The epistemic value of speculative fiction

This is the draft version of a paper that appeared in *Midwest Studies in Philosophy*, 2015, vol 39, 58-77.


Johan De Smedt & Helen De Cruz

1. Storytelling and philosophizing

In Daniel F. Galouye’s novel *Simulacron-3* (1964), a team of computer scientists employs an elaborate computer simulation to reduce the need for marketing research in the actual world. The agents in the simulation are conscious and do not realize they live in a simulated world. When Morton Lynch, one of the scientists, mysteriously disappears, his colleague Douglas Hall attempts to find out what happened to him, only to realize that nobody else even remembers the vanished man. Gradually, it transpires that the world Hall inhabits is also a simulation, and thus that their creation is a simulation within a simulation. The fact that Hall remembers Lynch is a computer glitch. The novel explores several philosophical topics: if the creator of a simulated world turns out to be a malicious sadist (as is the case in the novel), can he be considered a god to his creation? Do simulated beings have souls? If there is an afterlife, will only “real” people go there, or also simulated beings? Can you fall in love with a simulated being? How do people from the level above know they are living in the real physical world, or do they perhaps also inhabit a simulation? Schwitzgebel and Bakker (2013) explore similar issues in their short story *Reinstalling Eden*, focusing on the moral responsibilities of creators of simulations to their simulated entities.

Nick Bostrom (2003), in his philosophical paper *Are we living in a computer simulation?*, uses a weak indifference principle and probability calculus to argue that it is either highly likely that we are living in a computer simulation, or that future generations will never run any simulations (for instance, because they go extinct before reaching the technological stage when they will be able to simulate). Bostrom echoes Galouye’s conclusion, i.e., if you can simulate minds it is very likely you are living in a simulation. Galouye, Schwitzgebel, Bakker, and Bostrom engage in speculative thinking about computer simulations. But only the fictional stories explore broader ethical consequences of simulations, such as the moral responsibilities toward the simulated world, and metaphysical puzzles, such as the ontological status of simulated beings.

What are the similarities and differences between fictional stories and philosophical thought experiments¹? In this paper, we use insights from cognitive science to argue that analytic philosophy (especially philosophical thought experiments) and speculative fiction (such as science fiction and fantasy) rely on similar cognitive mechanisms. These include episodic future thinking (to consider future possibilities) and counterfactual reasoning (to imagine hypothetical possibilities). However, speculative fiction, unlike analytic philosophy, elicits

¹ We discuss philosophical thought experiments in detail in subsection 3.1.
transportation by drawing readers emotionally into a story and reduces the need for cognitive closure. As a result, speculative fiction allows for a richer exploration of philosophical positions than is possible through ordinary philosophical thought experiments.

In section 2 we highlight some of the cognitive mechanisms underlying philosophical thought experiments and speculative fiction, and consider the similarities and differences between these narrative formats. Section 3 argues that speculative fiction has a unique epistemic value: compared to the philosophical journal article or monograph, speculative fiction allows for a richer, more embodied, evaluation of philosophical positions. Section 4 applies the resulting framework to explore three novels, Robert A. Heinlein’s *The Moon is a Harsh Mistress* (1966), Jack Vance’s *Languages of Pao* (1958), and Stephenie Meyer’s *Breaking Dawn* (2008) and looks at how these works engage with enduring questions in political philosophy, philosophy of language, and philosophy of religion, respectively. In section 5, we summarize interviews conducted with philosophers who write speculative fiction, and examine how they engage in philosophy using the format of fiction. Section 6 provides a brief conclusion.

2. Cognitive mechanisms underlying philosophical thought experiments and speculative fiction

Telling stories is a universal human phenomenon, whether in written, oral, televised, balletic, or other forms. Stories may be cross-culturally universal because they are both sources of factual information (e.g., subsistence information in times of scarcity), and they allow for the risk-free exploration of what-if scenarios or hypothetical situations (Sugiyama 2011, Tooby & Cosmides 2001). Humans easily and spontaneously extract facts from fiction, often having the illusion they always knew facts they learned from stories. This is illustrated by Marsh et al. (2003), where participants claimed they knew a particular piece of information prior to reading the story in which it was presented, but that they could not possibly have known since it was invented for the sake of the experiment. In this section, we examine cognitive mechanisms underlying philosophical thought experiments and speculative fiction, focusing on future thinking, counterfactual reasoning, and transportation.

2.1 Mental prospection

Humans are unique in their ability to anticipate future events, to foresee the possible future outcomes of their current actions, and to examine alternative outcomes of past scenarios. These forms of thinking are collectively termed mental prospection. Neuroimaging studies have found that remembering autobiographical events, simulating the future, conceiving the viewpoint of others, and imagining hypothetical scenarios depend on an overlapping neural basis (Buckner & Carroll 2007). For example, people with damage to the parahippocampus not only have difficulties

---

2 This source amnesia (unawareness of the source of knowledge) is unsurprising, given that one needs extra cognitive resources to monitor sources.

3 Some other animals are able to consider future events, e.g., Western scrub jays, a species of caching bird, which can cache food in a location they foresee they will be hungry in the next morning (Raby et al. 2007). However, such capacities are usually domain-specific (food-caching in the case of scrub jays) and closely tied to the animal’s everyday ecological context.
remembering their personal past, but also to imagine themselves in hypothetical situations, such as what it would be like to lie on a sandy beach in a tropical bay (Hassabis et al. 2007). This brain area is also involved in the production and comprehension of narratives, presumably because people have to imagine elements of a fictional world to situate characters in (AbdulSabur et al. 2014).

Mental prospection is common to both speculative fiction and philosophical thought experiments. To take but two examples from speculative fiction: What if the Allies had lost the Second World War, and the Axis powers had triumphed, with Germany and Japan occupying most of the US (Philip K. Dick, The man in the high castle, 1962)? What if we lived in a genderless society, where we could temporarily assume a given gender, depending on whom we are in a relationship with (Ursula K. Le Guin, The left hand of darkness, 1969)? Examples from philosophy include: Would you choose to live hooked up to an experience machine, which would maximize pleasurable responses indistinguishable from real life (Nozick 1974)? Would you rather save five workers on a train track by diverting the train you are driving to a track with only one worker, or let the train run its course (Foot 1967)?

Thinking about the distant future is a creative and open-ended form of mental prospection. In one study, undergraduates were asked to imagine what the world would look like five years from now or 50 years into the future. Compared to the near-future participants, subjects who thought of the distant future were more creative in a subsequent design task (Chiu 2012). Mental simulations of distant events tend to focus on essential features of an experience, discounting the context in which it occurs, which leads to overestimations of how good a positive experience, or how bad a negative occurrence would be. By contrast, people tend to use low-level, concrete representations to think about the near future, e.g., what their day will be like (see Gilbert & Wilson 2009 for a review). In ordinary reasoning contexts, there thus seems to be a tradeoff between thinking about the near future versus the distant future: thinking of the near future encourages low-level, realistic representations with little creativity, whereas envisioning the distant future encourages high-level, unrealistic representations with high creativity. Speculative fiction combines the advantages of near-future and distant-future prospection: it retains the details from the former and the creativity found in the latter. For example, in The forever war (Haldeman 1974), humans wage war against the alien Taurans. Because of time dilation, centuries pass on Earth while the soldiers are only away for a few months; the combatants experience disconnection and alienation. Mundane things like food prices, sexual mores, medical technology, and language become increasingly unfamiliar after every tour of duty. In this way, the novel fleshes out some of the unpractical consequences of frequent far-away space travel.

People frequently engage in counterfactual thinking: they imagine hypothetical alternatives to past situations (e.g., “If I had bought a house five years ago, I’d have been able to obtain a mortgage”). Everyday counterfactual thinking is a form of thought experiment: it involves laying out a causal chain of events in a sequence, changing one element in the chain, and seeing what happens. Although counterfactual thinking could in principle depart radically from reality, people tend to remain quite close to reality when they envisage counterfactual situations. For example, they seldom change natural laws (few people would wonder “If only the plane fell up instead of down”). Typically, reasoners consider just one or a few alternative possibilities in more detail, rather than exploring the full range of alternatives (see Byrne 2002 for an overview).
Philosophical thought experiments and speculative fiction build on our natural ability for counterfactual thinking, but their elaborations go far beyond what people normally do. Like in everyday situations, philosophers and writers imagine just one or a few alternative possibilities\(^4\), but unlike in everyday forms of counterfactual thinking, these may include changes in natural laws. For example, the philosopher Carol Rovane (1990, 374) examines the consequences of personal identity if people’s reproductive systems were radically different:

Suppose it were a natural fact that people divided like amoebae, with the result that both descendants bore the relation of psychological connectedness to their common ancestor. Would we want to say even in that case that the ancestor would be unable to identify with both prospective future selves?

The science fiction writer Richard Garffinkle (*Celestial matters*, 1996) explores in great detail how space travel would work if we lived in the world of Aristotelian science, where spontaneous generation really occurs and where space travel is hindered by the celestial spheres. Hal Clement’s *Mission of Gravity* (1954), by contrast, obeys our universe’s laws of physics but its action plays out on the planet Mesklin, an oblate world with extreme variations in gravity (700 g at the poles, 3 g at the equator). The appeal of the novel lies in its fleshing out what creatures on such a planet would look like, and how they would deal with its gravitational challenges. In Roger Zelazny’s *Lord of Light* (1967), technology allows for reincarnation, by uploading people’s *atman* electronically into a new body. The advantage of such radical departures from reality is that they allow a detailed examination of states of affairs very different from our own.

2.2 Transportation and empathy

As we have seen, mental prospection is common to speculative fiction and philosophical thought experiments. Now, we will look at cognitive underpinnings of storytelling that are more prominent in the former than in the latter. *Transportation* is a metaphor coined by Richard Gerrig (1993) to describe the phenomenon of the reader being fully immersed and drawn into a fictional world. Stories are not the only cultural artifacts in which people can be absorbed. Games (especially open world computer games) and music can also provide a sense of transportation. However, stories are particularly effective in eliciting transportation. The reader (or listener, or watcher) is not only absorbed, but feels like they are part of the action, and experiences a reduced focus on the self.

Through transportation, fiction can elicit emotions and invite the reader to think along with the fictional characters’ mental states. Non-fiction works, such as pamphlets and sermons, can also elicit basic motions like anger, sadness, or disgust. For instance, Émile Zola’s open letter *J’accuse* (1898) proclaimed Alfred Dreyfus’ innocence against the charge of treason and accused the French government of anti-Semitism. Typically, such works elicit direct emotions, aimed at real-world people

\(^4\) There are a few exceptions. For example, in the *Chronicles of Amber* novel series (especially the first pentalogy, 1970–1978), Roger Zelazny depicts a fantasy universe where parallel worlds with their own history and natural laws are shadows cast by the quasi-eternal city of Amber.
and situations\(^5\); the letter stirred strong emotions among the French public and ultimately led to Dreyfus’ rehabilitation. Some cognitive scientists (e.g., Goldstein 2009) have suggested that fiction can elicit emotions by providing a safe, risk-free environment, where readers can experience emotions without these having a significant impact on their actual lives. For example, adults who compared their experience of a gloomy film to an unhappy personal event experienced similar levels of sadness, but felt considerably less anxious watching the movie compared to recollecting personal experiences (see also interview with R. Scott Bakker, subsection 5.2 on ‘crash space’ and ‘cheat spaces’). Bal and Veltcamp (2013) found that reading fiction increases empathy over time, but only if the reader feels transported into the story, and strongly identifies with the main characters. Non-transported readers, by contrast, become less empathic\(^6\).

Sometimes, one needs to be emotionally invested in a hypothetical situation to fully appreciate its consequences. In such cases, fiction seems a better tool to explore the consequences of particular philosophical views than philosophical thought experiments. Take utilitarianism, which requires that one should save the greater number of lives. In trolley scenarios, readers consider what the best action is, and typically judge that flipping a switch to save five people (killing one) is preferable to letting five people die and not intervening. However, in the format of thought experiments there is little transportation, and so it is hard to imagine how one might actually react in such a situation. In the first season of the reimagined Battlestar Galactica television series (2004), a fleet of spaceships carries the human survivors of the Cylon extermination of humanity. In the episode \(33\), the humans are relentlessly pursued by Cylon spaceships and make FTL-jumps to escape, only for the Cylon fleet to catch up with them every 33 minutes. The human civilian and military leaders suspect that the Olympic Carrier, a large passenger liner, is bugged, which allows their pursuers to catch up with them after each FTL jump. They reluctantly decide to blow up the ship with its 1,345 passengers\(^7\). This turns out to be the right (utilitarian) decision as the Cylon pursuit ends, and 47,973 human lives are saved. Still, it is a difficult decision, and a number of people feel guilty about it, e.g., president Roslin keeps a paper on her desk mentioning the ship’s name as a reminder of the tragedy.

3. The philosophical value of speculative fiction

3.1. Making context matter
Analytic philosophers rarely turn to fiction (a few exceptions, such as Williams 1973, notwithstanding), but employ their own particular format: thought experiments,

\(^5\) Pamphletary novels, especially popular in the late 19th century, also engage with real-world situations with the aim of improving them, e.g., Harriet Beecher Stowe’s Uncle Tom’s cabin (1852) was directed against slavery.

\(^6\) Transportation can also happen through non-emotional means, for instance, Kiera Cass’ The selection (2012) contains rapturous passages describing fine clothing and hair styles, a written-word equivalent of playing with Barbie dolls.

\(^7\) No one knows for sure if the passengers are still on board, or if the Cylons have evacuated them. On the DVD commentary accompanying this episode, Ronald D. Moore, the author of the episode, states the original script specified that the fighter pilots destroying the ship would see people inside. Under pressure of the network, this scene was cut but subsequently some ambiguous movement was shown behind the portholes.
comprising brief scenarios involving what Eleonore Stump (2010, 25) has aptly termed “the philosophical crash-dummies Smith and Jones.” In the method of cases, such stories serve to evaluate the plausibility of more general philosophical principles. Often the stories are outlandish. It seems highly improbable that one should ever find oneself in the situation of Alvin Goldman’s fake barns (1976), driving around in a neighborhood that contains mostly fake barn facades, rather than genuine buildings, that happen to be so positioned that one would not be aware of this fact. Nor would it be likely that one would find oneself linked up to a comatose violinist for nine months in a hospital bed (Thomson 1971). Other stories are more mundane and realistic. In order to argue that assertion does not require knowledge, Jennifer Lackey (2007) considers a creationist primary school teacher who, notwithstanding her beliefs, proclaims evolution to be true. To argue that disagreement with an epistemic peer should lead one to reevaluate one’s position, David Christensen (2007) has two people splitting the check after a meal and coming up with different figures.

Regardless of whether they are outlandish or realistic, philosophical thought experiments lack features that speculative fiction typically has, including vivid, seemingly irrelevant details that help to transport the reader and encourage low-level, concrete thinking. We do not empathize with the person who is connected, without her consent, to the violinist, or even with the person who faces the dilemma of diverting a train to kill one person in order to save five. While such decisions are difficult, we feel no emotional investment in making one choice rather than the other.

Through their lack of detail, thought experiments lead to a default mode of mental prospection where one focuses on essential features and disregards context. Philosophers might be attracted to this lack of consideration of irrelevant details, but these may be highly important for reaching philosophical conclusions. Professional philosophers, particularly ethicists, are as much influenced by framing and context as laypeople. Schwitzgebel and Cushman (2012) presented philosophers with scenarios in which one could save five people by sacrificing one by either diverting a trolley by means of a switch, pulling a lever to drop one person in front of the approaching vehicle, or physically pushing someone with a heavy backpack in front of it to save the five. They found that philosophers who were presented with drop or push before switch were more likely to judge the scenarios as equivalent than people who read switch first, and more likely to reject the doctrine of double effect⁸. This effect occurred in spite of expertise, training, or active reflection on the task (Schwitzgebel & Cushman 2015).

It seems problematic that the mere order of presentation of scenarios should make a difference as to how we judge them, which is why experimental results such as these are generally regarded as bad news for the method of cases (but see De Cruz 2015 for a different perspective). However, perhaps the high context sensitivity of philosophical thought experiments is philosophically relevant. After all, cognition is situated, and comprises embedded, embodied, enacted, and extended forms of thinking (see Menary 2010 for a brief discussion). It arises through interactions between humans and their environment, for example by taking advantage of objects in the external world to help perform cognitive tasks. Philosophical thought experiments already go some way toward this by fleshing out concrete scenarios to test abstract philosophical ideas, but their brevity limits how involved one can become in them.

⁸ The doctrine of double effect holds that it is morally permissible to perform an action that has a good result, but that also unavoidably has a bad, unintended side-effect, e.g., killing one person in order to save five.
Rather than trying to excise low-level, seemingly irrelevant details of thought experiments, we should introduce them and see what the difference is between the bald scenarios and the elaborate fictional cases. Narratives can achieve this in a way that thought experiments cannot, because they are packed with rich details that offer a contextual backdrop that allows one to realistically embed philosophical ideas in a make-believe world. They enable transportation and allow readers to emotionally engage with the characters, situations, and concepts.

3.2 Opening the space of philosophical possibilities
In analytic philosophy, there is a tendency to frame debates in terms of a few, well-outlined, often opposing positions. To give just three examples in epistemology, the philosophical debate on skills centers on whether knowing-how is somehow reducible to knowing-that (intellectualism versus anti-intellectualism). The debates on testimony are mostly focused on whether or not testimonial justification is reducible to other sources of knowledge, such as memory and perception, or is itself a basic source of knowledge (reductionism versus anti-reductionism). Arguments about disagreement probe whether one should revise one’s beliefs in the face of a dissenting peer (conciliationism versus steadfastness). As a result of chalking out opposing views early on in a philosophical debate, philosophers close off a large space of possible views. There are many other things one could say about skills, testimony, and disagreement, but they are mostly left unsaid.

Why do philosophers so frequently get stuck in a rut of opposing frameworks? One reason might be that the formats in which philosophy is typically practiced—the journal article or the paper presented in front of a not so lively audience—increase a need for cognitive closure, “the need to reach a quick conclusion in decision-making and an aversion to ambiguity and confusion. It encourages ‘seizing’ on an early statement or proposition in the process of acquiring knowledge, followed by rigidly ‘freezing’ on the seized item, and remaining impervious to additional information” (Djikic et al. 2013, 149). The need for cognitive closure is usually precipitated by a requirement to act or reach a decision. It impedes creative thinking and open-mindedness, because it encourages reasoners to disregard positions that people formulate later on in a debate. Given the format of a typical analytic philosophy paper, authors indeed need to come to a conclusion—rarely do we see that authors remain undecided at the end. In this way, the format of the analytic philosophy paper encourages close-mindedness.

Practicing philosophy through speculative fiction can provide a valuable antidote to the need for cognitive closure. In an experimental situation, people who read fiction (compared to those who read a non-fiction piece of similar length and reading difficulty) had a decreased need for cognitive closure (Djikic et al. 2013). For

---

9 Prestige is another reason for this practice of sticking to a few well-trodden paths. Philosophy is a prestige-driven discipline, with a strong belief in brilliance (see Leslie et al. 2015, Fig. 1b). Newcomers are expected to frame their discussions in terms set out by influential thinkers. As Kieran Healy (2015) notes, the majority of papers in what are considered top general philosophy journals are hardly or not cited. A tiny fraction of papers “become centers of gravity that define what a field is about... Success means you structure the substance of the field. It’s like an Olympic event where the path taken by the winners also sets the shape of the track for future competitors.”
instance, they were less inclined to agree with statements such as “I dislike questions which can be answered in many different ways” and “I don’t like situations that are uncertain”. Plausibly, fiction has this effect by not requiring immediate resolution or action on the part of the reader, and by encouraging empathy for the characters, which may open one’s mind to the viewpoints they represent. For example, one may feel resistance reading a monograph defending libertarian political philosophy, such as Robert Nozick’s (1974) *Anarchy, state and utopia*. Reading a novel that explores similar issues, such as Heinlein’s *The moon is a harsh mistress* (see subsection 4.1), puts one in a different frame of mind: one need not accept or reject the libertarian views of its author, but can go along with the story and see what a fictional libertarian society looks like.

While philosophical thought experiments are also narrative in format, they lack empathic engagement and transportation. As a result, they do not elicit creative open-ended thinking to the same extent as speculative fiction. To the contrary, influential thought experiments, such as Edmund Gettier’s (1963) scenarios of justified true belief that are not knowledge, beget a cottage industry of papers that discuss minor variations of the same thought experiment and the plausibility of the intuitions it elicits. By expanding their philosophical work outside of the confines of academic journals or books, for example by turning to speculative fiction, philosophers can expand their conceptual toolkit.

### 3.3. Assessing the consequences of philosophical positions

Given that fiction can increase empathy when readers feel transported, speculative fiction can help us assess the consequences of philosophical positions in which empathy matters. Discussing atrocious situations in a dispassionate manner, such as the permissibility of torture in political philosophy, or the soul-building qualities of evil in the philosophy of religion, are not conducive to empathizing with the victims of torture or evil. As Shaun Nichols (2006, 465) points out, “Philosophical thought experiments are often unaccompanied by the affective consequences that would result if we really believed the situation described in the thought experiment.”

Philosophical thought experiments are austere and focused on a particular philosophical problem, which may explain their lack of emotional engagement. For example, in Frank Jackson’s (1982) Mary’s room thought experiment, we are to wonder whether the experience of red is reducible to reading about color psychophysics, whereas one’s first thought would be to commiserate with Mary in her isolation. “We want primarily to solve the hypothetical problem and accordingly focus on the minimal conditions. In those cases, then, the disturbing inferences are not drawn out and hence they do not reach the affective system” (Nichols 2006, 472). What it would feel like to experience the color red is probably not very high up Mary’s list of priorities, having been just released from the confines of her black-and-white prison. After all, she has missed out on a wealth of interpersonal and other experiences ordinary people take for granted.

### 4. Three examples of philosophical ideas explored in speculative fiction

We will now look at three examples of speculative fiction that explore philosophical topics, respectively in political philosophy (libertarianism and anarchism), philosophy of language (the strong version of the Sapir-Whorf hypothesis), and philosophy of religion (metaphysics of persons in an embodied afterlife).
4.1. A libertarian society: *The moon is a harsh mistress*

Robert A. Heinlein (1907–1988) explored a wide variety of topics in political philosophy and ethics throughout his novels and short stories. For instance, *Jerry was a man* (1947) centers on civil liberties, particularly on the moral status of genetically modified chimpanzees. *Starship Troopers* (1959) presents an authoritarian military regime where only people who have been in active military service have the right to vote. *Stranger in a Strange Land* (1961) explores the fictional Church of all Worlds, which advocates free love among its adherents.

As Heinlein’s readers get transported into his stories, they get a sense of what it would be like to live in societies that radically differ from our own in their ethical norms and political organizations. *The moon is a harsh mistress* (1966) describes the successful revolution of Luna against the political and economic hegemony of the Earth. Originally, the moon was a penal colony where governments sent their criminals and political exiles. As a result of physiological adjustments to the lower gravity, Loonies (as the inhabitants of Luna call themselves) are unable to return to Earth. In the novel, many adults on Luna are third-generation and have never visited Earth. The story describes in detail an idealized libertarian society that is radically anarchistic. Luna does not even have laws for people to abide to, only customs, such as paying one’s debts. People who commit serious offences are thrown out of the nearest airlock without formal trial; this form of lynching is common. The liberties of Luna’s inhabitants are unconstrained, for instance, there are no impositions on how people can organize marriage. Heinlein expounds on line marriage, a long-term polyamorous arrangement between several men and women, with people able to opt into an existing marriage so that the line marriage can continue even after the death of its founding members. This system provides a social safety net, parental stability for children, and allows for the accumulation of wealth and resources.

As a political philosophy, libertarianism is a theory of justice. Libertarians hold that personal liberty is a supremely important good, and that others (including the government) should only interfere with it to the extent that exerting one’s personal liberties would infringe upon the liberties of others. When the revolutionaries are considering what the government of an independent Luna should look like, Bernardo de la Paz (the mouthpiece of most of the anarchistic libertarian ideas in the book), says “government is a dangerous servant and a terrible master. You now have freedom—if you can keep it. But do remember that you can lose this freedom more quickly to yourselves than to any other tyrant” (Heinlein 1966 [2008], 300), and “Comrades, I beg you—do not resort to compulsory taxation. There is no worse tyranny than to force a man to pay for what he does not want merely because you think it would be good for him” (Heinlein 1966 [2008], 302-303). While it may be that the new Lunar government can get by without taxation, it is unclear how its inhabitants would make sure people do not infringe on the rights and liberties of others, a problem of any anarchistic system. Heinlein solves this by appealing to institutionalized (but not official) vendettas of the line marriages and friend networks: “we figure this way: If a man is killed, either he had it coming and everybody knows it—usual case—or his friends will take care of it by eliminating man who did it. Either way, no problem. Nor many eliminations” (Heinlein 1966 [2008], 165-166).

Another problem of libertarianism is epistemic: present differences in wealth came about through mechanisms that are unjust, such as slavery, feudalism, looting during wars. Since it is hard to determine what right violations took place in the past, it is difficult to make out what justice would require in the present situation (Vallentyne & van der Vossen 2014). On Luna this epistemic problem hardly appears
because the Loonies have started on an equal footing in the recent past (as shipped convicts), and wealth accumulation is a fairly recent phenomenon. Thus, the lunar society avoids both the practical problem of how to ensure everybody respects the freedom of others and the epistemic problem of recognizing and rectifying past injustices. In this respect, Luna is an anarchist utopia, presenting in detail how an anarchist society could function.

4.2. Language and thought: The languages of Pao

Does language shape our thoughts, and if so, in what way? This is an enduring question in philosophy as well as cognitive science. Some authors propose that language shapes thought, for instance, by helping to structure cognitive tasks or by ordering our environment into concepts (e.g., Clark 2006). Others take the rich cognitive capacities in animals and infants as evidence that most thought processes are language-independent, and hold that language only serves a purely communicative function (e.g., Bloom 2000). If language shapes our thoughts in a non-trivial way, one natural consequence would be that different languages result in different thought patterns. This hypothesis is termed linguistic relativity or the Sapir-Whorf hypothesis. The linguist Benjamin Lee Whorf (1956, 213) speculated that speakers of different languages experience the world differently, “We dissect nature along lines laid out by our native language.” While there is little evidence for radical forms of the Sapir-Whorf hypothesis, according to which language would completely determine how speakers see the world (see Wolff & Holmes 2011 for review), there is substantial evidence for moderate Whorfian effects in several domains of cognition. For instance, people find it easier to see subtle differences in color hues when these correspond to basic color terms in their native language: speakers of Russian have basic color terms for light blue (goluboy) and dark blue (siniy), and they find it easier to distinguish between blues that are on opposite sides of the goluboy/siniy border than non-Russian speakers (Winawer et al. 2007).

Many works in speculative fiction put forward invented languages. J.R.R. Tolkien is unparalleled in his use of invented languages, such as the Elvish and Dwarvish language families in Lord of the Rings. Fewer works examine philosophical views about language. One example is Jack Vance’s The languages of Pao (1958) which examines the Sapir-Whorf hypothesis in its strong form. Pao is a peaceful planet with a mostly rural, homogenous population that speaks a single language, Paonese. The panarch, the absolute ruler of this planet, is killed by his brother, who then uses an off-planet consultant, Palafox, from the planet of Breakness, to help him with the military and economic reorganization of Pao. Palafox proposes to change the mentality of the Paonese by changing their language: “We must alter the mental framework, of the Paonese people—a certain proportion of them, at least—which is most easily achieved by altering the language” (Vance 1958 [1989], 57). He devises several classes of specialized linguistic groups, including warriors, engineers, and merchants, by developing languages that create the appropriate thought patterns, Valiant, Technicant, and Cogitant, respectively. “All these languages will make use of semantic assistance. To the military segment, a ‘successful man’ will be synonymous with ‘winner of a fierce contest’. To the industrialists, it will mean ‘efficient fabricator’. To the traders, it equates with ‘a person irresistibly persuasive’” (Vance 1958 [1989], 59). During this linguistic upheaval, Beran, son of the murdered panarch, studies linguistics on Breakness. Upon assuming leadership of the planet, he dissolves the language castes, and integrates them within the larger community for mainly political reasons: only by uniting the castes into a new linguistic community can he be
sure that their loyalties lie with Pao. Using his linguistic training, he instates Pastiche, a recently invented language that combines elements from Paonese and the three engineered languages, as the new language of Pao. In this way, he aims to harvest the felicitous properties of the artificial languages for the prosperity of the planet.

Unfortunately, there are almost no examples of linguistic expressions in Paonese or the other languages. The relevance of the novel for philosophy of language lies in the high-level exploration of the consequences of the strong version of the Sapir-Whorf hypothesis. If languages dramatically shape our thinking, then there are only linguistically-laden ways to look at the world. In the words of Finisterle, one of Beran’s tutors, “we know of no ‘neutral’ language—and there is no ‘best’ or ‘optimum’ language, although Language A may be more suitable for Context X than Language B... we note that every language imposes a certain world-view upon the mind” (Vance 1958 [1989], 83). The languages of Pao is also unique in its exploration of the ethical consequences of engineering language communities for political purposes. As the Sapir-Whorf hypothesis in its strong form is probably incorrect, we do not have real-world cases of languages being engineered to impose a particular worldview on its speakers. However, in The languages of Pao, we can engage in a thought-experiment that examines what could happen if the Sapir-Whorf hypothesis were true.

4.3. Persons and the afterlife: Breaking dawn

Stephenie Meyer’s young adult fantasy tetralogy Twilight, New Moon, Eclipse, and Breaking Dawn recounts the love story of human teenager Isabella (Bella) Swan and her vampire boyfriend Edward Cullen. The novels read as an exploration of Mormon afterlife beliefs. Meyer is a devout member of the Church of Jesus Christ of the Latter Day Saints (LDS), whose members are known as Latter Day Saints or Mormons. It is a post-Christian religion founded by Joseph Smith in the 19th century. She considers the LDS faith as having “a huge influence on who I am and my perspective on the world, and therefore what I write” (Meyer, cited in Baptiste 2010, 25-27).

The LDS church has an elaborate theology of the afterlife. Like most Christian traditions, it holds that humans will be resurrected in bodily form some time after death. However, it differs from Christianity in having a materialist, monist view on personhood and the soul. Everything in the world, including God, spirits, and angels, is material. Humans, like other intelligent entities, are eternal. Before their earthly existence, they were spirits, which are also material, albeit more ethereal than bodies made of flesh and bone. After their resurrection, humans will live their lives in a way that is continuous with their mortal existence: they will eat, drink, converse, walk, talk, and play musical instruments. Unlike in Christianity, where marriage is dissolved upon the death of one of the partners, Mormons can opt to be “sealed” in the temple so that their marriage continues after death. Parent-child relationships also continue into the afterlife (see also De Cruz & De Smedt in press for an exploration of Mormon afterlife theology).

---

There are several artificial languages, such as Bahasa Indonesia and Ivrit, that were imposed on the inhabitants of respectively Indonesia and Israel for nationalistic and political reasons, but these languages were not engineered to influence the thought patterns of their speakers.

Parley Parker Pratt (e.g., 1845) developed this materialistic ontology (see Park & Watkins 2010 for further discussion).
In the early chapters of *Breaking Dawn*, Bella dies after giving birth to her and Edward’s daughter Renesmee. She is saved by an injection of Edward’s venom into her heart, and transforms into a vampire. This transformation echoes Mormon portrayals of the afterlife. Edward and Bella continue as a couple, presumably into eternity, maintaining their parental relationship to their quickly-growing daughter. Bella is similar to her pre-vampire state, but better: her senses are more acute, she is more agile, stronger, more beautiful. When she speaks, her voice “rang and shimmered like a bell” (Meyer 2008, 363). A similar view of postmortem life can be found in Aquinas (Van Dyke 2014, 277n20), who had an indelible influence on Christian afterlife theology. According to Aquinas, “the glory of the risen bodies exceed[s] the natural perfection of the heavenly bodies so as to have a greater lightsomeness, a more stable incapacity for suffering, an easier agility, and a more perfect worthiness of nature” (Aquinas 13th c [1957], chapter 86, 6). Human Bella was clumsy, always tripping and stumbling; vampire Bella is in perfect control of her body: “weakness is what we experience in a body found wanting in the strength to satisfy the desire of the soul in the movements and actions which the soul commands, and this weakness will be entirely taken away then, when power is overflowing into the body from a soul united to God” (Aquinas 13th c [1957], chapter 86, 3).

Aquinas believed that people would not do much with these wonderful bodies, however—all they would do is stand on a crystal sphere, beholding God’s splendor (the beatific vision). There would be no further growing in knowledge, love, or human relationships (see Van Dyke 2014 for discussion).

By contrast, the Mormon afterlife is more in line with our premortal existence, filled with intellectual projects and social activities. With so much time on their hands (aided by the fact that they do not need to sleep), the vampires can develop their individual interests: Carlisle Cullen is a master surgeon, whereas his adopted son Edward is an accomplished pianist. As Edward says, “There’s a reason why I’m the best musician in the family ... read the most books, studied the most sciences, become fluent in the most languages ... I’ve just had a lot of free time” (Meyer 2008, 449).

Through the transformation of its female protagonist, *Breaking Dawn* attempts to give a preview of resurrection in the Mormon tradition. For example, Bella describes her newly-found superior visual perception as follows “Behind the light, I could distinguish the individual grains in the dark wood ceiling above. In front of it, I could see the dust motes in the air, the sides the light touched, and the dark sides, distinct and separate. They spun like little planets, moving around each other in a celestial dance” (Meyer 2008, 357).

Meyer also indirectly explores the so-called tedium of immortality, the philosophical suggestion that immortality ultimately becomes unbearable for creatures like us because eventually any activity would become tedious and boring. Revealingly, Bernard Williams (1973) who started the philosophical literature on this topic used a case from fiction, Elina Makropulos (a character in an opera by Janáček, based on a play by Čapek) whose life has become tedious and boring after centuries of living. By contrast, Mormons believe that people will never get tired of living (always inventing new projects and staying committed to their spouses). Carlisle Cullen, the patriarch of the family who was born in the 17th century, is slightly older than Elina Makropulos but does not seem to have lost his joie de vivre and his interests in medicine. What would happen after thousands of years, when the vampires have read every book, mastered every skill (Shea 2009)? Will boredom inevitably set in?
Meyer’s vampires evade the tedium of immortality by leading lives that are rich in social relationships, the pleasure of hunting big game, and intellectual projects. Stephenie Meyer would agree with John Fischer (1994, 261) as he objects to Williams (1973)

an immortal life could consist in a certain mix of activities, possibly including friendship, love, family, intellectual, artistic and athletic activity, sensual delights, and so forth. We could imagine that any one of these would be boring and alienating, pursued relentlessly and without some combination of the others... But of course from the fact that one’s life will be unending it does not follow that it must be unitary or unbalanced. That one’s life is endless clearly does not have the implication that one must endlessly and single-mindedly pursue some particular sort of activity.

In the Mormon afterlife, one continues to progress, eventually becoming another God who has his own creatures (there is no ontological difference between God and humans – God was like us, and we will be like him, but at that point, God will have moved on). So in Mormon theology, there is no tedium of immortality.

There is also an ethical dimension to the Twilight tetralogy: how does one live one’s eternal life in such a way that it contributes to the flourishing of both oneself and others? Most vampires in the novels are portrayed as thoroughly selfish creatures, regarding humans merely as food, and vying with each other for political power. The Cullens are exceptional: they form committed, long-term family relationships that transcend the opportunistic covens of other vampires, abstain from feeding on humans, and distance themselves from vampire politics. In this respect, they resemble the ideal, righteous LDS family and portray how such a family might enjoy an eternal blissful existence. By her exploration of the story of Bella and Edward, Meyer provides a keen insight into philosophical questions of immortality, including a sense of what it could be like to be immortal.

5. Interviews with philosophers who write fantasy or science fiction

To further examine our hypothesis that speculative fiction has unique epistemic value by its ability to transport the reader, elicit emotions, and allow for a better appraisal of the consequences of hypothetical situations, the second author conducted interviews with philosophers who write speculative fiction. We will first provide an overview of the philosophical topics these authors explore through their fiction. We then ask why they chose this medium.

5.1. Philosophical topics in speculative fiction

12 These authors have completed at least some graduate work in philosophy, and had at least one published novel or short story in science fiction or fantasy. Given that this is a small group, we relied on snowballing to obtain participants. Participants gave consent to have their interviews published online on the blog *Philosophical Percolations* (http://www.philpercs.com/), and to have excerpts appear in this paper.
All of the authors interviewed consistently explore a particular philosophical topic in their narratives. In his short stories, Mark Silcox (professor of Humanities and Philosophy at the University of Central Oklahoma) prospects how the [near] future might be if certain present trends continue. The images of the future they contain are often rather bleak and resolutely anti-humanistic, but I do also usually try my best to include at least a few decent jokes.

Eric Schwitzgebel (professor of Philosophy at UC Riverside) writes short stories about the moral status of robots and simulated beings.

One issue that really interests me—and which I think still has much more potential to be explored in science fiction—are the moral relations between beings in simulated worlds (“sims”) and the beings who run those sims, who have god-like power over the beings inside the simulated worlds. This is a central theme in some of my own work—for example in “Reinstalling Eden” (co-authored with R. Scott Bakker) and “Out of the Jar”. These issues connect with issues in theology (if we consider the sim-managers to be gods – as I think we should, taking the perspective of the sims), in animal rights and human enhancement, in the nature and value of personhood, and also in connection with the fundamental ethical question of what kind of world we aspire to live in.

R. Scott Bakker, professional fantasy author (e.g., *Prince of Nothing* series), had almost completed graduate school in philosophy at Vanderbilt University, but abandoned it in favor of a writing career:

I realized that I could turn the ‘man the meaning maker’ paradigm upside down, tell the story of a protagonist struggling to bring meaninglessness to an objectively meaningful world. Once I kicked that door down, the narrative and thematic possibilities seemed endless, and the books began to write themselves.

David John Baker (associate professor of Philosophy at the University of Michigan) notes

Most of the stories I write (more than half, anyway) take place in a future history that I consider pretty utopian. The humans in this setting are pretty far beyond *Homo sapiens* biologically. They’ve eradicated scarcity, cured senescence and live pretty peacefully in general. It seems to me that at this stage of development, if we ever reach it, the problem of nihilism and the meaning of life will loom large in people’s minds. They’ll start to wonder, what do we mean when we say that we’ve accomplished something good in building this civilization? How do we know there’s any such thing as goodness? I often find myself writing about characters who live in paradise and wrestle with this sort of doubt.
5.2. Motivations for writing speculative fiction

The interviewees agreed that speculative fiction has a unique philosophical value that cannot be found in a philosophy journal article or monograph. R. Scott Bakker draws attention to the richness of fantasy worlds as places to engage in counterfactual thinking (see also subsection 2.2):

Human cognition is bounded cognition, ecological cognition, only possessing enough exaptive ‘slop’ to mimic general cognition; the modern world as reorganized and represented by science, meanwhile, is drifting ever farther from the meaningful worlds we are adapted to solve. We live in what I call ‘crash space,’ environments where the problem-solving reliability of our basic heuristic toolkits can only degrade over time. On my view, fantastic secondary worlds provide readers with faux adaptive problem-ecologies, settings that can be reliably parsed and understood in psychological terms. They provide what I call ‘cheat spaces,’ places where we have learned, over time, to game our cognitive predilections. And in this sense, you can see fantasy fiction as a kind of cultural grave marker, a place to make-believe our dead relatives are alive. It shows what has become of meaning in the world.

Eric Schwitzgebel emphasizes the epistemic utility of emotions that fiction can elicit (see also our discussion of trolley cases in subsections 2.2 and 3.1):

Emotion seems especially valuable in thinking through moral issues. It’s one thing to think about the propriety of killing one person to save five others, if one does so in an emotionally disengaged way in the context of a three-sentence scenario involving a runaway boxcar. But one might have very different thoughts about the case if one was confronted with it in a full-length movie, with the protagonist on the tracks! I don’t think we should base our philosophy wholly on our reactions in such emotionally engaged cases. Sometimes, you need to set aside such emotions and be colder in one’s evaluations. But only sometimes, I think! At other times, emotional engagement with philosophical issues is probably exactly what you want.

Similarly, David John Baker remarks (in agreement with our discussion of Mary’s room in subsection 3.3):

What’s special about fiction is that it puts you inside the thought experiment. Parfit’s human fission examples are shocking when you first read them, but you don’t necessarily put yourself inside them and ask what it means for your life that you’re the sort of being who could in principle be split into multiple people. But in John Varley’s *The Ophiuchi Hotline*, we follow a character whose very humanity is threatened by being duplicated over and over, who feels guilty for the crimes of her duplicates. There’s a philosophical point you take away from hypotheticals like this—the propositional knowledge you gain from them, I suppose. But in fiction there’s a personal moral you take away. When it works, it works because the people in the fiction feel
the way real humans might feel when confronted with the hypothetical situation. In a philosophy paper, the characters in the thought experiments aren’t usually of much importance.

6. Conclusion

In this paper, we examined the epistemic value of speculative fiction. How can speculative fiction contribute to the evaluation and development of philosophical ideas? In some respects, speculative fiction and philosophical thought experiments are similar: both engage in mental prospection, including future thinking and counterfactual reasoning. Speculative fiction, in contrast to philosophical thought experiments, can transport readers and in this way elicit empathy as well as an in-depth examination of the consequences of philosophical positions. We looked at how three speculative fiction novels, The moon is a harsh mistress, Languages of Pao, and Breaking dawn vividly examine and bring to life philosophical positions in political philosophy, philosophy of language, and philosophy of religion respectively. Our interviews with philosophers who write speculative fiction confirm our hypothesis that speculative fiction allows for a better appraisal of philosophical positions due to its transportive qualities, attention to detail, and emotional engagement.

References


Byrne, R.M. (2002). Mental models and counterfactual thoughts about what might have been. Trends in Cognitive Sciences, 6, 426-431.


