Note to the reader/user: You are welcome to make use of this commentary as you deem appropriate. All I ask is that you make proper attribution of it in your work and use the following entry on your Works Cited page (include the URL, of course):

Buaas, John. "Some Comments on *Ex Machina*." *Domestic Issue*. 24 Mar. 2020.

Some Comments on *Ex Machina*

by

John Buaas

Professor of English and Humanities, Butler Community College

These notes are meant to serve as a kind of thematic companion for you as you watch or re-watch writer/director Alex Garland's 2014 film. I presume you have seen the film and know its plot, so I will take no precautions in hiding "spoilers" from you. Though I have some opinions about this film, these comments are not meant to offer up some sort of definitive "take" on it, nor are they intended to take the place of watching it—indeed, given film's nature as a visual medium, some of these comments will make little sense unless you're watching. But more to the point regarding this film in particular: while I'll be discussing some of its visual elements, *Ex Machina*, though officially a science-fiction film about robots, is most definitely not your typical action-packed, laser-blast kind of film. It takes very seriously the big themes at the heart of research in Artificial Intelligence and the implications of that research, which means that it has a whole lot of talking in it. Thus, it rewards close attention, as well as some thinking about those themes.

One final comment, and then we'll move on: for me, *Ex Machina*'s appeal for me lies in the fact that, as many times as I've seen it and discussed it with students now, its central question—whether the robot Ava truly is conscious in the same way(s) humans (say we) are—remains an open one. In the world of the film, Caleb has decided (and Nathan agrees with him) that she is, which dictates his actions in the film's final act. But in the world you and I inhabit . . . I don't know and, for that matter, I don't know Alex Garland's intentions regarding whether, in his creation of Ava, he has created a sentient machine. (I don't know if Garland has said one way or another, and I much prefer not knowing.) What I can say is that I think you can make a strong case either way, for reasons we'll touch on in these notes.

****

Cast (in order of appearance, with some observations regarding each):

Caleb Smith (Domhnall Gleeson): Caleb is an employee of Blue Book, Nathan's Google-like internet search-engine company, who has won the prize of a week at Nathan's house. It turns out that the real intent is for the winner to conduct a Turing Test. (We will learn later that Nathan has actually deliberately selected him.) We learn that Caleb is 26; single; an only child; originally from Portland, Oregon; presently living in Brookhaven, Long Island; and was orphaned when his parents were killed in a car crash (he was a passenger in the car). He picked up computer coding while recovering in the hospital from the accident. Caleb is more than just a coder, though; he is also well versed in high-end research on Artificial Intelligence, and shows some knowledge of theories of mind as well. Curiously, his musical tastes seem to run toward '80s-era New Wave: at one point, we hear him listening to "Enola Gay" by Orchestral Manoeuvres in the Dark, and he tells Ava that he prefers Depeche Mode to Mozart.

For what it's worth, my students have usually liked Caleb and are sympathetic toward him. That said, I have had some who dislike him. One, in fact, made the case recently that his being "a good kid," too innocent and/or naïve, and too deferential toward Nathan, are actually failings on his part. (Well, sure: innocence often gets people into trouble in various ways, but we don't usually think of it as a fault.) Another once just flat called him an "incel," but walked that back some by saying that Caleb, by wanting to help Ava escape, believes that he can certainly protect her and possibly even control her, and so is really no different from Nathan in that regard. While I personally would not make this argument, there's nothing in the film that argues against it. Perhaps Ava *is* right to be suspicious of him.

Nathan Bateman (Oscar Isaac): Nathan is in many ways the stereotypical uber-rich tech-bro. He is a computer-programming genius who started and owns Blue Book, the search-engine company that employs Caleb. He lives in near-isolation in a sleek house/research facility on an enormous estate. He is smart, and he is not shy about letting Caleb know he is smart. His overbearing self-confidence is a source of humor in the film, but it may also be his Achilles' heel: his programmer's mindset causes him to think he knows what Caleb is feeling and thinking, but I feel fairly certain he does not believe Ava (and certainly not Kyoko) would kill him. He is also enigmatic. He eats well (he appears to prefer a vaguely Japanese diet of fish and rice) and is something of a fitness freak, but he also drinks to excess; he is driven to create a truly intelligent machine, yet he seems oddly detached regarding his own role in advancing technology, seeming to imply his role in building Ava is analogous to that of a mutant gene's being passed on and expressed in succeeding generations. Nathan's tastes in music are eclectic: classical chamber music plays throughout the living-spaces of the house, but in the out-of-left-field dance scene with Kyoko, Nathan selects Oliver Cheatham's 1983 post-disco minor hit, "Get Down Saturday Night." (Like Caleb, Nathan's tastes in popular music seem not to have advanced past the 1980s.)

Nathan's last name, Bateman, is not revealed in the finished film but does appear in the shooting script. Perhaps (just speculating here) his surname is a pun on "bait-man": someone who baits, who sets a trap. In a film whose big themes are awareness and self-awareness, perhaps the man who

fancies himself to be more self-aware than anyone else ends up, with his creation of Ava, making himself the bait in the trap he has set for Caleb.

As you might imagine, not many of my students like Nathan. Some, though, do admire his intelligence and arrogance in that they see those attributes has having pushed Nathan to achieve what he has so far accomplished. They see Nathan as another in a long line of brilliant, strong-willed types who, by virtue of their at times brutal insistence on adhering to their vision, thereby advance the cause of science.

Ava (Alicia Vikander): Ava, whose name seems to echo the name Eve (the woman in the Judeo-Christian myth of creation in the Book of Genesis), is the robot who is the catalyst for the film's central plot of the Turing Test examination. In their first meeting, she asks Caleb if he would like to know how old she is and says, "I am one." When Caleb asks her, "One year, or one day, or . . . ?" she simply says, "One." She says that she has not had to learn language but has always known how to speak. Nathan will tell us that he has programmed Ava to be heterosexual, and she chooses to dress herself for Caleb in very feminine-looking clothes (though, for what it's worth, those appear to be the only clothing options Nathan has provided for her). Ava draws pictures "every day," she tells us, but she initially says she does not know what her subjects are. (Art—or, more precisely, the artistic impulse—is an important if understated theme in the film, about which I will have some comments later on.) She deeply dislikes Nathan ("Is it strange to have made something that hates you?" she asks him) and, even as she draws Caleb into her confidence in her effort to escape from Nathan's house, is ultimately deeply suspicious of him as well: "An advanced programmer . . . like Nathan." None of these things is a secret, yet Caleb's challenge as the tester (and ours as viewers) is what to make of all of this as we ponder the question of Ava's sentience. In his discussion of *Ex Machina*, the host of the YouTube channel *Lessons from the Screenplay* makes the astute observation that the viewer of this film is in a position analogous to that of the reader of *The Great Gatsby*. We learn what we learn of Ava and Gatsby as filtered through a mediating conscience, Caleb and Nick Carraway, respectively; thus, when it comes to determining the truth about those characters, the audience/reader is at something of a disadvantage. I would just add to this, though, that it is precisely that disadvantage that both makes it difficult to decide the question of Ava's sentience and keeps us engaged with this film through repeated viewings.

Moreover, Caleb has a past—a childhood, adolescence, and young adulthood—that he can recall and that all of us can recognize because of our shared humanity with him. But Ava, like the goddess Athena who sprang fully-formed and in full armor from the head of Zeus, has no such past; hence her answer of "One" when she says how old she is, and hence her never having learned language but always having known how to speak. Nathan says he has programmed her to be heterosexual and says in so many words that that—programming—is no different from how humans' sexuality comes about; but notably lacking from that characterization is, again, our past, evolving, physical and emotional engagement with our sexuality that, for better and for worse, has led us to where we are now. All these matters and more become even more complicated when we learn late in the film that Nathan has programmed Ava to try to persuade Caleb to help her escape from Nathan's house: is she indeed just "pretending to like" Caleb, as Nathan asks in a possibly-rhetorical-or-possibly-not manner. But (to leave the film's world and return to our own for a moment) these also happen to be (or should be) very real, live questions for researchers building intelligent machines designed to

interact with humans. For a solid, readable engagement with these questions from the perspective of developmental psychology, see the blogger ShrinkWrapped's post "Empathy and AI: Part IV" (though the entire five-part series is well worth your time if you're interested in these matters).

My students over time have tended to be evenly divided on the question of Ava's sentience; how they land on that forms the grounds for their attributing a motive for her abandoning Caleb to die. Those who think of her as only the sum of her programming see that decision as, well, something a machine would do. Those who think of her as sentient either think of her as being essentially autistic or socio-pathological in her manner, or they understand her action as a combination of both the cruelty toward those who show us kindness that all humans are capable of and Ava's lack of experience with humans (Ava, like Caleb, is innocent in her way, hers arising from a lack of experience: she equates Caleb-as-advanced-programmer with the only other human she's known, Nathan-as-advanced-programmer; she hates Nathan; therefore . . .). Later in these comments, I'll lay out the evidence the film presents us that argues in favor of and against her sentience (with the recognition that we can see some of that evidence as supporting either claim).

Kyoko (Sonoya Mizuno): The robot named Kyoko (Nathan's Japanese cook, server, occasional dance partner, and lover) literally never says a word during *Ex Machina*, but over the four years now that I have watched this film with my students, she has become more and more fascinating to me as a character—and not just because she stabs Nathan first at the film's climax. Nathan is at his most controlling and brutal and, even, inhumane where Kyoko is concerned. In the evening dinner scene when he explodes in anger at her for knocking over Caleb's wine glass and Caleb tries to reassure her, Nathan says that there's no point in talking to her since "she doesn't understand English" but then goes on to say that that means he can't tell her when he's angry with her: as if to suggest that her incomprehension of English is tantamount to incomprehension, period. We and Caleb of course do not at that point know Kyoko is a robot, so Nathan's behavior toward her is shocking. Yet I would make the argument, in light of Caleb's discovery of Nathan's videos of earlier iterations of robots, that Nathan's rendering Kyoko incapable of responding verbally is even more shocking: it is his attempt to deprive this particular intelligent machine, at least, of agency. Nathan, clearly having tired of the notion that sentient machines he's built might be angry that he won't let them leave his house, creates, or believes he has created, his version of a dream woman: she cooks, cleans, dances, and has sex with you whenever you want, *and* she literally cannot talk back to you.

But even though Nathan has done what he can to render Kyoko a non-entity, the film clearly does not regard her as such. As I point out to my students when we watch the crucial scene in which Caleb and Nathan discuss sexuality and AI, the men's immediate context for that discussion is Ava, but at crucial moments the camera is in tight on Kyoko's face as she cuts up a piece of fish, the men's conversing going on in the out-of-focus background. This discussion just as clearly applies to Kyoko, we will retrospectively realize, and we would be forgiven for wondering what she is thinking about as she listens. From that scene on, Kyoko appears with some frequency in the film: watching Nathan as he watches the session between Caleb and Ava in which Caleb tells her he is there to determine whether she has a conscience; staring intently at Nathan's Jackson Pollock painting; watching via a monitor as Caleb, after learning of the other robot prototypes (and that Kyoko is herself a robot), has something of a psychotic break and cuts open his forearm and punches the

bathroom mirror. Finally, it will be Kyoko who seeks out Ava in the film's last act, setting into motion the robots' alliance against Nathan.

My students universally express sympathy for Kyoko, but they usually are not otherwise as intrigued by her as I am. That said, one of the better papers I have received from a student on this film makes the argument that Kyoko is actually the subject of the film's Turing Test. Though I would not frame things quite that way, I do think one could make the case that the film shows Kyoko changing in ways that we don't see happening with the other characters.

What follows is not a scene-by-scene commentary but, instead, a discussion of some of *Ex Machina*'s recurring visual motifs and themes, listed in the order that they first appear in the film. If you read this straight through, then, you'll see some looping back and forth in the film's chronology.

The Title

The film makes use of or alludes to several different stories from Greek, Judeo-Christian, and Hindu mythic traditions, some of which I will discuss in more detail later on. Indeed, the basic story beneath its plot, a person's desire to create a human or human-like being from inanimate material, is, up to a point, that of the Greek myth of Pygmalion: a woman-hating artist who, despite that hatred, sculpts a woman so beautiful and life-like that he falls in love with it and asks the goddess Venus to endow it with life (Hamilton 108-111). But while those references are in the film, it's hard to know just how hard to insist that they somehow matter to the characters in the story, or whether they should matter to the viewer. The references are there, or appear to be, but *Ex Machina*'s world is clearly secular: when Caleb asks Nathan why he built Ava, Nathan doesn't say that to do what the gods can do has been a dream of humankind from time immemorial; he just shrugs and says, "Wouldn't you if you could?"

Surely someone has made the following argument before, but I have not seen it. So, to borrow a phrase from Nathan, "I'm just gonna throw this out there so it's said, okay?" but without any claim as to its originality. Many people have noted that the film's title comes from the ancient theatrical and literary device called "deus ex machina" (literally, "god out of the machine," in which a god (in Greek drama) or other person suddenly and often implausibly arrives to rescue the hero from a difficult situation) (Cuddon 216-217). My question, though, is, "This is such an obvious reference, so why is 'deus' removed from the title?" The answer, it seems, is equally obvious: there are no gods to rescue us. The gods have, quite literally, been removed from the machine that used to appear miraculous; now, the machines we have built, and the dilemmas they have created for us, are of our own devising. Whether we can rescue ourselves is also up to us.

There's a passing, joking reference to all of this when Caleb says that if Nathan has built a conscious machine, "That's not the history of man, that's the history of gods" and then later Nathan (intentionally?) misquotes Caleb as having said, "I'm not a man, I'm a god." More darkly, though,

the theme reappears later. As Caleb and Nathan discuss the latter's motives for building Ava and plans for the next iteration of Ava, Caleb quotes Robert Oppenheimer, the man who developed the atomic bomb, and his remembering the lines from the Bhagavad-Gita, "Now I am become Death, destroyer of worlds," when he witnessed the detonation of the first bomb at the Trinity Test Site in New Mexico. Nathan not only knows that quote, he'll quote still more from that same passage (which Oppenheimer also knew) in a rare moment of genuine, if drunken, reflection on what he has done—note in particular the last sentence, repeated three times: "In battle, in forest, on the precipice of the mountain, on the... the great, dark sea. In the sleep, confusion. In the depths of shame... The good deeds a man has done before defends him. The good deeds a man has done before defends him. The good deeds a man has done before... defends him" (qtd. in "*Ex Machina* Quotes"). Oppenheimer did not believe in the divine and thus could not know whether his actions conformed to a transcendent, cosmic order; the best he could hope for was that he would be remembered and honored for his other, good deeds (Temperton). Nathan is also very much in that same position.

No one refers to the German philosopher Friedrich Nietzsche in *Ex Machina*, but it seems appropriate to quote him here, from "The Parable of the Madman":

> God is dead. God remains dead. And we have killed him. How shall we comfort ourselves, the murderers of all murderers? What was holiest and mightiest of all that the world has yet owned has bled to death under our knives: who will wipe this blood off us? What water is there for us to clean ourselves? What festivals of atonement, what sacred games shall we have to invent? Is not the greatness of this deed too great for us? Must we ourselves not become gods simply to appear worthy of it?

Mirrors and glass

The very first image we see in *Ex Machina*, an interior shot of the Blue Book headquarters, is a very brief (perhaps one-second long) shot of a few people walking, their images refracted by several panes of glass. It is as though we are peering at them through a faceted diamond or quartz. That same image will be echoed—reflected, as it were—by the film's very last image: Ava's standing in a crowd in the city and then suddenly vanishing into a mass of people reflected in the glass of an office building. In between those two moments, glass walls (both transparent and translucent, rectangular-paned and asymmetrically-shaped, smoothed and fractured) and mirrors are just about everywhere we turn in Nathan's house, most notably in the space where Caleb's sessions with Ava occur. The film's set designers talk at some length about that particular space ("Through the Looking Glass")—it is clearly of great importance to them.

When I discuss that space with my students, I remind them that Caleb's goal, to determine whether Ava has a conscience, is complicated by the fact that in an empirical sense, we are not even sure that humans have a conscience. All we have to go on when we talk about human conscience is a combination of faith and our own and other people's say-so. The glass that separates Caleb and Ava in those sessions is like the dilemma we find ourselves in when talking about conscience: it seems to

be right there, real and tangible, but we cannot (not yet, anyway) get through that final barrier that keeps us from knowing with certainty.

The Minotaur

(What follows is one of those instances in which it's hard to know just how much to insist that this is a Thing in the film.)

Bull's heads are mounted on walls in three different places in Nathan's house. This fact, combined with Nathan's description of Ava as "a rat in a maze" and more than a few camera shots suggestive of entrapment, leads me to wonder if the film is encouraging us to think of the Greek myth of the Minotaur, the monster in the labyrinth on the Mediterranean island of Crete.

The Minotaur was half human and half bull, the monstrous offspring of King Minos' wife Pasiphaë and a bull that she had fallen in love with. (This bull had been a gift to Minos from Poseidon, the god of the sea. Rather than sacrifice it to Poseidon as the god had intended, though, Minos instead took it back with him to Crete. Poseidon, to punish Minos, caused Pasiphaë to fall in love with the bull.) Rather than kill the Minotaur when it was born, Minos had the famous architect and inventor Daedalus construct a labyrinth for the Minotaur, a maze so intricate no one once placed in it can find their way out, and every nine years Minos would demand a sacrifice to the Minotaur of seven virgins and seven young men from Athens. A young man named Theseus, with the ingenious help of Daedalus and a young woman named Ariadne, is able to kill the Minotaur and find his way out of the labyrinth (Hamilton, 151-152). Minos, determining that Daedalus must have helped Theseus, imprisons Daedalus and his son Icarus in the labyrinth. You probably know the rest of this story: Daedalus and Icarus fashion wings made of wax and feathers to fly out (the maze was open to the sky); Daedalus warns Icarus not to fly too close to the sun; Icarus, in his excitement, forgets the warning and his wings melt, causing him to fall into the sea and drown.

You can see the problems with insisting too strongly on exact parallels between *Ex Machina* and this story. Who is Minos, who is the Minotaur, and who are Theseus and Ariadne—or are there even Theseus and Ariadne figures in the film? There are no good, firm answers: even if we say, well, Nathan, like Minos, seems to be disobeying some grand cosmic edict, the boundaries between humans and gods (and monsters) in the myth remain intact—but one of the implicit assumptions of *Ex Machina* seems to be that those old boundaries no longer exist. As I said earlier in the discussion of the film's title, our culture is post-myth: while, yes, these stories still float about in our shared cultural consciousness, they are uprooted now from an understanding of culture whose social and moral codes are at least partly rooted in a shared religion. They work these days more like memes than like something that anchors us to a distant past via belief in and obedience to the divinity that our forebears had worshiped.

The Turing Test

Scenes like the one in which Nathan asks Caleb what a Turing Test is and Caleb tells us are known as *exposition*: they are in the film solely for the audience's benefit. Of course *they* both know what it

is, but it may be that someone watching the film does not. It's crucial that the audience knows, though, or else there goes the film's plot; and so the filmmakers have it in there, just to be sure. To their credit this is the only such scene in the film. Two signs of a good film are when its makers trust the audience to be attentive and, by withholding what would otherwise be obvious information, invite the audience to do some work as well by using their imagination to fill in that information. *Ex Machina* does both these things very well—see *Lessons from the Screenplay*'s discussion of the film to see two specific examples.

(Disclaimer regarding what follows: I am not an expert in Artificial Intelligence.) I have always thought that the Turing Test as usually configured is potentially flawed in its conclusion—what if the person who concludes that the computer is intelligent is themselves deficient in some way? What if that person is especially susceptible to suggestion, for example? Whether Nathan would agree with this, I can't say, but his version of the Turing Test is something like a response to this flaw: Nathan reveals that it's really Caleb who is the subject. Would he, "a good kid . . . with a moral compass," as Nathan describes him, be lured by Ava into helping her escape? (Nathan's getting Caleb to stop asking about how Ava works and instead focus on how he feels about her—his emotional response toward her—is part of all of this as well.) So, for reasons discussed elsewhere, both Nathan and Caleb conclude that because Ava was successful in persuading Caleb to act on her behalf, she does indeed possess AI. I remind you, though, that just because *they* reach that conclusion doesn't mean that *we* also have to.

Caleb alludes to something of a corollary to the Turing Test, or a still more profound version of it, when he tells Nathan before the second session that just talking with Ava is like testing a chess-playing computer by just playing chess: the game may make good moves, but that can't tell the tester if it knows it's playing chess, or even if it knows what chess is. Caleb decides that determining whether Ava indeed understands what she is saying is the test Nathan really wants him to conduct. This is the essence of John Searle's argument that true (or "strong") AI should be able to demonstrate "understanding" of what it is doing; in his 1980 essay "Mind, Brains, and Programs," Searle presents the thought-experiment known as the Chinese Room to demonstrate that just because a computer responds correctly or appropriately to data fed into it doesn't prove that it understands what it is doing or what its output means (3-4). Even though that essay is 40 years old now (almost the equivalent of pre-history in this field), it still provokes arguments in support of and in opposition to it among AI researchers and theorists—a sign that if you're at all interested in this subject, it's still a good essay to read and think about.

Art-making, AI, and intention

What causes people to make art is a recurring subject in *Ex Machina*. This may seem like an odd secondary theme to have in a film about Artificial Intelligence, but I have become more persuaded over time that this film invites us to see art-making and advancing technology as arising from the same basic, ultimately inexplicable impulse.

Surprisingly, Ava introduces the subject in the second session when she shows Caleb an abstract drawing she has made and asks him what it is, hoping that he would be able to tell her. In articles on the film, I have seen others make comparisons between it and either Nathan's Pollock painting

or the interior of the brain Nathan has designed for Ava; personally, I think it most resembles Ava's own mesh covering. Given that most of the time, the visual artist has some sort of intent with regard to the object or subject the work represents, Ava's saying "I make drawings every day, but I never know what they are" seems significant, and not just from the standpoint of aesthetic theory. There is nothing exactly wrong, by the way, with Ava's hoping that Caleb might know what she's drawn: there's a whole body of aesthetic theory that argues in different ways that, the artist's intention aside, it's the audience who makes meaning out of a work and thus "completes" it. Even so, her inability to discuss her drawings, even just to explain why she draws, seems rather odd. One way you can see this is as a sign of Ava's emerging consciousness—my three-year-old son could not, a while back, tell us what he was drawing, but now he draws with intention and can tell us that. However, the fact that she will make two other drawings, both of them the result of her interactions with Caleb (the second one of the tree in her room after he suggests that she choose something to draw; the third a portrait of Caleb that bears a striking resemblance to Caleb's image on his keycard(!)), can cause us to think of her and Caleb as a kind of Generative Adversarial Network, or GAN. As you read the paragraph below, think of Ava as the "generator model" and Caleb as the "discriminator model":

> GANs are a clever way of training a generative model by framing the problem as a supervised learning problem with two sub-models: the generator model that we train to generate new examples, and the discriminator model that tries to classify examples as either real (from the domain [a data set]) or fake (generated). The two models are trained together in a zero-sum game, adversarial, until the discriminator model is fooled about half the time, meaning the generator model is generating plausible examples. (Brownlee)[1]

"This is all so super cool," as Nathan might say. However, even though a GAN produces content based on a kind of dialogue between two machines, and even though that content is genuinely original, that does not mean that those machines understand (in John Searle's sense of that word) what they are doing.

The subject of intentionality in art-making comes up again when Nathan and Caleb discuss Jackson Pollock's approach to painting—a discussion prompted by their discussion of whether AIs need a gender and, by extension, sexuality (see below). Nathan describes that approach as "automatic art": not completely random, but not pre-determined, either. That leads to this exchange:

> **[Nathan:]** What if Pollock had reversed the challenge. What if instead of making art without thinking, he said, "You know what? I can't paint anything, unless I know exactly why I'm doing it." What would have happened?
>
> **Caleb:** He never would have made a single mark.

---

[1] Coincidentally, the first article describing GANs was published in 2014: the year *Ex Machina* was released. Also coincidentally, GANs have been used most often to generate images of various sorts: non-representational art, images of actual objects, and even pictures of people who have never actually existed. Regarding this last: in the world of the film, we learn that Ava's face, which Nathan created based on the online pornography Caleb had looked at, would be an example of what a GAN can do.

> **Nathan:** Yes! You see, there's my guy, there's my buddy, who thinks before he opens his mouth. He never would have made a single mark.
>
> [pause]
>
> **Nathan:** *The challenge is not to act automatically. It's to find an action that is not automatic.* From painting, to breathing, to talking, to fucking. To falling in love...("*Ex Machina* Quotes," my italics added)

So, what does all of this have to do with building AIs? Here in the italicized passage I think we see the link, at least from Nathan's point of view. I mentioned in both my discussion of his character and the "Title" section above that Nathan, though clearly driven to build an AI ("Wouldn't you if you could?"), presents himself as motiveless for doing it. He doesn't know exactly why he's doing it.

There is much, much more to be said here, but I will just let it be said for now that whatever else art is supposed to do, everyone except totalitarian leaders agrees that art should bring us pleasure and enjoyment but is otherwise not absolutely necessary to human existence—it doesn't have to "do" anything else—and it's fascinating to see the motive driving the pursuit of cutting-edge technology of the sort the film presents us with talked about in pretty much the exact same way art is talked about.

"Other minds" theory

In his conversation with Nathan after the important second session (Ava shows Caleb the first of her drawings, prompts him to share personal information about himself and, during the power cut, tells him that he shouldn't trust Nathan), Caleb says that Ava's making a joke was the best sign yet that she is conscious. He argues that she couldn't make the joke without having some awareness of her own mind, which means that she probably has an awareness of mine.

We all say and act as though we and others are conscious beings. However, in the absence of empirical evidence that there is such a thing as consciousness, as I said earlier, we cannot truly know whether other people have minds. You can imagine, therefore, the difficulties inherent in determining whether a machine is conscious. Yet, this is what the Turing Test and Searle's insistence that a computer have understanding in order to be considered intelligent add up to: the human conducting the test must conclude that they are not interacting with a machine but with another human.

This is the philosophical question known as the "other minds" problem:

> Philosophers, thus, find themselves saddled with the question, How do I know that others have minds? One can put the emphasis on the word "know" and raise a sceptical question parallel to the one Descartes raises in connection with our knowledge of the external world, or one can put the emphasis on the word "how" and question the source of our knowledge given that I am not in a position to have direct knowledge of the mind of another. (Avramides)

In case you are curious, Caleb, following the philosopher John Stuart Mill, seems to be making the "argument from analogy" case that Ava has a mind: Mill states, "First, [other people] have bodies like me, which I know in my own case, to be the antecedent condition of feelings; and because, secondly, they exhibit the acts, and outward signs, which in my own case I know by experience to be caused by feelings" (qtd. in Avramides). As Anita Avramides goes on to note in her essay on this subject, though, the chief flaw in this argument is the same as that which I argue is intrinsic to the Turing Test as traditionally configured: it begs the question by presuming that the "I" in Mill's formulation is himself/herself in possession of a mind.

AI and gender/embodiment

Caleb and Nathan's discussion of AI and gender is easily (for me) the most thought-provoking scene in *Ex Machina*. Before this scene, Ava has dressed for Caleb and says she wants to go on a date with him; he begins to feel something for Ava but, as his questions to Nathan make clear, he still thinks of her as a machine with physical features and programming thought of as "female" and "heterosexual." After this scene, though, he regards her as having, at the very least, a body that is analogous to his (he will later tell her that Nathan will re-format her mind, "which is the same as killing you").

To talk about everything raised in this scene in anything like depth would run to many pages, so here (and my "out" is that, well, *they* don't go into depth on these matters) I'm opting just to identify and briefly comment on the topics raised. A little thinking about any of them should show that just about all of them are interrelated with each other, and that underlying them all is the unspoken but crucial questions of how embodiment plays a role in consciousness, and what that might look like in a robot.

Caleb makes the argument that Nathan could have just made Ava "a gray box" instead of endowing her with a gender and sexuality. Nathan counters in a couple of ways; he begins by asking whether Caleb can give an example of consciousness, "at any level, human or animal, that exists without a sexual dimension?" Caleb replies, "They [meaning "animals;" curiously, he leaves humans out of this statement] have sexuality as an evolutionary reproductive need." Nathan leaves that for the moment and returns to Caleb's "gray box" image to critique it: Would a gray box have any reason to interact with anyone, or with another gray box? Can consciousness exist without interaction?

Nathan implicitly believes that the basic premise of embodiment is a crucial component in consciousness: that we learn about and understand the world (and ourselves) via the sense data we gather through our physical being (Wilson and Foglia). However, Nathan makes far more explicit his belief that gender and sexuality, since they are dependent in large measure on bodies, are unavoidable, indeed inextricable, components of consciousness. For him, then, gendering an AI is a foregone conclusion. Why he (and a sizable majority of roboticists in both the real world and myriad fictional ones) overwhelmingly choose to gender their AIs as female is another question; but while correlation is not causation, the vast majority of people working in this field are men (Penny; Turkle 287-290). Caleb's questions, on the other hand, reflect the idea behind Rene Descartes' famous declaration, "*Cogito, ergo sum*" ("I think; therefore, I am"): the mind receives and interprets

data from the senses but is otherwise separate from the body. Hence his assertion that AIs don't need a gender to be conscious.

In the second part of their discussion, Nathan pushes the notions of gender and embodiment for AIs to (what he regards as) their logical conclusions: "In between [Ava's] legs, there's an opening, with a concentration of sensors. You engage them in the right way, [it] creates a pleasure response. So if you wanted to screw her, mechanically speaking, you could. And she'd enjoy it" (qtd. in "*Ex Machina* Quotes").

Even if you have not had sex, you still know via your body what the words "pleasure" and "enjoy" feel like—you just know those things from other experiences you have had. But what do those terms mean for a robot? **Can** those terms have any meaning at all for a robot? Nathan seems to assume that *all* bodies, whether living or mechanical, can experience pleasure if equipped in the right way (this, by the way, extends to his saying that Kyoko "likes" to dance); but, to be more precise and, yes, sinister about this: because he has designed Ava and Kyoko (and, probably, the earlier AIs as well) in this way, Nathan not only thinks they *can* experience pleasure, he thinks they *should* experience pleasure. He may believe that "sexuality is fun, man," but determining whether the AIs he has built can in any way understand that idea, even on their own terms, will be made difficult by the fact that we have only the experience of being human to draw upon.

"If a lion could speak, we would not understand it."—Ludwig Wittgenstein, *Philosophical Investigations*.

Evidence for/against Ava's sentience

What follows is a mix of things already discussed with one or two things not mentioned earlier, along with comments as appropriate. You may also find other moments that argue for or against her sentience.

Evidence for her sentience:

1) Her drawings: over the course of the film, she changes from not being able to say what they are (or why she draws at all) to drawing with intention.
2) Her joke (throwing back Caleb's line about being interested in seeing what he would choose to tell her): Caleb interprets this as a sign she has an awareness of her own mind and, thus, an awareness of his.
3) Her ability to cause the power cuts: ***IF*** Nathan has not programmed this ability in her, Ava has truly learned how to do this on her own (that is, it is new information of Ava's making), and she can replicate it when she chooses. (See also #3 in "Evidence against her sentience.")
4) She tugs at the sleeves of the sweater she puts on when she first dresses for Caleb: what makes this stick out for me is that a) we get a close-up of her hands as she does this: a signal from the filmmakers that this gesture matters; b) she does this out of Caleb's line of sight; it's not done in front of him to try to trick him in some way. If we saw a human doing this, we would know instantly that they are nervous or anxious—clear emotional states.
5) "Is it strange to have made something that hates you?"

6) Just before the final encounter with Nathan, Ava whispers something in Kyoko's ear and at the same time taps some kind of code on her forearm. While I have no idea what is happening there, it's a safe bet that Nathan did not program Ava to do precisely these things. (Side note: Kyoko already has a knife with her when she comes downstairs, perhaps indicative of her own decision-making independent of programming.)
7) Her decision to leave Caleb behind: She makes a conscious decision based on her real, if limited, experience with men who are also advanced computer programmers (n=2): if Nathan is a jerk, Caleb must be also.

Evidence against her sentience:

1) Her drawings: she never explains *why* she draws, which suggests that she is programmed to do so on a regular basis; moreover, as mentioned earlier, she and Caleb form a kind of GAN—she draws the second picture (the tree in her enclosure) in response to his request that she pick something to draw, and the third one (the portrait of Caleb) as a further enticement to encourage him to help her escape (which Nathan has programmed her to do).
2) Her wanting to visit a traffic intersection on her hypothetical date with Caleb: Nathan tells us that Blue Book is Ava's software. Search engines are hungry for data; Ava's expressed desire here is really no more than a reflection of her software's need for data.
3) Her ability to cause the power cuts: ***IF*** Nathan has programmed this ability in her, this is nothing special, even if she can choose when to cause them. What makes this hard to know is that we have no clear evidence one way or the other of the extent of Nathan's knowledge about the cause of the cuts. The first time the cut happens (Caleb's first night at Nathan's house), Nathan seems evasive when explaining to Caleb what's going on and/but reassures him that everything is fine; later on, though, when Ava triggers another cut, Nathan looks around as if he's startled by it.
4) Her programming: Nathan tells Caleb that Ava was designed to use various means at her disposal to persuade Caleb to help her escape. Among those he lists is "empathy," which seems unlikely: Ava has a high degree of *emotional intelligence*—that is, she has been programmed to respond appropriately with both words and facial expressions when, for example, Caleb tells her that his parents had been killed in the car accident—but as a non-biological robot, one who has not experienced and can never experience the loss of parents or other loved ones, she cannot be truly empathetic. (For a good, clear discussion of the difference between "emotional intelligence" and "emotion," see rdigitalife.)
5) Her decision to leave Caleb behind: Nathan tells Caleb that perhaps Ava is only pretending to like him in order to achieve her goal of escaping. Nothing in her programming obligates her to reward Caleb for doing so.

Works Cited and Consulted

Avramides, Anita. "Other Minds." *Stanford Encyclopedia of Philosophy*. 2 May 2019. plato.stanford.edu/entries/other-minds/. Accessed 23 Mar. 2020.

Brownlee, Jason. "A Gentle Introduction to Generative Adversarial Networks (GANs)." *Machine Learning Mastery*. 19 Jul. 2019. machinelearningmastery.com/what-are-generative-adversarial-networks-gans/. Accessed 23 Mar. 2020.

Cuddon, J. A. *The Penguin Dictionary of Literary Terms and Literary Theory*. 4th ed. Penguin, 1999.

"*Ex Machina* (2014)." *Internet Movie Data Base*. Amazon. n.d. www.imdb.com/title/tt0470752/?ref_=nv_sr_srsg_0. Accessed 19 Mar. 2020.

"*Ex Machina* Quotes." "*Ex Machina* (2014)." *Internet Movie Data Base*. Amazon. n.d. www.imdb.com/title/tt0470752/quotes/?tab=qt&ref_=tt_trv_qu. Accessed 22 Mar. 2020.

Garland, Alex. "Ex Machina" [clean shooting script]. *The Daily Script*. n.d. www.dailyscript.com/scripts/exMachina_script.pdf. Accessed 19 Mar. 2020.

Garland, Alex, writer and director. *Ex Machina*. Lionsgate, 2015.

Hamilton, Edith. *Mythology*. Mentor, 1952.

Lessons from the Screenplay. "*Ex Machina*: The Control of Information." *Lessons from the Screenplay*. YouTube. 20 Dec. 2016. www.youtube.com/watch?v=1Ko9mWdqW-M. Accessed 19 Mar. 2020.

Nietzsche, Friedrich. "Parable of the Madman." *Modern History Sourcebook*. Fordham University. Aug. 1997. sourcebooks.fordham.edu/mod/nietzsche-madman.asp. Accessed 26 Feb. 2020.

Penny, Laurie. "Why do we give robots female names? Because we don't want to consider their feelings." *New Statesman*. 22 Apr. 2016. www.newstatesman.com/politics/feminism/2016/04/why-do-we-give-robots-female-names-because-we-dont-want-consider-their. Accessed 20 Mar. 2020.

rdigitalife. "Rosalind Picard on Robot Evolution." *rdigitalife*. YouTube. www.youtube.com/watch?v=n32KaMMJuI8&t=15s. Accessed 23 Mar. 2020.

"ShrinkWrapped." "Empathy and AI: Part IV." *ShrinkWrapped*. 4 Feb. 2010. shrinkwrapped.blogs.com/blog/2010/02/empathy-and-ai-part-iv.html. Accessed 20 Mar. 2020.

Searle, John. "Minds, Brains, and Programs [1980]." *Cogprints*. 14 Aug. 2003. cogprints.org/7150/1/10.1.1.83.5248.pdf. Accessed 20 Mar. 2020.

Singler, Beth. "Controlling AI, Controlling Fictions." *Dr. Beth Singler*. 23 Nov. 2019. bvlsingler.com/2019/11/23/controlling-ai-controlling-fictions/. Accessed 17 Jan. 2020.

Temperton, James. "'Now I am become Death, the destroyer of worlds'. The story of Oppenheimer's infamous quote." *Wired UK*. 9 Aug. 2017. www.wired.co.uk/article/manhattan-project-robert-oppenheimer. Accessed 22 Mar. 2020.

"Through the Looking Glass: Creating *Ex Machina*." *Ex Machina*. Lionsgate, 2015.

Turkle, Sherry. "The Robotic Moment." *Up for Discussion: Readings, Rhetoric, and Research*. Edited. by Andrea McCaffree-Wallace et al. Butler Community College English Department, 2019, pp. 285-296.

Wilson, Robert A. and Lucia Foglia. "Embodied Cognition." *Stanford Encyclopedia of Philosophy*. 8 Dec. 2015. plato.stanford.edu/entries/embodied-cognition/. Accessed 23 Mar. 2020.