In fact, most seem confident that superintelligence will be achieved in the relatively near future. Turing, for example, predicted that superintelligence would be achieved by 2020. Obviously, his prediction was a bit hasty, but perhaps not off by an order of magnitude. In a survey of half of the world's leading AI scholars, over 50% suppose that human-or-greater intelligence will be achieved by the turn of the 22nd century. For the vast majority of secular scholars, the question of superintelligence is not an 'if,' but a 'when.'

It also follows that, given a physicalist understanding of human intelligence, an artificial intelligence could have a composition that allows its capabilities to *greatly* exceed those of the traditional human mind. Ray Kurzweil, for example, argues that, "once a computer achieves a human level of intelligence, it will necessarily roar past it. Humans will no longer have intellectual advantages over machines."[29] Rather than pit humans *against* machines, however, many scholars see the integration of artificial intelligence with the human as a more likely scenario. Currently, the capacity of human intelligence is bound to the biological limits of the homo sapien body; integration with artificial intelligence, however, shares no such bounds. Such integration would create a "posthuman." With the posthuman, as Katherine Hayles explains, "there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals."[30] Advocates of posthumanism also argue that such advancements are but the natural progression of human evolution. Hayles, again, notes that "the posthuman view thinks of the

[29] Kurzweil, Ray, and George F. Gilder. *Are we spiritual machines?: Ray Kurzweil vs. the critics of strong AI*. Discovery Inst, 2002.
[30] Hayles, N. Katherine. "How we became posthuman: Virtual bodies in cybernetics, literature, and informatics." (2000): 464.

body as the original prosthesis we all learn to manipulate, so that extending or replacing the body with other prostheses becomes a continuation of a process that began before we were born."[31] To that end, the technological expansion of traditional biological limitations is a continuation of evolutionary adaptation. Nick Bostrom, a leading secular AI Ethicist, sees the shedding of traditional hardware limitations and the transition from human to posthuman as a near inevitability.[32]

Given a physicalist, Neo-Darwinian account of intelligence, certain aforementioned Christian perspectives become immediately problematic; namely, those scholars who hold intelligence to be the key property of the human soul. In asserting intelligence as inherently *(i)* divine and/or *(ii)* unique to humanity, proverbial alarm bells ring louder and louder as AI development hastens. For this reason, those who view human ontology as independent of, or beyond specific thresholds of intelligence have a much more resilient theological perspective in the face of an intelligence explosion. Bjork, for example, does not deny or fear superintelligence, and even accepts a darwinian account of the human mind, positing that "God did not create abstract intelligence—he created physical brains, evidently using an evolutionary process, which incorporates features that closely resemble those in the brains of lower creatures."[33] He prioritizes "not what humanity is, but rather God's relationship to us based on his purpose for making us." This purpose, in Bjork's view, would not change with the advent of superintelligent AI. The general view, then, of man-as-shepherd of God's will remains theoretically consistent with a Neo-darwinian one. Issues still arise, however, in the view of man-as-transcendent. Though Niebuhr and Schurmaan, for example, do not hold rationality itself to be a transcendent phenomena, their holistic view of human divinity still conflicts with Wilson's assertion that *all*

[31] *Ibid.*
[32] Bostrom, N. *Superintelligence: Paths, dangers, strategies.* Oxford University Press, 2014.
[33] Bjork

aspects of humanity can be explained by our genetic history–intelligence or otherwise. This strong anthropocentrism, however, could never be rebutted on rational grounds, as the axiomatic framing of human ontology as a supernatural phenomena makes rational discussion definitionally futile. As such, so long as supernaturalist scholars do not list *specific* properties of human uniqueness, and hold steadfast in their assertion of *holistic* human superiority, their theological perspective remains equally resilient to an intelligence explosion.

AI, Ethics, & Moral Standing

I now wish to shift the conversation to one surrounding the ethical implications of such views of human ontology, intelligence, and AI. Namely, I am curious how the various perspectives we have discussed inform one's determination of the moral standing of hypothetical artificially intelligent beings. In exploring such a curiosity, it is necessary to examine the various criteria for moral worth that have been put forward, whether such criteria be on utilitarian, deontological, or theological grounds.

Robot Church Membership? - Christian POV

In 2018, the Ethics and Religious Liberty Commission released a statement of principles pertaining to artificial intelligence. Over 70 co-authors and co-signers from various evangelical congregations were involved in the production of the document. What should be stressed, however, is the social and political difference between the ERLC and more traditional academic theologians. American Evangelicals are those that, as mentioned earlier, tend to wield their faith both *(i)* less critically and *(ii)* more politically than members of the academy. That said, I find

their views still worthy of exploration, with particular emphasis on how they interact with the academic frameworks we've discussed thus far. The very first principle stated is the following:

> "We deny that any part of creation, including any form of technology, should ever be used to usurp or subvert the dominion and stewardship which has been entrusted solely to humanity by God; nor should technology be assigned a level of human identity, worth, dignity, or moral agency. Genesis 1:26-28; 5:1-2; Isaiah 43:6-7; Jeremiah 1:5; John 13:34; Colossians 1:16; 3:10; Ephesians 4:24"[34]

Here, these scholars wield the man-as-shepherd conception of a unique human teleology to justify the normative dominion and stewardship over other beings, thereby precluding the possibility of artificial life being granted equal moral status in perpetuity. The commission also wields our unique ontology as justification for higher moral status. They "deny that AI will make us more or less human, or that AI will ever obtain a coequal level of worth, dignity, or value to image-bearers." In describing humans as image-bearers, the authors wield creation as sufficient evidence for our superior ontology. Given that AI are *not* created imago Dei (an argument taken as self-evident), they cannot be considered of equal value. Beyond rejecting *equal* value, the commission also rejects that *any* level of love or respect ought to be granted to intelligent machines. In their words, artificial intelligence is "not worthy of man's hope, worship, or love."[35] To them, the love of AI would interfere with both *(i)* human-human connection and *(ii)* human-God connection. Both of these arguments are present among other scholars, as well.

On the subject of love, we can return to Peters' argument that agape love is the true marker of humanity's uniqueness. As such, this love could not be extended to other beings, lest it be unrequited and unfulfilling. He emphasizes the importance of the "person-in-person

[34] Staff, Ethics & Religious Liberty Commission. "Artificial Intelligence: an Evangelical Statement of Principles." Southern Baptist Convention (2018).
[35] *Ibid.*

relationship," granting such personhood only to humans. Herzfeld, similarly, warns that "we must ever be wary that we do not substitute relationships with machines, or even relationships with each other that are totally mediated by machines, for the real thing."[36] The "real thing," in her view, is of course the unique connection between human-human and human-God. For both Peters and Herzfeld, there is some discrete quality X that serves as the singular criterion for the granting of human moral status, and thus the validity of one's relationships. For Peters, quality X is the capacity for agape love. For Herzfeld, quality X is the rational soul. Both of these views lean into strong anthropocentrism, positing that the substandard quality of human-machine relationships is self-evident, given the machines' absence of quality X. Even if, by all prima facie accounts, one's relationship with an AI could make one laugh, cry, feel disgust, or feel affection, Herzfeld and Peters would argue that such relationship was infinitely lesser than that with another human, though each hold different qualities of humanity to be the indicators of superiority. Other scholars lean into this argument, holding that *any* display of emotion or social interaction by an AI is a mere simulation thereof. That is, machines can only "represent emotions without having emotions."[37] As such, these representations ought not be given the same moral weight as the 'genuine' emotions of humanity.

AI's potential interference with the unique God-human relationship is also cause for concern among Christian scholars. Reichenbach warns against a scenario in which AI developers and genetic engineers are worshipped rather than God:

> "On this scenario, God is no longer relevant to the origins of the new being. Humans now are free to create the new humans in their own image. The imago Dei will be replaced

[36] Herzfeld, Noreen. "Artificial Intelligence and the Imago Dei." *The Blackwell Companion to Science and Christianity* (2012): 500.
[37] Wallach, Wendell, and Colin Allen. *Moral machines: Teaching robots right from wrong*. Oxford University Press, 2008.

> with the imago māchinae, or better, imago indiciōrum. One looks to the replicators for guidance as to what the new being will remember, think, feel, desire, and believe, for downloading and replication can now be selective, taking from the individual human brain what it wants and discarding the rest."[38]

To secular transhumanists, this scenario does not necessarily carry a negative normative connotation. To Christian scholars, however, the worship of technology as a false God is a major problem. Schurmaan details and critiques such transhumanist movements in the following passage:

> "Anthony Levandowski, a former Google and Uber engineer who is working to 'develop and promote the realization of a Godhead based on Artificial Intelligence' and that 'through understanding and worship of the Godhead, [to] contribute to the betterment of society.'x The transhumanist Zoltan Istvan suggests that this new AI deity 'will actually exist and hopefully will do things for us.' This is essentially a form of idolatry, replacing a trust in the creator with technology."[39]

Beyond a normative disapproval of such idolatry, would Schurmaan and Reichenbach advocate against the near limitless potential benefits that AI could bring to society? Those at the Ethics & Religious Liberty Commission argue that, "Humanity should not use AI and other technological innovations as a reason to move toward lives of pure leisure even if greater social wealth creates such possibilities." This is an interesting turn. Beyond the rejection of AI as individual persons, the strongest viewpoints reject even the potential instrumental value that AI could bring to humanity. It is a peculiar combination of the classic 'struggle makes us stronger' mentality, mixed with a fear of idolatry and techno-theism.

---

[38] Reichenbach
[39] Schurmaan

Are there prominent Christian scholars that are, simply put, pro-AI? My tentative answer seems to be: sort of. It is difficult to find an academic work that explicitly argues that artificial intelligence, or artificial life in general, is *certainly* worthy of a moral status equal or greater to that of humans. That being said, there are authors who have produced a theological ethic that leaves the door open for such potential moral value.

James Gustafson, in his seminal *Ethics from a Theocentric Perspective,* critiques the general tendency of modern Christian Ethics as being far too anthropocentric. He argues against a normative hierarchy that places humanity at the top. He advocates instead for the treatment of all beings as creations possessing unique relationships with God, rather than mere instruments of humanity's purpose. His argument is an explicit rejection of the man-as-shepherd view of human superiority. Turning to the text, He posits that "if God is 'for man,' he may not be for man as the chief end of creation. The chief end of God may not be the salvation of man."[40] As such, I argue that Gustafson's ethic allows for the potential recognition of artificially intelligent beings as possessing moral worth. If all beings are part of creation, and all beings have relationships with God, who are we to deny the God-AI connection? Similarly, Gustafson rejects the notion of a knowable human telos, instead offering a dynamic conception of human flourishing in the following passage:

> "When we seek to discern and articulate the necessary conditions, for example, for the flourishing of a human society, our observations are less precise, more generalized (and perhaps even quite formal). These conditions change over the course of human development. Some ordering of sexual relationships seems to be required; precisely what that will be is subject to many biological, historical, cultural, and social factors."[41]

[40] Gustafson, James M. *Ethics from a theocentric perspective: theology and ethics*. Vol. 1. University of Chicago Press, 1983
[41] Gustafson, 240

Gustafson recognizes that the nature and conditions–both material and philosophical–of human flourishing are historically and socially contingent. To that end, the humans of the middle ages, renaissance, or industrial revolution could not possibly theorize what 21st-century flourishing could or *should* look like, beyond a rudimentary comprehension of basic biological need. I would argue, thus, that Gustafson's dynamic view of human flourishing leaves the door open for the full integration of artificial intelligence into human society.[42] This view would, of course, stand in direct opposition to the aforementioned arguments that heed warning against technological integration and dependence. Bjork, while maintaining that humanity has a unique teleology, does not espouse fear or disdain of superintelligence or its societal integration. In his view, "there is no inherent theological conflict between a biblical view of personhood and work in artificial intelligence, nor would successes in this field undermine human value or the doctrine of the image of God."[43]

Secular AI Ethics

Let us now turn to salient secular perspectives on the same topic. Like under the umbrella of Christianity, there exists a good degree of variance within secular circles apropos of the moral standing of AI, or lack thereof. AI ethics is, of course, an entire field of Philosophy and Policy, thus it would be impossible to fully analyze all possible points of view. I will focus, thus, on those that most heavily borrow from a Neo-Darwinian or physicalist philosophy of mind to inform their ethic. To the physicalist, any quality of a living creature–human or otherwise–could theoretically be replicated artificially, as we have established. Thus, any quality X in a biological

[42] I should note that I am extrapolating from a work that was published long before the horizon of superintelligence, and that this is *my* reading of Gustafson.
[43] Bjork

creature that serves as a criterion for moral status could be equally present in an artificial being. As such, much of the discussion has been in the determination of quality X, or quality set X.

Peter Singer, for example, is known for his non-anthropocentric utilitarian framework of ascribing moral status. Singer's primary criterion for moral value is an agent's capacity for suffering and/or happiness. As such, the interests of countless beings besides humans ought to be considered in any moral calculation. Practically speaking, this ethical perspective finds Singer and his followers passionately advocating for animal welfare and liberation, particularly from the suffering induced by factory farming, hunting, poaching, and meat consumption in general. As Singer argues, it is undeniable that a cow, pig, or chicken experiences tremendous suffering when bred for slaughter and consumption.[44] Not only are the deaths painful and gruesome, but the living conditions are outright torturous. Why, however, does Singer prioritize the suffering of advanced mammals like cows and pigs more than the, say, billions of *insects* that are killed every year? Under Singer's framework, capacity for suffering and cognitive capability are roughly commensurate spectra. That is to say, the intelligence or sentience of an agent is directly proportional to its capacity for pain and pleasure.

This assumption is not arbitrary, as Singerians take a physicalist approach to understanding pain and pleasure; namely, they pay attention to the neurological, anatomical, and psychological factors that contribute to a creature's capacity to experience pain and pleasure. Given that this assumption is, thus, scientifically accurate, it follows that a cow's neurochemistry is much more complex than that of a centipede, allowing for the cow to have a more intense experience of pleasure and pain. The cow, thus, should be given more weight in whatever utilitarian moral calculations we may conduct. This framework is what justifies, according to Singer, the weighing of human interests as greater (per capita) than those of less complex

[44] Singer, Peter. "Animal liberation." In *Animal rights*, pp. 7-18. Palgrave Macmillan, London, 1973.

animals. I say per capita because, though the interests of one human outweigh those of one cow, questions arise when we compare one human to 100 cows, 1000 chickens, or 1 million insects, ad infinitum. Following this justification, it would follow that any being possessing equal sentience to a human–thus equal capacity for suffering–ought to be given equal moral weight. Further, a being possessing *greater* sentience ought to be granted *greater* moral standing. For this reason, many scholars who operate under Singer's utilitarian framework advocate for granting of moral standing to artificially intelligent beings, commensurate with the level of sentience thereof.

One might cast doubt on the proposition that an AI could experience sensations like suffering and happiness, however. That is, though an AI could achieve equal or greater intelligence, it could not feel pleasure or pain as we do. To this, I have two responses. First, many scholars, both Christian and secular, mentioned in this paper and beyond, hold that emotional experience, the presence of subjective desires and interests, a sense of embodiment and the self-preservation thereof, etc., are necessary properties of human intelligence. As such, to say that an artificial intelligence would be "equal" to human intelligence would imply an equal capacity for said properties. Second, we must recall that we are working with Neo-Darwinian assumptions. That is to say, the biological processes that allow for an animal's experience of pleasure and pain, or happiness and suffering, cannot be more than the evolutionary history thereof. If human suffering is "real," it is naturally explainable, and can be explained by its physical composition. Operating under this physicalist framework, the artificial replication thereof is indeed theoretically possible.

If moral standing is taken as commensurate with cognitive capacity, things grow more complicated as artificial intelligence, as Kurzweil puts, "roars past" human intelligence. Nozick, all the way back in 1971, ran with this hypothetical and coined the idea of a "utility monster."[45]

[45] Nozick, Robert. 1974. *Anarchy, state, and utopia.* New York: Basic Books.

The utility monster is a hypothetical artificial intelligence that is designed with the sole purpose of maximizing its own utility. Given its computational power and self-replicating nature, the monster would be able to consume resources such that its utility:suffering ratio was orders of magnitude beyond any other being–artificial or biological. Under a Singerian utilitarian framework, the monster would be most deserving of moral weight, given its superlative capacity for experiencing pleasure and pain. The morally righteous choice, still within this framework would be to devote *all* resources to the continual replication and flourishing of the monster, such that all other life would go extinct. This hypothetical is exemplary of the "convergence" phenomenon in AI theory; put simply, the idea that an AI's directives, unless given specific bounds in the development stage, will necessarily be interpreted as boundless.[46] Obviously, this is a troubling scenario, and one that calls into question the idea that sentience should be the primary criterion for moral standing. Utilitarianism usually has a way of inspiring troubling hypotheticals when its principles are taken ad absurdum.

Moving on from visions of human apocalypse, there is merit in Singer's non-anthropocentric approach to moral standing. Operating within physicalist bounds, it would be arbitrary to posit that the experiences of pleasure and pain could not be replicated by artificial intelligence. Beyond utilitarian frameworks, *any* property or criteria of moral status found in humans or other creatures (whether it be dignity, agency, etc.) could theoretically be reproduced by an AI. Thus, no matter the ethical framework, a physicalist cannot write off the moral standing of artificial beings on ontological grounds alone.

Cockelbergh, another secular scholar, seeks to do away with discrete and specific criteria for moral standing altogether. Rather than listing some quality X or quality set X as indicative of moral standing, Cockelbergh argues instead for a "relational approach." In this approach:

[46] This discussion of this phenomenon is most attributed to Nick Bostrom

> "The idea is that an entity cannot be defined without reference to its relations—both social and natural relations (and more precisely, also mixtures of these). For example, a particular animal has its place in the ecosystem and in the webs of social relations with other animals. These relations also have a history and are tied to specific places, habits, and things. To define moral standing in isolation from these relations is itself a moral violation, since it takes as its departure an abstract 'entity' with "properties.'"[47]

As such, the moral status of artificial beings ought to be determined given its relations with other beings, rather than on the basis of some cognitive capability. Cockelbergh also critiques the notion that the ascription of moral standing could ever be considered an epistemologically infallible practice. Given that we, humans, are the ascribers of moral status, our limited subjective perspective will undoubtedly lead to biased and, more often than not, self-serving designations of moral value. How person X interacts with another being (read: their relation thereto) is critical to how person X determines that being's moral standing.

Thus, it is critical to be ever-cognizant of the ways in which our subjective experience influences our attitude towards AI. Plenty of science fiction, for example, paints a picture of apocalypse caused by superintelligence. Other artistic works, however, paint AIs as potential companions, friends, lovers, or even Gods. Once AI becomes more integrated into the quotidien of you and me, the specific nature of the human-AI interfacing will be instrumental in shaping our attitudes and ultimately ascription of moral value. He brings up Levinas's concept of the "face of the other," and asks whether or not an artificial being could have a face worthy of ethical obligation. Would an AI be able to replicate the power of the "human presence" that Levinas describes? What does this "presence" feel like with beings other than humans? Would an AI that

[47] Coeckelbergh, M. The Moral Standing of Machines: Towards a Relational and Non-Cartesian Moral Hermeneutics. Philos. Technol. 27, 61–77 (2014). https://doi.org/10.1007/s13347-013-0133-8

replicates human aesthetics receive more favorable treatment than those that don't? Are we biologically more responsive to beings that portray a similar aesthetic to that of our own kin? Intuitively, the answer seems to be yes. This does not mean, however, that such aesthetic biases are ethically sound. To this end, Cockelbergh points out the epistemological gap between abstract moral theory (or "moral science" in his words), and empirical experience. Any theoretical prescription, thus, of a "proper" way to grant moral status–*that lacks direct experience*–is insufficient.

This framing of moral status as a socio-relational byproduct could help to elucidate why some Christian scholars are so theoretically antagonistic towards the prospect of superintelligent AI. That is to say, in the disdain for technological idolatry, or in the fear that AI will make human existence too leisurely, such feelings are projected onto these scholars' theoretical ethical arguments. The same could be said, turning to the other coin, for transhumanists and secular scholars (e.g. myself!). In the pursuit of expanding their ethical perspective, they determine those moral assumptions that would benefit their work and their personal interests. Of course, lines between personal interest, supposed ethical obligation, and religious faith are always going to be ultimately indistinguishable at the granular level. Nonetheless, Cockelbergh's nuanced conception of moral standing is illustrative of the shortcomings of both Christian and secular scholars who cling to a strict properties-based approach.

Conclusion

The purpose of this thesis is not to interject my personal ethical arguments into existing literature. Rather, I have attempted to discern, compare, and critically analyze the leading scriptural interpretations, theological principles, and philosophical arguments put forward by

Christian scholars that seek to answer major questions regarding the nature of intelligence–both human and artificial–and the ethical implications thereof. I have also juxtaposed these arguments with those of leading secular scholars and schools of thought, attempting to highlight areas of potential conflict or congruence. As I have mentioned, it is my view that anthropocentric theological views that are *conditional* upon a discrete set of unique human properties are theoretically weaker and less convincing than those that are *unconditional.* That is to say, plenty of scholars' possess a conditional perspective of human ontology that is wholly incompatible with the recognition of superintelligent AI as having equal-or-greater rational faculties. This incompatibility is often explicitly stated. Dembski, for example, states that "I fully grant that my theology would crumble with the advent of intelligent machines."[48] Reichenbach, similarly, argues that "strong AI presents serious challenges to a Christian understanding of personhood, including our freedom to choose, moral choice itself, self‑consciousness, and the relevance of God to our beginning, being, and ending."[49]

The issue of Christian perspectives' congruity (more specifically, the lack thereof) with contemporary scientific and/or technological achievements is not a new one by any means. Though these historical events are often dramatized in primary school classes, the Copernican Revolution and subsequent trial of Galileo, along with the Darwinian Revolution of the late 19th/early 20th century come to mind as historically salient. In either case, there were theological assertions that were undercut by more naturalist hypotheses backed by empirical evidence. Despite the ostensible incompatibility (though it was not across the board), the theological perspectives that were consistent with Copernican or Darwinian thought came more into the forefront over time. Of course, there still exist creationists, and perhaps even some geo-centrists,

[48] Dembski, William. "Conflating Matter and Mind," Perspectives on Science and Christian Faith 43, no. 2 (June 1991): 109
[49] Reichenbach

yet these individuals are often ridiculed and their assertions are not met with any legitimate respect. I might argue that these historical examples may be illustrative of the coming reckoning some Christian scholars may have to face if the intelligence explosion comes to pass. Some theological perspectives will, as mentioned above, "crumble," yet some can stand strong even with the theoretical advent of superintelligent AI. Given the diversity of opinion, and versatility of scriptural interpretation, it is likely that the theological perspectives that are consistent with superintelligent AI would become more culturally and socially dominant. Niebuhr and Schurmaan's holistic view of human transcendence; Bjork and Gustafson's emphasis on God-creation *relationships* rather than ontology; and Peter's focus on agape love (among others) would be more congruous with the intelligence explosion. Those views that see AI as an impossibility, or those that see rational faculties as solely human, would more likely fade into the background as creationism and geo-centrism once did.

The inundation of ethical issues post-explosion is also likely to cause a reckoning among both theological and secular scholars. Once again, I would argue that those beliefs which draw hard and fast ethical boundaries, placing humanity upon an axiological pedestal, are the most likely to be replaced if intelligent machines display a capacity for emotion, suffering, and happiness. I could be wrong, however; general anthropocentrism is far more resilient than any one religious invocation thereof. We may see similar invalidation of machines' moral standing as we do with animals today. There also exists historical precedent, however, for misplaced anthropocentrism. That is to say, there have been countless erroneous conceptions of what is 'human' that have led to erroneous normative ascriptions of moral worth. Throughout human history, people of different genders, races, religions, physical abilities (the list can go on and on) have been subjected to oppression under the assumption that these groups are less-than-human.

Plato’s political philosophy was built upon the axiological assumption that women were a sub-species of man, for example. In fact, the basis of his entire Republic illustrates a zoological taxonomy of different human groups, all assigned different moral value and societal roles. Who's to say that our contemporary criteria for personhood, however, is not as arbitrarily strict as that of Plato? Will 23rd-century philosophers ridicule *our* self-preferential ethics?