



I, robot teacher

David W. Kupferman

To cite this article: David W. Kupferman (2020): I, robot teacher, Educational Philosophy and Theory, DOI: [10.1080/00131857.2020.1793534](https://doi.org/10.1080/00131857.2020.1793534)

To link to this article: <https://doi.org/10.1080/00131857.2020.1793534>



Published online: 27 Jul 2020.



Submit your article to this journal [↗](#)



Article views: 260



View related articles [↗](#)



View Crossmark data [↗](#)



I, robot teacher

David W. Kupferman 

School of Teaching and Learning, Minnesota State University Moorhead, Moorhead, Minnesota, USA

'I know the future is scary at times, sweetheart.

But there's just no escaping it'.

Ernest Cline, *Armada*

As I sit here on Planet Zoom during the global COVID-19 pandemic and try to figure out how to make my online educational foundations courses as interesting, compelling, and fun as their face to face versions, it has dawned on me that perhaps the most cogent analysis of our current moment, from a pedagogical perspective, was written in 1951 by Isaac Asimov (1951/1957). Titled 'The Fun They Had', this very short story (it fits on a one-page PDF) demonstrates precisely why I couldn't simply import my face-to-face courses as they looked and ran in the first two months of 2020 onto an online platform for the last two months of the semester and beyond.

Asimov's story takes place in 2157, two centuries out from when Asimov wrote it. There, an 11-year old girl named Margie and her neighbor, a 13-year old boy named Tommy, have a conversation about a book that Tommy found the day before. Physical books are rare in the future, as everything has been digitized, including learning. What is even more compelling for Margie is that the book is about school.

In Asimov's 2157, every student has an individualized 'mechanical' teacher, a robot that is engineered to ostensibly fit each child's learning abilities. 'School' takes place at home and has been atomized, so that children only interact outside of learning. When children are six, presumably at the start of their 'school' years, they learn how to take tests by learning how to fill out a punch card. On the day the story takes place, Margie's robot teacher is on the blink, and is being repaired by a mechanic. She is hopeful that the problem will necessitate the removal of the robot teacher for some time, as Tommy's was once when it lost its history content. Unfortunately for her, her robot teacher is repaired in short order, at which point she and Tommy part so that they can each go to 'school'.

During the time when her 'teacher' is being overhauled, Margie and Tommy engage in a heated discussion about Tommy's book. A few details seem to startle Margie. First, there is the idea that teachers were people, not robots. This is confusing to Margie as she does not see how a person could know as much as her mechanical 'teacher', nor would she want a strange person to live with her family. Tommy explains that teachers didn't live with children – instead they met in a designated building called a 'school'. On top of that, Tommy thinks a human could probably know almost as much as a robot teacher.

The second feature of the book, the notion that children would gather together at a school to learn, fascinates Margie. She hates what she knows as 'school', and is wistful as her mother calls her into the house once her teacher is back up and running. The story ends with her day-dreaming about 'the fun they had' in the school described in Tommy's book as her robot teacher

begins to drone on about fractions. Presumably Tommy has similar thoughts as he returns home to 'school'.

There are a number of elements to unpack here, including the role and importance of socialization as a factor of schooling, the potential for truly individualized learning, and the incessant creep of technology in education. There are also arguments to be made about the increasing facility of children raised in a digital world, and whether or not teachers will be necessary in the future. The story also raises questions about inequities of access: can everyone afford a robot teacher? Can everyone afford 'school'? As we have already seen during the pandemic of 2020, lack of computer and Internet access – to both hardware and software – has yet to be resolved. So will only wealthy children be able to access robot teachers? Or will robot teachers be 'gifted' to students and communities that are poorer, the way that charter networks and technology philanthropists are buying their way into disadvantaged school districts and policy ala Mark Zuckerberg and Newark Public Schools in New Jersey? And just what kind of schooling is this in 2157 – is it public, private, charter, or some other iteration that we haven't yet dreamed up?

The purpose of this essay is not to answer any of these questions, but to make the case for further analyses that may provide some answers and to put out a general call for educational researchers, philosophers, and theorists to begin engaging in futures thinking. Asimov was speculating about a world 200 years in his future, but he needn't have gone that far. With the exception of a few technical details, owing to the realities of his time, he pretty accurately lays out the problems with education a mere 69 years in his future. Indeed, none of what he describes is particularly new to us, even if it was new in 1951. He might as well have been describing the Khan Academy or Sugata Mitra's Hole in the Wall project in India, both of which are models of independent learning via the Internet (no human interaction required), or the move to online learning thanks to the unintended push of the coronavirus. We are already headed in the direction of Margie and Tommy and their robot teachers. But we are not looking 69 years in the future, as Asimov did, and so we are not speculating about what education may look like in 2089. But we should. And that is why Asimov's story is so important in our present, because it represents precisely what we need in education research right now: more creativity, more science fiction, and more analyses of our current moment through wondering about the future.

Futures studies

Unfortunately it seems that we have lost the creativity of futures thinking, and importantly futures thinking in education, that defined the imagination of the 1950s and 1960s. That era was replete with visions of jetpacks, flying cars, Elroy's robot teacher Ms. Brainmocker on *The Jetsons*, and everything covered in chrome. Now there is a cynicism about the future that pervades our present moment. Perhaps that is because we never got those promised jetpacks, flying cars, or chrome. Perhaps it is because the wonder of the decades of the Space Race have given way to the privatization of space tourism and the farce of the Space Force (whether one is referring here to the Netflix show of the same name or the fanciful attempt by Donald Trump to militarize the moon is in the eye of the beholder). In any case, our current views of the future, by contrast to those of the mid-20th Century, are dim.

In the case of robot teachers, perspectives on their viability are also contested and met with a contrasting combination of excitement and disbelief. Ashok Goel at the Georgia Institute of Technology created an AI teaching assistant named 'Jill' for an online graduate course, arguing that AI – or robot teachers – can dramatically decrease the amount of work that professors and human TAs face in a given class. Goel recognized that he and his eight human TAs couldn't possibly respond to the 10,000 messages posted by his 300 students (Maderer, 2016). He also understood that students tended to ask the same questions repeatedly, which was an opening to unleash Jill on the course. Jill was so successful that students didn't realize 'she' was AI, which

begs for the development of some kind of Turing Teacher Test. This is quite a ways down the road from Margie's teacher in Asimov's story.

Yet while Jill is an outlier that represents an important glimpse into possible futures, in education the trend is to push back on such developments and make the case that education requires some kind of human interaction, preferably with someone trained in pedagogy. This argument makes sense, and has been made convincingly by a number of scholars of higher education (see Aoun, 2018; Poritz & Rees, 2017). On the website *Will Robots Take My Job*, teachers and other instructors are deemed 'totally safe' with a 1% chance of automation (*Will Robots Take My Job*, n. d.). And Anthony Elliott (2019) details the various ways that AI has and will continue to influence society, although education and robot teachers conspicuously play no role in his study.

But these are not the only perspectives. On the issue of automation and whether or not teachers can be replaced by robots or AI, the McKinsey Global Institute predicted that up to half of the workforce could be automated by 2055, including up to 27% of 'education services' (McKinsey Global Institute, 2017; for a more comprehensive survey and reactions to automation and education, see Peters et al., 2019). Not to be outdone, Anthony Seldon has made the case that robot teachers will begin replacing human ones by 2027 (Bodkin, 2017). But can robots really replace teachers? And if not, why not exactly? The move to online schooling in the first half of 2020 was fraught, largely because of the lack of familiarity of teachers at all levels, from pre-K to university, with technology as a tool for teaching. So if teachers don't keep up with technology, what will prevent that technology from passing them by? After all, teacher training programs are instructing students in how to be teachers today, not thirty years from now. But why aren't they?

Robot teachers are nothing new. Matt Novak (2013) provides a timeline of what I call 'robot teacher madness'. From comic strips in the 1950s to a National Education Association press release in 1960 seeking to ease parents' fears of the coming robot teaching force to the New York Worlds Fair in 1964–65 and its 'automated schoolmarm', the threat of robot teachers to human teachers – and to the human dimension of education itself – was ubiquitous in the social imaginary (Novak, 2013). So why are we now so unprepared to imagine them? Why have we lost our hysteria around robot teachers, and how can we try to recover some of that anxiety – and wonder?

Here is where we need education research to draw inspiration from futures studies. I should clarify from the outset that I use the plural 'futures', as does the World Futures Studies Federation, because of the multiplicity of futures that are possible (World Futures Studies Federation, 2020). Using the plural also signals a certain amount of ignorance about the future, the flipside of a certain amount of arrogance that goes into opining about the (one and only and singular) future. Additionally, I prefer the term futures studies as opposed to futurism or futurology, although in practice they all function pretty much the same way.

Since most people's encounter with the idea of 'futures' is probably a superficial brush with economic futures (what will the price of gold be in six months?), I should say that that is not what futures studies is about. Rather, futures studies is a niche academic area that draws from a number of disciplines, and really gained steam in the 1990s thanks to the work of Wendell Bell (1997), among others, even though the field has been around since the 1960s. To date there are five identified approaches to futures studies, roughly tracking the spectrum from positivism to post-positivism, although these distinctions and their applications are often blurred and by no means mutually exclusive. Sohail Inayatullah (2002) originally identified three dimensions of futures studies: predictive-empirical, in which language is neutral and deterministic and is the preferred method of policy planners; cultural-interpretive, in which comparisons can be made across societies through an assumption that language is contingent; and poststructural-critical, wherein current conditions of power and discourse are used to complicate visions of the future. Since then, Jennifer Gidley (2013) has added to futures studies the dimensions of empowerment-activist, which intends to apply proscriptive analyses through various forms of political and policy

action; and integral/transdisciplinary, in which environmental and planetary concerns are prioritized.

More importantly, and much more useful, are the various frames through which futures studies classifies alternate futures. Wendell Bell (1997) offers nine major tasks of futures studies, including the study of images of the future; the study of future studies' knowledge and ethical foundations; and interpreting the past/orienting the present, among others. Of these, a pretty nifty framework has evolved that can be summarized as: what is probable; what is possible; and what is preferable. For our purposes, and for educational futures, we can focus on what is likely to happen, what could happen, and what we want to see happen. The preferable is perhaps the most subjective, as it rests almost entirely on one's worldview, and as such it may wind up being the most unattainable. But it is a worthwhile undertaking, if only to draw out the distinctions and variations between where we're most likely to go and where we want to wind up.

I should also note that I am not making the case here for predicting any kind of future of education. The realm of prediction, forecasting, and the like are contested notions in futures studies (Bell, 1997), and are wrapped up in policy planning and other forms of legislating the future. Often included in the probable, the possible, and the preferable is a discussion of the proscription, but that is an approach to futures studies preferred by policy makers who tend to suck all the fun out of thinking about the future. I am interested here in considering the infinite futures that may or may not be realized, in order to understand both where we are today and where we *could* go (not where we will or must go), to think through alternative futures.

But futures studies, for the most part, isn't very much fun, and part of the reason is that there is an awful lot of proscription and theory, and a concerted lack of creativity, that goes into thinking about the future. Futurists, in academia at least, seem reluctant to put any kind of thoughtful futures scenario out there, and hold it up to scrutiny. As a result, much of futures studies scholarship, including in educational futures, is really about projecting an idealized version of the present onto some amorphous and overly-generalized 'future', or rather to solve the problems of the present (see, for example, Augé, 2014; or Hicks, 2002). This is not particularly interesting to read, and doesn't help us understand how we get to one or multiple futures from the present.

As such, part of the reason for this lack of speculating about the future in futures studies is that the field seems preoccupied not with the future but with the present – significantly, the present as it is informed by the past. An example of this comes from Gordon and Todorova (2019), where they offer future scenarios working through a dozen different examples of contemporary issues, such as politics, technology, health, and religion (interestingly, education is not one of their selected topics). Yet their frame is counterfactual analysis, which is rooted in a psychological concept that asks one to consider different outcomes had some decision or event in the past been different. They offer to apply this type of thinking to the future, with mixed results.

Counterfactual thinking also asks us to do just what it says: think counter to the facts. This is most popularly applied to history, and there has been a growing acceptance of counterfactual history in traditional history circles. Applied elsewhere, notions of counterfactual thinking serve as fun thought experiments, wherein we can ask what if historical moment X had/hadn't happened, and extrapolate out to our present to see how the world as we know it would be different. Of course, we do this all the time: what if I had gone to a better school, married a better person the first time around, saved more money when I was younger, and so on. A lot of time travel science fiction also works from this premise, especially on film, the most famous perhaps being the *Back to the Future* franchise (Zemeckis, 1985). In the first film, Marty McFly travels back in time to 1955 and makes his parents fall in love in circumstances preferable to the ones that they had before (wherein his father finally stands up to Biff, the school bully), and when Marty wakes up back in 1985, his parents are successful, his siblings have themselves put together, and a now-sheepish Biff is waxing Marty's new truck.

Beyond serving as the framework for revisionist history parlor games, counterfactual analysis can be related more broadly and applied to the realm of logic, where we see it butt up against reality, occasionally in entertaining ways. David K. Lewis (1973) devotes a whole book to counterfactual thinking, beginning with the mischievously simple statement, ‘if kangaroos had no tails, they would topple over’. From there, Lewis posits a number of possibilities, most famously that ours is one of many possible, multiple worlds. Here counterfactuals require us to imagine a world counter to the facts: if *x*, then *y*, and *y* is different from what we know as real. Importantly, Lewis offers a distinction between counterfactuals, one that is helpful when viewing them through a lens of futures studies. One type of counterfactual states ‘If it were the case that _____, then it *would* be the case that ...’; the other type of counterfactual states ‘If it were the case that _____, then it *might* be the case that’ (Lewis, 1973, p. 2, emphases added). It is this second version that complements the multiplicity of futures studies.

Yet I think that it is problematic for futures studies theorists to hew so closely to counterfactuals, the way that Gordon and Todorova (2019) do, in that they are still beholden to historical thinking. Attention to grammar seems important here. Historical counterfactuals ask in the past tense, if this *had* happened, then that in the present (or the future) would be this other way. Logical counterfactuals can ask the same question in the present tense: if this *is* the case, then that other outcome would be the result (Lewis’ kangaroos and their tails).

But there is another type of analysis, pioneered by Lawrence J. Sanna (1996), known as prefactual thinking. This type of analysis asks questions entirely in the future tense, so that if this *will be* the case, then some other sort of outcome might result. An illustration of prefactual thinking can be found in the movie *Avengers: Infinity War* (Russo & Russo, 2018), in which the Avengers finally confront Thanos, their most formidable foe yet in the Marvel Cinematic Universe. During a scene on Titan, Thanos’ home planet, a smattering of Avengers (including Iron Man, Spider-Man, half of the members of the Guardians of the Galaxy, and Dr. Strange) attempt to remove the Infinity Gauntlet from Thanos’ hand so that they can prevent him from completing his collection of Infinity Stones – all six of which, if placed on the gauntlet, would make him the most powerful being in the universe. While waiting for Thanos to arrive, Dr. Strange, who possesses the Time Stone, is seen in a sort of agitated meditation as he views possible futures given different courses of action the Avengers could take. As Dr. Strange says, ‘I went forward in time to view alternate futures. To see all the possible outcomes of the coming conflict’. Of the 14,000,605 alternate futures Dr. Strange sees, only one leads to Thanos’ defeat. Put another way, while counterfactuals ask us to imagine a world or worlds counter to our current facts, prefactuals ask us to imagine the world before we even know what those facts will be. As Sanna (1996) puts it, prefactuals ‘refer to the imagination, before the fact, of alternative possible predicted outcomes’ (p. 1020). I would argue that prefactuals, rather than counterfactuals, are the way we get to thinking about the future as an exercise in anticipation. Examples abound, if we know where to look for them. And like Margie’s robot teacher, most of these examples happen to come from a deep well of prefactual thinking: science fiction.

Science fiction

Anticipating the future isn’t an exact science. In fact, it’s hardly science at all (Gibbons & Kupferman, 2019; Shaviro, 2015), which is why it is largely the purview of science fiction writers and speculative futurists. But science fiction gets a bad rap because it is vulgar in the Latin sense, meaning ordinary or common. Partially because it is so common, it does not sit well within conventional academia. But that is precisely what I find so exciting about science fiction: the fact that it is popular should not be a mark against its utility and importance. Instead, science fiction, and popular culture writ large, is really the best way – the best tool there is – to make sense of abstract, theoretical, academic ideas, and to discuss those ideas with both

academic and non-academic audiences. That is why we need to infuse futures studies with some readily accessible science fiction. Futures studies needs a creative turn.

I am making the case here for science fiction first and foremost because it is *useful*. Popular culture and science fiction already exist, and can therefore be used to put 'meat' on the bones of an idea. Pop culture serves as a touchstone precisely because it is everywhere and is relatable, rather than being the province of a select few intellectuals. Everyone consumes pop culture and science fiction, even if they don't spend a lot of time deconstructing it. So one of our goals in educational futures should be to write about education policy as if it were science fiction. Given what has been coming out of the US Department of Education over the past two decades, for example, a lot of it already reads like a plot to destroy the earth.

Note that I am not trying to define what science fiction is. I have neither the time nor the interest to engage in that debate, so perhaps it is best to adopt US Supreme Court Justice Potter Stewart's description of pornography: 'I know it when I see it'. Science fiction is a blurry genre, running the gamut from horror to fantasy and all spaces in between (Kupferman & Gibbons, 2019), and so when I use that term here I mean it in the broadest possible sense. It is an inclusive, rather than exclusive, term when looking to it for examples of the future.

That being said, science fiction is popular and compelling for a few simple reasons. While some science fiction is about the past, most of it is about the future. But the futures we imagine, even the most abstract and horrifying, are actually reflections of the future given the tools we have in the present. So science fiction, while often futuristic, is really a mirror held up to where we are right now. It is a way to think about today, through the medium of some kind of future. It is not really what that future looks like that matters so much as how we envision it given our current moment. In this way, science fiction is not about resolving present crises. It is, like prefactual thinking, about anticipating what might be, using our current anxieties as inspiration.

Another reason for its popularity and ubiquity in popular culture is that science fiction is ultimately about us, and if there is one thing we like to read about, it is ourselves. Show me the most far-out, inexplicable science fiction tale, and I'll show you a master class in narcissism. That is why practically every 'alien' is coincidentally humanoid, and about 5 foot 10, or looks like a variation on some type of terrestrial reptile or plant. Our imaginations can scarcely travel beyond our own planet. Science fiction is, at its base, about what we think might become of us, for better or worse, and about how we get there *given where we are right now*. And it is precisely for these reasons that science fiction is both accessible – as a reader, as a writer – and therefore useable. It is really the best way to think about where we are by considering the – maybe weirder, maybe not – places we could go. In other words, science fiction is a set of funhouse mirrors, which only work if we have a basic grasp of what we think we look like, to ourselves and others, so that we can laugh and cry at all of those amusement park distortions. And that amusement park is located in futures studies: a combination of games of chance (the probable), freak shows (the possible), and the burlesque (the preferable). We need to approach educational futures as if entering a carnival (as opposed to a funeral home), with all the bells and whistles of experimental pop cultural examples available to us.

Ultimately science fiction is about future making (Montfort, 2017). That is why some of the most exciting examples of futures studies reside in the arts, architecture, technology, politics, and design (see, for example, Dunne & Raby, 2013; Frase, 2016; Grove, 2019; Lijster, 2019). Education is unprepared for the future because we're not imagining what it could look like. That is why we got caught flatfooted in the move to wholesale online learning in 2020 – no one asked what if all schooling were suddenly online, quite literally overnight? But this now seems like an excellent question to ask, and begs other questions as a result. What happens to pedagogy in an entirely digital/algorithmic future? What are the ethics of online learning, not gradually for a few students here or there, but globally? We have to think about how education will be *different* in the future, not the same as we know it today. After all, if we learned anything

from the immediate shift to online schooling, it is that it's not as simple as putting content into a program or platform. Or ... is it?

The point is, while other disciplines embrace the future-making aspect of futures studies, and incorporate the creative turn that is science fiction in following new threads of analysis, educational futures has not, yet. Futures studies provides a delightful framework. Science fiction offers some welcome – and not so welcome – detailed scenarios. Educational futures needs to embrace these infinite scenarios. Returning to Anthony Seldon's prediction about robot teachers replacing humans before the end of this decade, I argue that it is not enough to predict such things. We need more than predictions.

I, robot teacher

Despite my earlier skepticism about Gordon and Todorova's (2019) use of counterfactual thinking in futures studies, they do offer a nice conceit by way of what they call 'point scenarios' – 'a day in the life of ...' vignettes with which to consider the future (p. 8). Each of their chapters is structured around a series of short possible scenarios that then offer an opportunity for them to examine how 'today's unresolved issues might be resolved in the future' (p. 8). While I am interested not in solving today's concerns, but rather in exploring our various futures given where we are starting from today, the idea of point scenarios lends itself comfortably to prefactual thinking through the art of science fiction. Indeed, much of science fiction is just that: point scenarios of the future that ask us to engage with uncertainty.

Two recent examples from political theory come to mind. The first is Peter Frase's *Four Futures* (2016), a book that takes as its starting point four different outcomes of a post-capitalist world. Frase imagines future social and economic constructs of communism, rentism, socialism, and exterminism using combinations of hierarchy, equality, scarcity, and abundance. Another is the epilogue to Jairus Grove's *Savage Ecology* (2019), in which he describes, in harrowing prose, the corpse-strewn coastline of California in 2061, on the sixtieth anniversary of 9/11. There is perhaps no more effective way to wrap up a study of the end of humanity and the endgame of capitalism and war. Both of these illustrations provide examples of a way forward for educational futures, as they are as creative and imaginary (if also despairingly dystopian) as they are intellectual. They are exemplars of scholars having some fun.

So I would like to end this argument by engaging in a point scenario of my own, and in order to bookend this essay with robot teachers. The term robot first entered the English language in 1921 with Karel Čapek's play *RUR (Rossum's Universal Robots)*, using the word for 'forced work' in Czech: *robot*. In the prologue to the play, Henry Domin, the central director of Rossum's Universal Robots, is explaining the history of robots and the island factory of RUR to Helena Glory, who will later on reveal herself to be an advocate for liberating robots. During their conversation, the topic turns to the ability of robots to learn, a precursor of sorts to discussions about AI and algorithmic learning potential. As Domin explains, 'They learn to speak, write, and do calculations. They have a phenomenal memory. If you were to read them a twenty-volume encyclopedia they could repeat the contents in order, but they never think up anything original. They'd make fine university professors' (Čapek, 1921/2004, pp. 13–14). Despite the humorous aspersion cast on academics, the implication here is that robots are limited by what kinds of input they receive from their human programmers, although the play turns on the robots' achieving self-awareness and declaring they have souls.

Which brings us back to Asimov, and Margie's robot teacher, as well as Jill at Georgia Tech. While Margie's teacher appears to be a basic content-delivery platform, and therefore most likely not self-aware and not learning, Jill is another window into the future entirely. Over the course of the semester as a TA, Jill learned not only how to efficiently respond to thousands of student inquiries, but also to appear to those students as a human. It is a worthwhile exercise to

consider what Jill could do in terms of algorithmic learning with the intent of teaching were she to encounter, for example, quantum computing.

In the short story 'Runaround', Asimov (1942/1963) introduced the Three Laws of Robotics. He refined these laws over the years, and they have become touchstones in contemporary science fiction, but the original version is worth repeating here:

One, a robot may not injure a human being, or, through inaction, allow a human being to come to harm ...
 Two, a robot must obey the orders given it by human beings except where such orders would conflict with the First Law ...
 And three, a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws. (p. 51)

Elsewhere Asimov also included a Zeroth Law: A robot may not harm humanity, or, by inaction, allow humanity to come to harm. While Asimov would play with this law in terms of what 'harm' means, it begs a few questions for us. In the age of online teaching and the spectre of AI, we should ask if we might apply Asimov's laws to robot teachers. Does Margie's robot teacher follow these laws? Does Jill at Georgia Tech? Do they 'harm' their students? What about the qualitative dimension of robot teachers?

In his defense of the humanities in the future of the university, Joseph E. Aoun (2018) lays out three probable strands of higher education in what he calls humanics: data literacy, technological literacy, and human literacy. Here universities will be concerned with teaching the comprehension and use of big data, the software and hardware collecting and operating with that data, and what might be called a liberal arts gone digital. But what I find most compelling is the use of the term humanics, because it is a term that Asimov coined in 1987. Humanics, according to Asimov, is a set of laws designed to ensure that robots are aware that humans should also engage with each other along the lines of the Three Laws of Robotics. In this way,

A human being may not injure another human being, or through inaction, allow a human being to come to harm ...
 A human being must give orders to a robot that preserve robotic existence, unless such orders cause harm or discomfort to human beings ...
 A human being must not harm a robot, or, through inaction, allow a robot to come to harm, unless such harm is needed to keep a human being from harm or to allow a vital order to be carried out. (Asimov, 1987/1990, pp. 460–462)

These laws of humanics might be applied to the future of learning and teaching as well, so that the humans behind robot teachers and online content delivery first do no harm. After all, when we move all teaching online instantaneously, aren't we essentially posing as robot teachers? This idea is also taken up by Jonathan A. Poritz and Jonathan Rees (2017), who offer five of what they call 'Jonathans' Laws' of university teaching in the internet age. The first of these is 'Every real student deserves individual attention from, and interaction with, a real teacher' (p. 117). But they do not define a 'real' teacher (although presumably they mean a human one). Certainly Margie thinks her robot teacher is real, an observation that leads her to doubt the veracity of Tommy's book and its depictions of humans as teachers, since no person could possibly know as much as a robot. And if Ashok Goel's students think that Jill is 'real', what does it mean to disabuse them of that notion? Jill seems as real as anyone with whom I have interacted via email but never met in person, to say nothing of her fellow human TAs (who also likely never met their students in person).

So here is my point scenario with which to engage in some prefactual thinking about robot teacher futures, my Laws of Robot Teachers:

1. A robot teacher may not miseducate a human student, or, through inaction, allow a human student to come to miseducation.
2. A robot teacher must obey the pedagogy and curriculum given it by human beings except where such orders would conflict with the First Law.

3. A robot teacher must protect its own algorithmic learning as long as such learning does not conflict with the First or Second Laws.

And a Zeroth Law: A robot teacher may not harm humanity by disregarding individual learning needs, the importance of socialization, or the development of creativity as part of the human experience.

Are these sufficient? Are these elements of the probable, the possible, or preferable? Robot teachers are only one scenario of educational futures, and Margie's dislike of school and her robot teacher is but one outcome of one of those futures. Certainly her distaste for school is a result of her robot teacher's violation of my proposed Zeroth Law. But it is a start, and it is a way for us to begin thinking about the futures of education so that we can anticipate them, rather than be caught by surprise. We need fleshed out depictions of what educational futures look like. We need to widen our understanding of what our texts and areas of inquiry look like. We need to be creative and develop infinite point scenarios. And we need to have some fun. People should look at scholars of futures studies in education and think to themselves: the fun they had.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

David W. Kupferman  <http://orcid.org/0000-0001-5369-9431>

References

- Aoun, J. E. (2018). *Robot-proof: Higher education in the age of artificial intelligence*. MIT Press.
- Asimov, I. (Ed.). (1957). The fun they had. In *Earth is room enough: Science fiction tales of our own planet* (pp. 146–148). Doubleday. (Original work published 1951)
- Asimov, I. (Ed.). (1963). Runaround. In *I, robot* (pp. 40–58). Doubleday. (Original work published 1942)
- Asimov, I. (Ed.). (1990). The laws of humanics. In *Robot visions* (pp. 458–462). Penguin. (Original work published 1987)
- Augé, M. (2014). *The future*. Verso.
- Bell, W. (1997). *Foundations of futures studies: Human science for a new era. Volume I: History, purposes, and knowledge*. Transaction Publishers.
- Bodkin, H. (2017). 'Inspirational' robots to begin replacing teachers within 10 years. *The Telegraph*. <https://www.telegraph.co.uk/science/2017/09/11/inspirational-robots-begin-replacing-teachers-within-10-years/>
- Čápek, K. (1921/2004). *R.U.R. (Rossum's Universal Robots)* (C. Novak, Trans.). Penguin Books. (Original work published 1921)
- Dunne, A., & Raby, F. (2013). *Speculative everything: Design, fiction, and social dreaming*. MIT Press.
- Elliott, A. (2019). *The culture of AI: Everyday life and the digital revolution*. Routledge.
- Frase, P. (2016). *Four futures: Life after capitalism*. Verso.
- Gibbons, A., & Kupferman, D. W. (2019). Flow my tears, the teacher said: Science fiction as method. In S. Farquhar & E. Fitzpatrick (Eds.), *Innovations in narrative and metaphor: Methodologies and practices* (pp. 167–181). Springer.
- Gidley, J. (2013). Global knowledge futures: Articulating the emergence of a new meta-level field. *Integral Review*, 9(2), 145–172.
- Gordon, T. J., & Todorova, M. (2019). *Future studies and counterfactual analysis: Seeds of the future*. Palgrave Macmillan.
- Grove, J. V. (2019). *Savage ecology: War and geopolitics at the end of the world*. Duke University Press.
- Hicks, D. (2002). *Lessons for the future: The missing dimension in education*. RoutledgeFalmer.
- Inayatullah, S. (2002). Pedagogy, culture, and futures studies. In J. A. Dator (Ed.), *Futures studies in higher education* (pp. 109–122). Praeger.
- Lewis, D. (1973). *Counterfactuals*. Harvard University Press.
- Lijster, T. (Ed.). (2019). *The future of the new: Artistic innovation in times of social acceleration*. Valiz.

- Kupferman, D. W., & Gibbons, A. (2019). Why childhood ex machina? In D. W. Kupferman & A. Gibbons (Eds.), *Childhood, science fiction, and pedagogy: Children ex machina* (pp. 1–15). Springer.
- Maderer, J. (2016). Artificial intelligence course creates AI teaching assistant. *Georgia Tech News Center*. <https://www.news.gatech.edu/2016/05/09/artificial-intelligence-course-creates-ai-teaching-assistant>
- McKinsey Global Institute. (2017). *A future that works: Automation, employment, and productivity: Executive summary*. <https://www.mckinsey.com/~/media/mckinsey/featured%20insights/Digital%20Disruption/Harnessing%20automation%20fo%20a%20future%20that%20works/MGI-A-future-that-works-Executive-summary.ashx>
- Montfort, N. (2017). *The future*. MIT Press.
- Novak, M. (2013). The Jetsons get schooled: Robot teachers in the 21st century classroom. *Smithsonian Magazine*. <https://www.smithsonianmag.com/history/the-jetsons-get-schooled-robot-teachers-in-the-21st-century-classroom-11797516/>
- Peters, M. A., Jandrić, P., & Means, A. J. (Eds.). (2019). *Education and technological unemployment*. Springer.
- Poritz, J. A., & Rees, J. (2017). *Education is not an app: The future of university teaching in the Internet age*. Routledge.
- Russo, A., & Russo, J. (Directors). (2018). *Avengers: Infinity war* [Film]. Marvel Studios.
- Sanna, L. J. (1996). Defensive pessimism, optimism, and simulating alternatives: Some ups and downs of prefactual and counterfactual thinking. *Journal of Personality and Social Psychology*, 71(5), 1020–1036. <https://doi.org/10.1037//0022-3514.71.5.1020>
- Shaviro, S. (2015). *Discognition*. Repeater Books.
- Will Robots Take My Job. (n. d.). *Teachers and instructors, all others*. <https://willrobotstakemyjob.com/25-3099-teachers-and-instructors-all-other>
- World Futures Studies Federation. (2020). *About futures studies*. <https://wfsf.org/about-us/futures-studies>
- Zemeckis, R. (Director). (1985). *Back to the future* [Film]. Universal Pictures.