Science Fiction and the Boundaries of Philosophy:

Exploring the Neutral Zone with Plato, Kant, and H.G. Wells

Andrew Fiala

California State University, Fresno

Abstract

In this paper, I consider the difficulty of distinguishing between science fiction and philosophy. The boundary between these genres is somewhat vague. There is a "neutral zone" separating the genres. But this neutral zone is often transgressed. One key distinction considered here is that between entertainment and edification. Another crucial element is found in the importance of the author's apparent self-consciousness of these distinctions. Philosophy seeks to edify, and philosophers are often deliberately focused on thinking about the question of the borders that distinguish genres. Science fiction is more interested in entertainment, and narrative authors tend to care less about policing the border. This distinction is a pragmatic one. And canonical authors often violate the boundary. I examine key authors to make this point, including Plato, Kant, and H.G. Wells. My discussion shows that these canonical authors wandered into the neutral zone in their own work, reminding us that the boundaries we draw around genres are arbitrary and subject to transgression.

In his new book, *Star Trek's Philosophy of Peace and Justice*, Joseph Orosco argues that science fiction is a useful tool for imagining a progressive future. He connects science fiction and fantasy to a much longer tradition of what he calls "moral fables," which he extends back to Plato's allegory of the cave. Philosophers have often made use of imaginative fictions. In addition to Plato, Orosco cites Nozick's "experience machine," Rawls's "original position," and Thomson's "trolley problem." Orosco's book is part of a growing area of interest in professional philosophy, which has given birth to a number of books focused on philosophy and popular culture, including volumes focused on the science fiction of *Dr. Who*, *The Matrix*, *Inception*, *Star Wars*, etc. There is even an online journal focused on "Sci Phi" and a new *Journal of Science Fiction and Philosophy*. In an issue of *Sci Phi Journal*, David Kyle Johnson explains that allegories and fictions have often been employed by philosophers to open up new vistas of thought (Johnson, no date).

And yet, philosophers remain somewhat uncomfortable with science fiction. In the inaugural issue of the *Journal of Science Fiction and Philosophy*, Alfredo Mac Laughlin, the founding editor of the journal, attributes this to the feeling that science fiction is not "serious literature" (Mac Laughlin 2018). There are other concerns. Perhaps we are pandering to our students or to the mob, when we (academic philosophers) make overt attempts to make connections between academic philosophy and pop culture trying to be relevant and, well… popular.[1] We might also worry that pop culture encounters can be preoccupied with superficial and trivial matters that seem far removed from ethical, epistemological, or metaphysical inquiry. The general worry is, I suppose, that we can end up allowing our philosophical focus to be corrupted by Hollywood and the publishing industry. We can also be distracted by the bells and whistles of the stories and lose track of the philosophical questions. When we get too focused on pop culture, literature, and film, we end up doing something that is not properly philosophical.

Of course, in thinking about the value of science fiction, philosophy, and popular culture in general, we find ourselves confronting the question of what counts as philosophy. Such a question is answered by surveying the boundaries of philosophy and looking for features that distinguish philosophy from its other. Or, to use a science fiction metaphor, we find ourselves surveying the neutral zone. The neutral zone, in Star Trek, is the border region separating rival powers. Of course, such a boundary is somewhat arbitrary and not easy to police. This problem occurs repeatedly in Star Trek. There are areas of influence but there are also incursions and relations across the neutral zone. And a boundary, once drawn, can be redrawn—as the neutral zone was redrawn, bringing the Klingons into the Federation, while still excluding the Romulans. The science fiction example may seem a bit silly. But the point remains true in politics, the arts, and the life of the mind: we draw borders and make distinctions for various purposes; but these borders are also transgressed and occasionally redrawn. This is also a problem in attempting to define and police the borders that separate philosophy from literature and from science fiction. It turns out that it is not as easy to define these boundaries (and defend them) as we might think. My thesis here is that when we ask about the difference between philosophy and fiction (including science fiction), we end up questioning how we understand philosophy. And we are in fact doing philosophy when we ask that kind of question. Philosophy occurs when we ask the question, "But is it philosophical?"

Part 1: Defining Philosophy, with a Glance at Plato

So, what exactly is philosophy? We might begin with a stipulative definition of philosophy that simply rules out fiction. We might define philosophy as a set of rational arguments that avoids fictional scenarios, counterfactual speculation, utopian hope, poetic rhetoric, and other excesses of emotional and non-rational language. According to this paradigm, philosophy ought to occur as a kind of systematic treatise such as we find, perhaps, in Thomas Aquinas's *Summa Theologica*, Hegel's *Encyclopedia of the Philosophical Sciences*, or Wittgenstein's *Tractatus Logico-Philosophicus*.

This narrow definition of philosophy does not stand up to historical scrutiny. Aquinas had a profound appreciation for theological poetry (see Murray 2013). Hegel thought art and poetry were manifestations of absolute spirit, along with religion and philosophy. And Wittgenstein suggested in a journal entry in 1933 (that appears in *Culture and Value*) that "philosophy ought to really be poetic" (Wittgenstein 1980, 23; my translation). We might also consider the poetic, fictitious, and dramatic work of authors such as Kierkegaard, Nietzsche, Camus, Sartre, and De Beauvoir. Behind all of that is Plato, whose work is poetic, dialogical, and rhetorically complex. That Plato was concerned with poetry is not surprising. Diogenes Laertius tells us that Plato was originally engaged in writing poetry and drama. But, so the story goes, when Plato was preparing to offer a tragedy at the theater of Dionysus, he happened to hear Socrates speaking, "after which he burned his poems" and became a disciple of Socrates (Laertius, 2018, 99). Plato offered philosophical accounts of poetry, rhetoric, tragedy, and comedy in works such as *Phaedrus* and *Symposium*. Plato's *Republic* contains much that is allegorical, speculative, and even poetic. In addition to the Allegory of the Cave, there is the tale of the Ring of Gyges. We might also cite the beginning and end of *Republic* as especially important. *Republic* begins with a narrative that takes Socrates down to the Piraeus and eventually puts him into confrontation with Thrasymachus. And Book 10 contains the "Myth of Er," which is a speculative and poetic tale. Of course, Plato warns against poetry and prose within *Republic* itself, in Book 3, where Socrates says the following:

> What poets and prose-writers tell us about the most important matters concerning human beings is bad... I think we'll prohibit these stories and order the poets to compose the opposite kind of poetry and tell the opposite kind of tales (*Republic*, 392b).

Plato offers us an interpretive strategy here: the poetry and prose that the philosopher-king thinks of as "bad" will be banished; but other kinds of poetry and prose will be allowed—so long as they are "not bad." Socrates suggests that the bad kinds of stories encourage viciousness, while good stories encourage virtue and wisdom. But—and here is the point—the judgment about what counts as good and bad stories must somehow be explained, defended, and accounted for. And as Plato shows us in the structure of *Republic* itself, we probably need poetry and prose, as well as dialectical philosophy, in order to make that judgment. Plato does not come right out and say this. But we see it cleverly enacted in the narrative and poetic elements that occur at the beginning and end of *Republic*. Plato seems to tell us that dialectic is insufficient; philosophy also needs narrative and poetry.

The history of philosophy is full of texts that transgress the supposed boundaries of philosophy and make incursions into the neutral zone. There are many canonical texts that contain poetry, prose, fantasy, and indeed science fiction. Parmenides takes us on a poetic journey. Theological speculation is an exercise in poetic creativity, with Augustine explaining the cosmos in terms of an extended metaphor of two cities (the city of God and the city of man). Social and political philosophy often includes imagined utopias (from Plato's *Republic* itself through Marx's communism and the laissez-faire idealism of libertarian capitalism). Even the supposed "rationalism" of a philosopher such as Descartes

includes the strange fiction of the evil deceiver who plays such a crucial role in the *cogito* argument in *Meditations*. The state of nature of Hobbes, Locke, and Rousseau is a creative fiction. And so on.

So, one easy point to make is the following: there is no firm or clear distinction between philosophy and other forms of writing and speaking. And yet, we can make some pragmatic distinctions that help us draw a boundary. This "helpful pragmatic distinction" does not aim to describe a Platonic paradigm of either philosophy, fiction, or poetry. Rather, this is a useful set of generalizations that can guide our thinking, as we wander in the middle of the continuum that includes philosophical fiction, as well as philosophical writing that employs narrative and fiction. Two useful generalizations are the following:

- Fiction, narrative prose, poetry, and drama, including science fiction are often focused on descriptive details, moral exhortation, and speculative idealism that is offered in the context of a concrete narrative and from the point of view of a character/personality. This may include arguments made either by the narrator or by specific characters. But the point in a work of fiction is not to argue from a universal/objective point of view. Rather, if there is an argument, it is presented from a certain perspective and vantage point and in connection with a specific narrative moment. And in addition to offering edification, speculation and exhortation, an important goal of the aesthetic form of communication is to entertain.

- Philosophical activity involves a much larger degree of universality and argumentation that aims for a kind of universality that is free of any concrete perspective (i.e., what Nagel called "the view from nowhere"). It considers and replies to objections—and is self-conscious and self-critical of the limits of what it is able to prove. The project often includes voices and points of view (as in a Platonic dialogue or a treatise that considers objections). This shows us that to attain "the view from nowhere" we must include elements that are located "somewhere." Indeed, philosophical communication ought to include some level of awareness of this fundamental problem, which is that the "view from nowhere" is itself a kind of useful fiction. And while philosophy can be provocative and exhortative, it tends to shy away from being merely entertaining, intending instead to be edifying.

With this in mind, we can see a difference between a text like Plato's *Republic* and another text like Sophocles' *Oedipus Rex*. In *Republic*, Plato provides arguments. There are characters and personalities in the text (e.g., Thrasymachus, Glaucon, and Er). But the point is not merely to represent characters engaged in a narrative struggle. Rather, the philosophical text is focused on the arguments made by these characters—aiming at an overall evaluation and conclusion of these arguments that ought to lead to the reader's edification. By "edification" I mean an effort that aims at improvement through knowledge, the production of wisdom, and a movement toward enlightenment. In more overtly

monological treatises, the process of argument and evaluation is provided by the voice of the author, who appears to speak from a kind of "nowhere" vantage point. Of course, we see argument and dialogue in Sophocles and in other works of narrative fiction. The Chorus provides some space for reflection and critique in Sophocles, as do other characters and narrators in other works of drama and fiction. And the characters in Shakespeare or Dostoevsky engage in dialogue and make arguments.

But—and here is an important point—Plato informs us in the text itself that there is a difference between philosophical argument and poetic play. He lets us know that he understands that difference—and that there are philosophical questions that arise even in making this distinction. This kind of critical self-conscious methodology can also be found in many paradigmatic works of "philosophy," which can result in the kind of circle of critique and justification that is found in Kant, Hegel, and Wittgenstein. Some speculative poetry, narrative prose, and dramatic works also include this kind of self-consciousness— especially in some "modernist" poetry and prose. So, the distinction is not hard and fast— and there is an overlap between philosophical fiction and works of philosophy that contain elements of narrative, fiction, and poetry. But again, the point is that the philosophical text is typically more explicitly self-conscious and less caught up in narrative details and the need to entertain.

Part 2: Two Paradigmatic Examples of Philosophy and Science Fiction

Now let's move directly toward a consideration of science fiction. For our purposes here, science fiction is narrative writing that includes scientific and technological elements that create circumstances that are essential to advancing the narrative. We'll consider this working definition more carefully later in the paper. But let's begin by noting that philosophical writing often makes use of science fictional elements; and literary/cinematic science fiction occasionally becomes explicitly philosophical. But the difference (again a "pragmatic" distinction) is found in the degree of self-consciousness of the author—and the degree to which the author's point is to edify or to entertain. One form of this distinction depends upon whether the text articulates a self-consciously philosophical argument or whether it merely offers a descriptive account of an alternative reality. Another consideration is whether the text provides moral exhortation by way of parables and cautionary tales, or whether it simply intends to amuse.

Again, these distinctions are limited and pragmatic. The neutral zone is often transgressed and re-established. And the boundaries between the genres are often blurred. To make this point about the blurring of the boundaries, consider the following two paradigmatic examples. The first is from the work of Immanuel Kant and the second is from H.G. Wells. These are among the most influential canonical figures in each genre. If we can show that Kant and Wells strayed back and forth across the neutral zone, this provides us with some evidence that the boundary is blurry and that it has been transgressed by key figures in the canon of each genre.

1. **Kant and the Extra-Terrestrials**. In the 1780s and 90s, Kant speculated about the philosophical need to consider extra-terrestrial life. In the *Critique of Pure Reason*, Kant wrote, "If it were possible to settle by any sort of experience whether there are inhabitants of at least some of the planets that we see, I might well bet everything that I have on it. Hence, I say that it is not merely an opinion but a strong belief (on the correctness of which I would wager many advantages in life) that there are also inhabitants of other worlds" (*Critique of Pure Reason*, B 853). In his *Anthropology* lectures, he stated that in order to fully understand humanity, which he defines as a "terrestrial rational being," we would need to compare humanity to some "non-terrestrial rational being" (Kant 2016, 225; see Szendy 2013).

   Kant is not the first philosopher to consider extra-terrestrial life. The ancient Epicureans had speculated that there was a plurality of worlds and that these worlds might be populated by creatures such as we find here on Earth (Epicurus in Diogenes Laertius, 2018). This view was rejected by Plato and Aristotle who insisted that our world was the only world—as in Plato's *Timaeus* (see Crowe 1999). And closer to Kant, Leibniz had also speculated about extra-terrestrial life. In his *New Essays Concerning Human Understanding* (written in 1704), Leibniz imagined what would happen if we met a rational being from the moon. Leibniz reflects on this as a thought experiment intended to clarify that the concept of "rational being" does not necessarily depend upon being human or terrestrial. And—what is significant for our purposes here—Leibniz refers to a book by Dominique Gonsales in making his point. "Gonsales" is a pseudonym used by the English author Francis Godwin. The book referred to is considered to be among the first examples of science fiction, *L'homme dans la Lune* (or *The Man in the Moon*), published in 1638 (see Roberts 2016). It features a flying machine that takes the narrator to the moon.

   By the time Kant was writing, the kind of speculation found in Leibniz was already in the air. Kant had speculated about extra-terrestrial life since he was a young man, writing about this in one of his first essays (from 1755), "Universal Natural History and Theory of the Heavens" (Kant 2012). In that essay, Kant speculates about the passage of time from the point of view of "more perfect creatures" who may exist on Jupiter (Kant 2012, 302). This idea foreshadows the question of the phenomenology of time that Kant would consider in more detail in *The Critique of Pure Reason*. He also engages in ethical speculation, wondering whether "sin exercises its domination in other spheres of the solar system as well or whether virtue alone holds there" (306). Kant imagines that the inhabitants of distant celestial bodies are "too noble and too wise to lower themselves to the foolishness that resides in sin" on the "lower planets" (306). He pines for the opportunity to leave the earth, imagining that by the time human beings are capable of doing this, we will have perfected our own natures. He says, "Who knows, perhaps the satellites orbiting around Jupiter will light our way to the future?" (307). He concludes by suggesting that when we speculate about the heavens in this way, "the view of the starry sky on a clear night gives one a kind of pleasure that only noble souls feel" (307). This

foreshadows the famous phrase associated with Kant (and which is found as the epitaph on his tombstone). This phrase is from the conclusion of *The Critique of Practical Reason*: "Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily we reflect upon them: the starry heavens above me and the moral law within me." Kant may have had in mind here not only the mechanical motion of the starry heavens but also the kind of edification that occurs when we speculate about extra-terrestrial life.

Of course, Kant did not write fiction. His goal is edification, not entertainment. The method of critical philosophy clarified the limits of thought from the vantage point of "the tribunal of reason." But the idea of the tribunal of reason is itself a metaphor (see Fiala 2002). To be clear, this metaphor and Kant's speculation about alien life are supposed to enlighten. Kant is not engaged in constructing a world of fantasy for the sake of amusement. In his speculations about extra-terrestrial life, Kant did not construct elaborate details about what extra-terrestrial life would look like, how we would make contact with extraterrestrials, or what the aftermath might be. Those endeavors were taken up by authors writing in another genre.

2. **Wells and the Martians**. And so, one hundred years later, we come to H.G. Wells, an author whom many consider to be the father of science fiction—despite the earlier work of authors such as Godwin (Gonsales), mentioned above. We might pause for a moment to flesh out an important detail about Godwin: he was a bishop of the Church of England who told the story of the Man in the Moon in order to explore themes in Christianity. Wells also used his fiction to explore philosophical themes. In his fictional works, Wells speculates about anthropology, ethics, metaphysics, and consciousness. In *The Time Machine*, for example, from 1895, Wells returns us to a problem as old as Parmenides: whether time and change are real. The opening chapter of *The Time Machine* offers a brief account of time as the fourth dimension in connection with the possibility of traveling along that fourth dimension. But the point of Wells's book is not to prove that this is possible or to consider the kinds of paradoxes and problems in the philosophy of time that have obsessed philosophers from Aristotle to Heidegger and McTaggart. Rather, Wells uses time travel as a convention that allows him to entertain us, while also offering speculation about anthropology and morality. Wells was himself a socialist and a pacifist who used his fiction to describe the absurdity of war and the need for social reform. This becomes obvious in his later fiction, such as *The World Set Free* (1914)—republished in 2022. Written at the outset of the First World War, Wells imagines the invention of atomic bombs that will destroy civilization, and usher in a new era of human development that abolishes war and unites humankind under a world government. The narrative in that book is not very compelling, since the book is often overtly didactic. It uses the fictional setting as a means for Wells to expound (through the voice of several characters) upon his vision of a better world.

Wells seems to have understood himself explicitly in relation to philosophy. This is very clear in the 1898 book, *The War of the Worlds* (Wells 2018). The narrator of this text describes himself as a "speculative philosopher" (Wells 2018, 171) who was writing a paper on "moral philosophy" (170) when the Martians invaded. Indeed, we might interpret the text itself as the very essay on moral philosophy that the narrator was writing. The text begins with a very general philosophical account of the place of humanity in the cosmos that is articulated along with a critique of human vanity and the merciless manner in which "humanity" conquered "inferior animals" and "inferior races" (6-7). Toward the end of the text (170), the narrator reveals some words from the essay he had been writing, quoting himself: "In about two hundred years, I had written, we may expect—"

The narrator's speculative text is interrupted by the appearance of the Martians. And so, on a careful reading of Wells's book, the author gives us a hint that the whole enterprise is really a thought experiment, an allegory used to make a point in moral philosophy. There is a kind of self-consciousness expressed here that points in the direction of the philosophical. But rather than unpacking a philosophical argument, Wells is mostly engaged in concretely imagining what an actual encounter with extra-terrestrials would be like. He seems to think that this might teach us important lessons about humanity and morality. But the lessons are part of an entertaining work of art. From Wells's vantage point, the lessons are not hopeful. The human beings in the story behave badly. They panic and go wild. The narrator himself bashes in the head of a priest in order to save himself. And in the end, humanity is not saved by human ingenuity or military prowess. Rather, the human race is saved by natural selection. It is the bacteria that kill the Martians, whose immune systems did not evolve on Earth.

With these examples in mind, let me return to my point about the blurring of the boundary between science fiction and philosophy. These two famous examples show us (1) that philosophical speculation (in the case of Kant) wanders into the realm of science fiction and (2) that science fiction (in the case of Wells) contains an interest in philosophical speculation, anthropology, and moral philosophy—and also includes a degree of philosophical self-consciousness. These examples show us that there is a continuum between philosophical fiction and more properly philosophical prose that wanders into the realm of fiction.

Part 3: Estrangement, Wonder, and the Utopian Imagination

Let's turn now to a consideration of a significant point of overlap between science fiction and philosophy. Joseph Orosco, in his book *Star Trek 's Philosophy of Peace and Justice*, provides us with an explanation of the link between science fiction and philosophy that focuses on what he calls "cognitive estrangement." He explains that cognitive estrangement occurs when we feel out of place, alienated, and at a distance from self and world. He states:

By encouraging cognitive estrangement, science fiction encourages us to reflect about what is actually real, valuable, and important in comparison to the speculative world. Understood in this way, science fiction shares intellectual goals with philosophy, in that both aim to inspire critical self-examination that can hopefully lead to better lives and relationships (Orosco 2022, 13).

This statement gives us an account that includes both means and ends. Orosco suggests that the end or goal of philosophy and of science fiction is reflection on what is real, valuable, and important. This encourages "critical self-examination," which can then lead us to live better. Let's underline this point here: on Orosco's account there is a kind of hope and admonition found in both science fiction and in philosophy. The point is to improve things and help us live better lives. The work of cognitive estrangement is not speculation for its own sake. Rather, it has a progressive agenda. Orosco makes it clear that a progressive view of social justice is contained in the original Star Trek series and its offshoots. Orosco has also emphasized the progressive view of social justice found in Ursula K. LeGuin and other sci-fi authors. This progressive idealism is fairly easy to find in many works in the sci-fi canon, most notably in the work of H.G. Wells himself. Even the dystopian and trippy sci-fi of authors such as Huxley, Vonnegut, P.K. Dick, and others seems to contain a progressive agenda hidden behind the dystopia. In the case of Huxley, for example, the utopian vision is revealed in the difference between the overt dystopian vision of *Brave New World* (1958) and the more optimistic (but also sadly dystopian) book *Island* (1962). *Brave New World* warns against authoritarianism, while *Island* attempts to flesh out a more affirmative vision of an ideal world of love and happiness, while warning that even this ideal will fail.

Orosco explains that the means by which this progressive aspiration is supposed to occur is through "cognitive estrangement." In other words, cognitive estrangement is an instrument for inspiring enlightenment and the work of progress. But this is not a recipe or a concrete agenda for social change. Rather, in the experience of cognitive estrangement, we become conscious of our ideals, and their limitations. And as self-consciousness develops there is progress. But this is not the overt progress of political activism. Rather, it is, as Orosco suggests, part of a process by which self-consciousness can "hopefully lead to better lives and relationships" (as quoted above).

I think we might connect this to a traditional philosophical term, what Plato called "wonder." A familiar account of wonder occurs in Plato's *Theaetetus*. You'll recall that this passage occurs as Socrates is discussing paradoxes with Theaetetus. At one point (*Theaetetus* 155c), Theaetetus responds:

I often wonder like mad what these things can mean; sometimes when I'm looking at them, I begin to feel quite giddy.

The word that is translated as "giddy" here is "*scotodinio*" a word whose etymology connects it to darkness. This is the same term (*scotodinian*) employed by Plato in Book 10 of *Laws*, where the dialogue is discussing the nature of the gods and the soul. The Athenian warns in *Laws* that his partners in dialogue may grow dizzy in thinking about questions

they are not used to discussing. This is similar, it seems, to the feeling of dizziness and vertigo that comes from being in the dark, i.e., the feeling of not knowing your way around. There seems to be some parallel here with what occurs in Plato's cave, when a prisoner escapes and is blinded by the light.

This dizziness and blindness are linked to wonder. For it is from being lost in wonder that the dizziness occurs. In discussing this in *Theaetetus*, Plato provides a playful genealogy. He tells us that the Greek goddess of color and the rainbow—her name is Iris—is the daughter of Thaumas, the god of wonder, known for his link with the "wonders of the sea." And here Plato famously writes (*Theaetetus* 155d):

Wonder is the feeling of a philosopher, and philosophy begins in wonder.

This creative genealogy is a reminder that poetry and metaphor are central to understanding philosophy, which we might claim is itself a child of the rainbow (see Fiala 2020). There is a mystery contained in colored light dancing for a moment upon a mist that hovers somewhere in the emptiness. The rainbow is a metaphor for the transitory experience of appearances. We make meaning and see colors in a physical world. These meaningful appearances hover for a moment and then disappear. Such is life: it is a rainbow dancing in the ether. But the fact that it exists at all—that beauty and love and justice and thought exist—prompts further thinking. From the rainbow to vertigo to wonder. But we should also note that wonder is not the endpoint of philosophy. Rather, wonder (dizziness, vertigo, and blindness) are the first steps on the path out of the cave. The point is not to stop dumbstruck by the wonder of things—but to begin asking questions and to climb higher.

We may seem to have lost the thread here, in this aesthetic account of the rainbow and phenomenology. But that's the nature of philosophy and the process of cognitive estrangement: one question leads to another, as we wonder. In philosophy we are confronted with questions and more questions. And in asking those questions, we are in fact doing philosophy. To wonder about wonder and about the phenomenal appearances of things is something that occurs in poetry and in art. But wonder becomes philosophy when we keep asking questions—about the nature of wonder, about the value of art, and about the process of representation that occurs in poetry. We see this in Plato's genealogy of wonder, which is a kind of fiction and poetry that stimulates further thought. This kind of playful reflection is found throughout Plato's writings.

All of this is part of the kind of "cognitive estrangement" that Orosco discusses. Star Trek can cause us to wonder about things—and to ask deeper questions. When Kant imagines what time would be like for intelligent beings living on Jupiter, something similar happens. When Wells takes us on a time machine into a "futurity" (as he puts it) in which Morlocks prey on the Eloi, we encounter cognitive estrangement. This is also what happens when Descartes conjures up his evil deceiver and when Hobbes and Locke imagine the state of nature. And so on. From the aesthetic moment of cognitive estrangement, questions arise. And in pursuing those questions, we find ourselves doing philosophy.

Orosco adopts the concept of cognitive estrangement from Darko Suvin, the literary theorist who, in the 1980s, explains science fiction as involving cognitive estrangement. Suvin begins his 1979 book, *Metamorphoses of Science Fiction*, with an epigraph taken from Raymond Ruyer:

> *Les chose pouraient être autrement.*

> Things could be different.

Ruyer himself connected this phrase with utopian speculation, myth-making, and religious dreaming. As Ruyer explains, even the dream of a "Golden Age" is a way of saying "things could be different" (Ruyer, 1950, 4-5). Ruyer shows that this is common in all utopian works. He includes as an example, Plato's account of Atlantis in *Critias*, as well as other authors including Rousseau, Huxley, Tolstoy, and even Gandhi.

Of course, utopian speculation imagines that things could also be better—not just different but better. And so, we should also note that there are ways of imagining a world that is not better at all. Let's note, in passing, that dystopian speculation also says, "things could be different," not as a way of offering hope but as a warning against decay and dissolution. Here we find the science fiction of Huxley, Orwell, Vonnegut, and others—even including Wells. But I think Orosco is correct in suggesting that behind these warnings is usually a progressive social agenda. Even in Star Trek, there are both utopian and dystopian elements. As Orosco tells us, in the Star Trek universe, the 21st Century is a terrible time. The utopian future imagined for the Federation emerges only after a dystopian 21st Century.

And now let's return to Suvin and his account of cognitive estrangement, which he suggests as a defining feature of science fiction. Suvin argues that science fiction is "the literature of cognitive estrangement." He continues: "This definition seems to possess the unique advantage of rendering Justice to a literary tradition which is coherent through the ages and within itself, yet distinct from nonfictional utopianism, from naturalistic literature, and from other non-naturalistic fiction" (Suvin 1979, 4). Suvin notes that there is a common thread throughout all of literature in the experience of estrangement. He suggests that we see estrangement in the epic of Gilgamesh and even in the Hebrew Bible's account of Eden. It is found in Jonathan Swift's *Gulliver's Travels* and in the *Utopia* of Thomas More. But for something to count as *science* fiction, the cognitive element is key. By "cognitive element," Suvin appears to mean a rational account of how and why the estrangement occurs. Science fiction does not merely make up strange worlds and introduce gods and heroes. Rather, it offers a logical, rational, and technological account of how these worlds are discovered, how they operate, and what the humans (or other protagonists) do within those worlds. In other words, there must be some cognitive plausibility to the estrangement found in science fiction. Said bluntly, science fiction is *scientific* and not merely magical or fantastic.

This kind of distinction between science fiction and fantasy is noted by sci-fi authors themselves, who often find themselves confronting the neutral zone and feeling the need to reflect on the question of genres and classification. H.G. Wells described his own books as "scientific romances." He even suggested that his own work was more properly understood as "fantasy." Wells contrasted his work with that of Jules Verne, suggesting that Verne was more interested in thinking about a world of scientific possibilities. Verne's books, *Journey to the Center of the Earth* (1864), *Twenty Thousand Leagues Under the Seas* (1870), and *Around the World in Eighty Days* (1872) imagined technological innovations that would lead human beings to undertake new explorations and adventures. Wells's books such as *The Time Machine* (1895), *The Island of Doctor Moreau* (1896), *The Invisible Man* (1897), and *The War of the Worlds* (1898) were less focused on the details of the technological possibilities than in imagining alternative realities that could disclose moral and psychological problems. Wells explains that his works and others like them—he names *Frankenstein*, for example—"are all fantasies; they do not aim to project a serious possibility; they aim indeed only at the same amount of conviction as one gets in a good gripping dream" (Wells 2017, 13). Wells explains that the writer of what he calls "scientific romances" engages in a conjuring trick that encourages the reader to imagine some alternative reality. But the point for Wells is not the actual possibility of that alternative reality; rather, it is to inhabit that alternative in order to encourage reflection on the self, the world, and on ethics and politics. And yet, what helps us distinguish science fiction from other kinds of fantasy is some "realistic" (i.e., scientifically plausible) account of how and why the human beings who inhabit these alternative realities got there. Robert Heinlein—another sci-fi author—includes this as the most important aspect of his five-point definition of science fiction (Heinlein 2017, 19-20).

1. Science fiction must involve a "different" reality from the here-and-now
2. That new reality must be important to the story
3. The plot must focus on a human problem (and not merely be a technical account)
4. The human problem must be created by the "different reality"
5. And: the story must explain how that different reality occurred in some scientific or technologically plausible fashion.

The last point distinguishes science fiction from stories that focus on magic, wizards, superheroes, and gods. There must be some plausible, technological account underlying the alternative world imagined by science fiction. This point can help us eliminate from a "sci-fi canon" texts such as the Epic of Gilgamesh, Plato's Myth of Er, Cicero's Dream of Scipio, and the Bible's account of Eden. Those texts are magic, mystical, and supernatural. But they are not grounded in a plausible scientific explanation of how one gets to the alternative reality.

But in objection to the claim that the Epic of Gilgamesh, et al., are not works of science fiction, let me suggest that the question of what counts as "realistic" or "scientifically plausible" is itself in question here. Novels about rocket travel and artificial intelligence seem "scientifically plausible" to us today. But to the ancients these would have seemed like magic. And in the ancient world, the idea of soul-travel—either in dreams or in the afterlife—may have been more plausible than it is in our world.

Conclusion: Philosophy at the Neutral Zone

In making this last point we return again to philosophical questions about what we mean by plausible, realistic, and so on. These are deep philosophical questions. And it is in philosophy that we really delve into questions about what is realistic, magical, fantastic, and so on. These questions are theoretical, epistemological, and metaphysical. One way of putting this is to say that, as non-fiction, philosophy asks questions about what is plausible, real, or fantastic—while science fiction (and fantasy and fiction in general) simply assume something about the real, the plausible, and the fantastic. In other words, philosophy asks the question about what counts as reality, while science fiction assumes a standpoint and tells a narrative within an assumed framework of the real. This helps us firm up part of a distinction between edification and entertainment. Edifying discourse (e.g., philosophy) is supposed to be in some sense true: it ought to disclose something that is real. But entertaining works of fiction need not be based on any account of reality or truth: as fiction, they make no claims about truth and reality. But of course, as philosophers know, there are even disputes within philosophy about the meaning of truth and reality, fact and fiction. There are "facts" contained within fiction (for example, it is a "fact" of *The War of the Worlds* that the Martians invaded the countryside around London—and not Colorado— even though there are "in fact" in the real world no Martians). And fiction, poetry, and drama can disclose truths about ethics (such as that we ought to "pity those witless souls that suffer our domination," as Wells puts in *The War of the Worlds*, p. 145). As Picasso says, in a quote that has been widely warped and misattributed: "Art is a lie that makes us realize truth."[2] And those who have considered the philosophy of fiction—from Plato to Searle—remind us that the distinction between truth and fiction is not as firm as we might believe (see Kroon and Voltolini 2019).

And so, allow me to return to the worry with which I began—that in discussing literature, poetry, and popular culture we can end up wandering into the neutral zone and straying from philosophy itself. The solution to this concern involves remaining committed to a certain kind of questioning and interpretive work. Let us return to H.G. Wells to make this point. In discussing *The War of the Worlds*, we could get sidetracked into questions about how long the Martian rockets took to get to the Earth, why they chose London as their first point of assault, and why the Martians were so clueless about the terrestrial bacteria that ultimately killed them. Those questions are asked within the framework of the world created by Wells. A kind of "philosophical questioning" could further be applied to the text. We could ask about the analogy between the Martians and European colonialism or about the implicit critique of meat-eating and our treatment of the lower animals found in the text. Or we could wonder whether human nature really is as pathetic as Wells makes it out to be. In those questions, we have moved beyond the text to general questions about morality and civilization. Beyond that, we encounter even more philosophical questions, such as the question of whether literature itself is a useful method for encouraging us to ask those questions. And beyond that we encounter a further question—about why we ask these kinds of questions, what kinds of answers we are willing to accept, and whether in answering these questions we are actually thinking philosophically. At some point on this

journey of questioning, we may feel a kind of cognitive estrangement. This may manifest as a kind of dizziness and vertigo that is akin to wonder. And, as we know, when we get to that level of cognitive estrangement, we are well on our way to climbing out of the cave (as Plato would put it) or kicking away the ladder (a la Wittgenstein). But notice here in closing, that we have returned to metaphors in order to make this point—and find ourselves again hovering in the neutral zone.

✦

Works Cited

Crowe, Michael J. 1999. *The Extraterrestrial Life Debate, 1750-1900*. New York: Dover.

Diogenes Laertius. 2018. *Lives of the Eminent Philosophers*. Translated by Pamela Mensch. Oxford: Oxford University Press.

Epicurus. 2018. "Letter to Herodotus." In Diogenes Laertius, *Lives of the Eminent Philosophers*. Translated by Pamela Mensch. Oxford: Oxford University Press, Book 10.

Fiala, Andrew. 2020. *Simplicity*. Fresno, CA: Palm and Sierra.

Fiala, Andrew. 2002. *The Philosopher's Voice: Philosophy, Politics, and Language in the 19th Century*. Albany, NY: State University of New York Press.

Heinlein, Robert A. 2017. "On the Writing of Speculative Fiction." In *Science Fiction Criticism: An Anthology of Essential Writings*. Edited by Rob Latham. London: Bloomsbury.

Johnson, David Kyle. No date. "What Sci Phi Is All About." *Sci Phi Journal.* URL = < https://www.sciphijournal.org/?s=what+sci+phi+is+all+about >.

Kant, Immanuel. 2012. "Universal Natural History and Theory of the Heavens." In *Kant: Natural Science*. Translated by Beck et al. Cambridge: Cambridge University Press.

Kant, Immanuel. 2016. *Anthropology from a Pragmatic Point of View*. Translated by Robert Louden. Cambridge: Cambridge University Press.

Kant, Immanuel. 1998. *Critique of Pure Reason*. Translated by Paul Guyer. Cambridge: Cambridge University Press.

Kroon, Fred and Alberto Voltolini. 2019. "Fiction." In *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition). Edited by Edward N. Zalta. URL = < https://plato.stanford.edu/archives/win2019/entries/fiction/ >.

Mac Laughlin, Alfredo. 2018. "The What and Why of this Journal." *Journal of Science Fiction and Philosophy*, vol.1 (2018).

Murray, Paul. 2013. "Aquinas on Poetry and Theology." *Logos: A Journal of Catholic Thought and Culture*, vol. 16 no. 2, 2013, p. 63-72.

Orosco, José-Antonio. 2022. *Star Trek's Philosophy of Peace and Justice*. New York: Bloomsbury.

Roberts, Adam. 2016. *The History of Science Fiction*, 2nd edition. New York: Palgrave.

Ruyer, Raymond. 1950. *L'utopie et les utopies*. Paris: G. Monfort.

Suvin, Darko. 1979. *Metamorphoses of Science Fiction*. New Haven: Yale.

Szendy, Peter. 2013. *Kant in the Land of Extraterrestrials: Cosmopolitical Philosofictions*. New York: Fordham University Press.

Wells, H.G. 2005. *The Time Machine*. London: Penguin Book.

Wells H.G. 2017. "Preface to the Scientific Romances." In *Science Fiction Criticism: An Anthology of Essential Writings*. Edited by Rob Latham. London: Bloomsbury.

Wells, H.G. 2018. *The War of the Worlds*. New York: Vintage Classics.

Wells, H.G. 2022. *The World Set Free*. Cambridge, MA: MIT Press.

Wittgenstein, Ludwig. 1980. *Culture and Value.* Translated by Peter Winch. Chicago: University of Chicago Press.

Notes

[1] I say "we" here because I am an academic philosopher who has struggled with this problem in my own teaching, writing, and in public lectures.

[2] Also misattributed to Albert Camus and others in a variety of places. See https://quoteinvestigator.com/2019/10/29/lie-truth/.