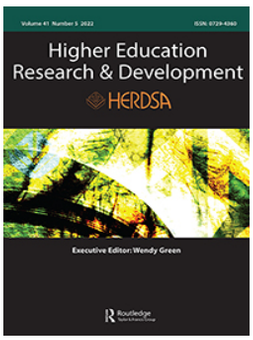

Higher Education Research & Development

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cher20

Criteria of quality in fiction-based research to promote debate about the use of AI and robots in Higher Education

Andrew Cox, David Cameron, Alessandro Checco, Tim Herrick, Maria Mawson & Richard Steadman-Jones

To cite this article: Andrew Cox, David Cameron, Alessandro Checco, Tim Herrick, Maria Mawson & Richard Steadman-Jones (2022): Criteria of quality in fiction-based research to promote debate about the use of AI and robots in Higher Education, Higher Education Research & Development, DOI: 10.1080/07294360.2022.2087603

To link to this article: https://doi.org/10.1080/07294360.2022.2087603

Published online: 22 Jun 2022.

HIGHER EDUCATION RESEARCH & DEVELOPMENT
https://doi.org/10.1080/07294360.2022.2087603

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Criteria of quality in fiction-based research to promote debate about the use of AI and robots in Higher Education

Andrew Cox [a], David Cameron[a], Alessandro Checco [a], Tim Herrick [b], Maria Mawson[c] and Richard Steadman-Jones[b]

[a]Information School, University of Sheffield, Sheffield, UK; [b]School of Education, University of Sheffield, Sheffield, UK; [c]University Library, University of Sheffield, Sheffield, UK

ABSTRACT

AI and robots have the potential to transform Higher Education (HE) but pose many ethical and implementation challenges. To ensure the widest debate about our choices for the future of HE with these technologies, engaging ways to present the issues are needed and this article is part of an exploration of the potential of fictional narratives to do so. Specifically, the purpose of this article is to enrich understanding of quality in such fiction-based research, through analysing responses to a collection of fictions from a group of expert readers. A starting point was synthesising previous attempts to articulate notions of quality. The discussions with the readers suggest that the key qualities were substantive contribution, credibility, resonance, ambiguity and aesthetics; rich rigour and sincerity need also to be considered. Fiction has a place in educational research because it enables one to imagine vividly different possibilities, presents issues in an open-ended way, and is engaging.

ARTICLE HISTORY
Received 3 August 2021
Accepted 1 April 2022

KEYWORDS
Artificial intelligence; robots; fiction; research quality

Introduction

Selwyn (2019, p. iv) suggests the use of AI and robots could be 'one of the defining educational challenges of the next twenty years'. Yet much of the literature about AI in HE has been written by computer scientists without much reference to either pedagogy or ethics (Zawacki-Richter et al., 2019). This is such an important gap because AI and robots stir deep societal controversy. Jobin et al. (2019) identifies no fewer than 84 recent publications containing guidelines for ethical AI from international bodies, national governments, professional bodies and tech companies. Common concerns are around privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility and promotion of human values. A growing body of work from critical data studies raises inter-related concerns around datafication, commodification and surveillance linked to AI (e.g., Jones, 2019; Williamson & Eynon, 2020).

CONTACT Andrew Cox a.m.cox@sheffield.ac.uk Information School, University of Sheffield, Level 2, Regent Court, 211 Portobello, Sheffield S1 4DP, UK

Yet neither AI itself or the ethics of AI is easy to grasp. The current wave of excitement about AI is not related to a single technology, whose effects might be easy to delimit, but a bundle of technologies such as machine learning, natural language processing and voice recognition, combined with 'big data' and computing power. Indeed, rather than technologies, both AI and robots are ideas, with a long history and powerful and evolving public imaginaries.

There is clearly a need for a wider engagement with the issues raised by the use of AI and robots in HE beyond technologists, despite the complexity of the subject. One plausible way to promote debate is through the vehicle of fiction because it can instantiate the issues in an accessible, relatable way. Design fiction creates descriptions of imaginary technologies to provoke debate about societal issues, and thus offers a suitable approach (Blythe, 2017; Dalton et al., 2016). Yet the use of fiction in academic writing remains relatively rare. Therefore, there is a need to reflect on what quality in fiction as a research method is (Leavy, 2016). In this context, the objective of this article is to make an original contribution by reflecting on the criteria that define quality in such fiction-based research.

AI and robots in HE

To establish the context for the study, this section offers an overview of the literature on AI and robots in HE. It is fair to say that the majority of such literature focuses on learning, with distinct bodies of literature relating to specific classes of AIEd applications, such as Intelligent Tutoring Systems, Automated Writing Evaluation and Conversational agents (chatbots) (Holmes et al., 2019). Uses in student administration are also emerging. There is also a discrete literature around the role of AI on the smart campus (e.g., JISC, 2019). AI and robots are also likely to have an impact on research, e.g., through the robot scientist that autonomously conducts experiments on a mass scale and through the application of text mining to different forms of data and published literature (Jones et al., 2019). Little of the literature encompasses the gamut of these applications to reflect on the potential impact on HE as a whole. While most literature on using AI and robots in HE is about developing and evaluating systems (Zawacki-Richter et al., 2019), some key authors seek to imagine how AI can address the grand challenges of education in the twenty-first century (Holmes et al., 2019; Luckin et al., 2016). They believe that applying AI to education can address many of the inefficiencies and injustices of the current global education system such as ineffective assessment systems and educational inequality.

Counter-balancing these bodies of literature there is a longstanding tradition of a humanistic, critical response to applying technology in HE. It typically focuses on the fear that technology will dehumanise the learning experience. Selwyn's (2019) recent book summarises many of these familiar arguments by applying them to AI and robotics in an attempt to shift debate away from what is possible to what *should* be. Chiming with this is a growing body of work from critical data studies that challenges issues around bias, explainability and privacy in AI, some specifically in the educational field (e.g., Jones, 2019; Williamson & Eynon, 2020). Crawford and Joler (2018) reveal the environmental and societal impacts of seemingly 'virtual' technologies. Much of this work is distinctly dystopian about three potentially interlinked trends: the impact of datafication

and the metricisation of social life; the commodification of learning and the impact of Edtech; and surveillance and the promotion of neo-liberal ideologies through technologies. But not all authors from a critical tradition see AI as inherently hostile to human values. For example, Bayne's (2015) paper on 'Teacherbot' reflects on the design of a chatbot informed by critical pedagogy.

Another substantial body of literature around technologies in HE analyses the challenges of implementing them in practice as an organisational change management issue. In this literature, while technologies are seen as potentially beneficial, emphasis is placed on the barriers to adoption and implementation. In the case of AI, we can draw on Tsai et al.'s (2019) work which reveals why learning analytics have not yet found extensive practical application. Cultural issues figure heavily among these barriers. Given that AI is typically trained on big data, by extension it can be argued that the same barriers will probably apply to AI. This somewhat pragmatic literature articulates an important response to AI by recognising the barriers to its implementation.

Work bringing together the range of applications in HE and representing the contrasting responses is lacking. One vehicle to do so and to widen debate might be fiction.

Fiction as a means to explore the impact of new technologies

There are many genres of writing emerging around the current wave of AI and robots (Avin, 2019). This article explores the specific affordances of fiction in promoting understanding and debate about the use of AI and robots in HE. In some ways this is not a radical departure. Much of our expectation about future technologies have been set by fiction, e.g., science fiction (Dourish & Bell, 2014). Since the future is unknown any statement about the future is at some level a fiction: we are exposed to many fictions in corporate brochures. Further, it is fairly common for academic commentators on AI to use fictional scenarios to help present how they imagine technologies will develop in a concrete way (e.g., JISC, 2019; Luckin & Holmes, 2017; Selwyn et al., 2020). In terms of generating debate, fictions are accessible and engaging.

The gap between fiction and social science writing is not as great as often assumed and Leavy (2016) has charted historical connections between them. Both can generate knowledge about society and seem to be increasingly used by social scientists (Watson, 2022). Fiction can portray the complexity of social life and promote empathy and reflection (Leavy, 2016). Leavy (2016) suggests that fiction is particularly powerful for raising critical consciousness because of its ability to build empathy with those who have different experiences from ourselves. It also has the obvious advantage of making academic texts more accessible and engaging. Nevertheless, the value of fiction for educational research is still contested (Spindler, 2008).

In the specific context of AI, Avin (2019) goes so far as to say fictional treatments are over-abundant, although he is referring to those in the popular media, rather than academically sourced ones. His other criticisms arise from this but may be relevant to thinking about all forms of fiction. He suggests that they need to place too much stress on entertaining. He also suggests that in fiction there is a pressure to embody issues in a way that makes it difficult to represent more intangible issues, such as the workings of algorithms. He argues that such fictions are lacking in accountability, because the

feedback they get is limited to editors and readers. He also questions the diversity of voices expressed through these media.

In order to promote a more accountable form of fiction this article adopts the design fiction approach, which is rooted in information system design practices. At a very practical level, information systems design scenarios are often used to develop use cases and enable stakeholders to give input in articulating user requirements. Yet, as Nathan et al. (2007) comments, traditional design scenarios tend to represent the technology being used solely as intended by the designer. They also typically only include immediate contexts of use, rather than reflecting on the wider societal impacts of its pervasive use. In response, a growing body of work within Human–Computer Interaction studies adopts the use of speculative or critical narratives to destabilise assumptions through 'design fictions' (Blythe, 2017; Dalton et al., 2016), which create a speculative space in which to provoke questions about the desirability of a particular technology and the assumptions lying behind its design, but also to query our whole relationship to technologies. They offer an antidote to technological determinism and solutionism that appears in many presentations of technology futures. Fictions could be texts but also other media or embodied in material objects and prototypes, and are often produced collaboratively.

Fiction is increasingly recognised as a social science method (Leavy, 2016) and as one of a number of arts-based methods that can be used in educational research. Nevertheless, its use in an academic context remains unusual. There is therefore a need to think more about what quality in fiction-based research looks like. Thus, the purpose of the study was to work with a range of experts to reflect on what constitutes quality in a collection of fictions developed from a literature review.

Materials and methods

The fictions

A wide-ranging literature review on AI and robots in HE was conducted through a mixture of systematic and snowballing searches and including grey literature as well as academic texts (Cox, 2021). From the understanding gained, which is summarised in a previous section a collection of eight fictions was composed by the first author (amounting to a total of around 5000 words). The full text of the fictions can be downloaded from the University of Sheffield data repository, ORDA, at https://doi.org/10.15131/shef.data.12826076.v1. They are summarised in Table 1. Each imagined a somewhat different technological scenario (column 3): some, from a purely technical perspective, quite remote, some potentially realisable today. They were set across the different domains of higher education: teaching, administration and research (column 5). They were deliberately written in different styles (column 4). Despite being set in different storyworlds, the intention was that by virtue of being in the collection and through thematic interlinking, they could be read against each other, so revealing multiple layers to the issues presented. For example, the earlier fictions focus on the technologies in action in the classroom or campus. Fiction 5 refocuses on the role of commercial players and attendant issues of commodification of HE. Fiction 6 shifts attention to the organisational culture barriers to adopting AI. Fiction 7 reveals benefits and issues around datafication hinted at in earlier fictions. Fiction 8 poses issues around often hidden

Table 1. Summary of the fictions.

1. Title	2. Synopsis	3. Technologies involved	4. Genre	5. Area of application to HE	6. Central issue explored
Fiction 1: AIDan, the teaching assistant	Imagines a possible classroom of the future from a student's point of view	Intelligent tutoring systems, adaptive pedagogical agents, use of sensors to allow affective/embodied adaptivity	Traditional design scenario	Teaching	Optimistic view of AI as a solution of grand challenges in education through continuous evaluation, personalisation
Fiction 2: Footbotball	Imagines a future form of football which combines human and robotic players as a team	Robots	Soliloquy	Extra curricula activity	Potential for humans to have physical and emotional attachment to robots
Fiction 3: CriticalBot in conversation	An imaginary conversation between a group of students trying to do an assignment and a chatbot	Conversational agent	Dialogue	Teaching	AI that makes no pretence to be human yet makes a distinctive contribution to learning
Fiction 4: The intelligent campus app	Part of a day in the life of a student on a smart campus	Smart campus: wayfinding, nudging	Mundane, day in the life	Estates management/ Teaching	AI as mundane part of the infrastructure, nudging behaviour
Fiction 5: Research Management Suite TM	A suite of technologies that support all aspects of the research process and career	Text and Data Mining, auto summarisation, auto writing	Marketing and PR material	Research	Role of Edtech companies
Fiction 6: Verbatim minutes of University AI project steering committee: AI implementation phase 3	A discussion between university leaders relating to the cultural obstacles to implementing AI	Not defined	Meeting minutes	Any	Political and cultural barriers to adoption
Fiction 7: Dashboards	The speaker explains to a visitor how data from all sorts of sources is managed at the university level	Data mining, conversational agents	Soliloquy	Administration/Teaching	Potential of data to improve learning but also cause datafication
Fiction 8: Minnie, the AI admin assistant	A voice-based assistant answers all administrative queries yet its clumsiness is an annoyance. In the second part the protagonist reflects on the much wider societal and environmental impacts of AI	Conversational agents	Surreal, cyberpunk dystopia	Administration, Wider social infrastructure	Societal and environmental impacts of the technology

environmental and societal impacts of technology that again are not foregrounded in the earlier fictions, but could be seen as relevant to them. The collection sought to offer a balance of optimistic and pessimistic perspectives, and seek to pose the dilemmas without 'taking sides' too strongly.

Criteria of quality for fiction-based research

Leavy (2016) has proposed eight criteria of quality in fiction-based research, though she stresses that this is not intended to be an exhaustive or prescriptive list. This section reflects on these criteria in relation to Tracy's (2010) influential summary of eight criteria of quality in qualitative research, with a view to producing an initial set of criteria of quality for reflecting on the fictions produced in this research.

While it is not the first criterion on her list, Leavy (2016) proposes that a fiction must make a 'substantive contribution' to the topic it deals with. This equates to two elements in Tracy's schema: 'worthy topic' and 'significant contribution'. It seems reasonable to argue that this is the over-riding purpose of research. The other elements relate to how this is achieved.

Leavy (2016, p. 80) places stress on 'verisimilitude' in research fiction. People and settings must be presented 'realistically, truthfully and authentically'. Indeed, she regards this, rather than contribution, as the primary standard by which fictional research should be judged. The nearest equivalent in Tracy's (2010, p. 842) schema is the term 'credibility' which refers to the 'trustworthiness, verisimilitude and plausibility' of a work and is achieved through techniques such as thick description, crystallisation and multivocality. Tracy's wider term seems more expressive.

Leavy (2016) also places emphasis on 'empathy'. Creating empathy is a particular strength of fictional writing because it can draw on techniques such as writing from the point of view of a character, revealing their interior dialogue and rich characterisation, in a way that non-fiction does not usually do. Tracy (2010) refers to 'resonance', but it is clear from what she writes about this term that empathy and an emotional connection are a large component of her concept. 'Resonance' as the wider term might be better.

A particularly valuable quality of fiction for Leavy (2016) is its ability to create and exploit ambiguity. Rather than close down meaning, fiction opens up multiple interpretations. It can pose dilemmas in an open-ended way (Spindler, 2008). Presumably this is somewhat in tension with credibility and with coherence, which is another criterion for Leavy (2016). The content and form should be 'coherent': e.g., that the formal decisions about structure fit the topic. Similarly, Tracy (2010) refers to 'meaningful coherence' as a key element of good qualitative research.

One of the strengths of fiction as a method is the way it can draw on different genres and can privilege 'aesthetic' appeal (Leavy, 2016). As Caulley (2008) argues, academic writing is often dull because it is so conventional and factual and it could usefully borrow some of the techniques of journalistic creative non-fiction. Tracy (2010) does not identify this as a specific criterion, though aesthetic merit is referred to as a means to achieve resonance.

Leavy (2016) also suggests that good research fiction has a 'personal signature' because the author should be present in their writing as an expression of their commitment to the topic. This can be linked to what Tracy (2010) refers to the importance of 'sincerity'

which refers to being both transparent and reflexive, to be authentic. The transparency aspect seems important to research: At some level this is possible for research presented in fictional form, through descriptions of the writing process. However, as an imaginative process, the writing of fiction can be described in outline but the working of creative inspiration cannot be fully described. Sincerity is the preferred term here.

Tracy (2010) suggests two further criteria that do not have an equivalent in Leavy (2016). She uses the term 'rich rigour' to refer to the abundance of data or theory in use. This contrasts with the precision implied by rigour for positivist research. Fiction can be rigorous too in the sense it can be based on exactly the same process as other more familiar methods, e.g., through interviewing participants or potentially some form of systematic literature review.

A final criterion in Tracy (2010) that does not appear in Leavy (2016) is 'ethical'. The same must apply to research fictions, e.g., to protect confidentiality of any participants such as interviewees who participated in creating them. Often this may be precisely why fiction is the best way to capture experience without exposing the confidentiality of participants. More fundamentally, good research could be ethical because it has a normative stance and/or informs us about ethical issues relating to the topic at hand. It seems important to emphasise this aspect and follow Tracy (2010) in adopting it as a distinct criterion.

From this discussion it is proposed that the following could be nine potential criteria relevant for this study:

(1) Substantive contribution
(2) Credibility
(3) Resonance
(4) Ambiguity
(5) Coherence
(6) Aesthetics
(7) Sincerity
(8) Rich rigour
(9) Ethics

These should not be seen as entirely discrete criteria. Substantive contribution feels like an outcome of other elements. It is hard to disentangle aesthetic choices and their effect, e.g., in creating resonance. Also, some criteria are in tension with each other, such as ambiguity and coherence. Leavy (2016) refuses to claim that her list is comprehensive or prescriptive. In this spirit, the article seeks to present an insightful discussion of the criteria in use, without seeking to produce a reductive checklist.

Engaging collaborators

In order to explore these criteria in depth with the fiction collection produced by the first author in this research, five participants who are co-authors of the article were invited to participate in reading and discussing them:

(1) Dave Cameron (Information School) for his expertise in robotics.

(2) Alessandro Checco (Information School) for his expertise in AI and machine learning.
(3) Tim Herrick (School of Education) for his interest in pedagogy and technology enhanced learning. He also likes reading science fiction.
(4) Maria Mawson (University Library) for her interest and a professional services/library perspective and insights into managing institutional change.
(5) Richard Steadman-Jones (School of English) for his expertise in narrative writing.

This group were selected to represent a diverse range of relevant expertise. Those from a technical background could comment on the plausibility of the technologies, in so far as that was deemed relevant, while Tim could comment more on the plausibility of the learning dimensions. Richard could comment from a literary perspective; while Maria brought a more practical focus from the point of view of an academic library.

They participated, firstly, by writing extensive comments on the fictions; secondly, by participating in an individual interview discussion via VoIP; thirdly, by reviewing drafts of the current article. Andrew coordinated the process, including analysing the data from discussions to produce drafts of the article iteratively.

Results and discussion

Substantive contribution

The participants agreed that the fictions were a worthwhile contribution to an important topic. The rich engagement prompted by the fictions can be illustrated by some of the discussion around the relation between humans and AI and robots that they sparked (albeit that the two with subject expertise already knew more than Andrew about this). In fiction 1 the AI monitors learning behaviour with a view to personalising activities to learner state and mood. Maria challenged whether given the choice students would opt into the monitoring of their performance as imagined. Equally, Alessandro questioned how feasible it would be in reality to opt out, if a critical mass opted in. Tim reflected that there might be issues with the monitoring of activity to personalise learning in this way:

> Undoubtedly there are visible signs of low motivation, but (1) it strikes me they can be easily confused with other states of being (anxiety, fatigue, being hungover for example); and (2) their origins are likely to be obscure – am I lower in motivation because of the topic, the teaching situation, an incident with my flatmates last night, an ongoing concern about a relative?

Dave expressed the concern that if AI was too helpful in nudging a student to motivate them there would be an impact on self-determination. We learn from making our own mistakes. Alessandro expressed a similar concern that AI would have many unintended effects, both perhaps making us rely lazily on AI, but also through the unpredictable effects of constantly being under surveillance, particularly on the ability to express disagreement. He went on to ask if we would be replacing what could be accomplished by a simple human interaction (a chat to see if a person was OK) by an infrastructure of surveillance for metrics that are a poor measure of learning, with many unintended impacts.

In reflecting on productive relations between AI and humans, Fiction 3 emerged as representing the most positive portrayal of human/AI relationships, with its humorous tone and sense of 'playful resistance' (Alessandro). Participants appeared to like this best of the fictions. Dave commented: 'this is a really interesting piece of writing because it brings out the non-humanness of the bot while nevertheless suggesting that it has something valuable to contribute'. Tim wrote:

> The role of CriticalBot here is relatively passive and reflective – reminding the students what they should be doing, the resources open to them to support them and modelling the kinds of question that might be productive. This seems relatively benign, and quite attractive – it's task-focused, non-directive.

The irony might be that CriticalBot could teach criticality when it could not actually itself know anything in a true sense.

Thus, all the readers felt that the fictions had provided material for further reflection, even for those with a strong existing knowledge of the subject. Participants were asked to suggest further stories, partly as a way to explore how far the fictions had successfully captured all the main issues they thought there were. Both Tim and Dave thought the fictions erred on the side of representing general AI, echoing the human ability to reason across contexts, and felt more emphasis could have been given to narrow AI, which operates only on a very constrained or structured task. Dave in particular thought that automation was likely to be more in the background. However, on balance, the fictions appeared to instantiate many if not all the issues that participants felt they should raise, providing evidence that they made a substantive contribution.

Credibility

Credibility was important to how the group viewed the fictions, but this criterion had multiple dimensions. Naturally Alessandro and Dave, as experts on AI and robots, respectively, felt confident to evaluate the credibility of the technologies and the issues they raised. Alessandro had some questions about whether from a technical point of view some fictions were consistent. But Tim considered fiiction 6 about a research system based on AI credible 'because it seems a logical extension of the support and kind of market analysis that we have at the moment'. Similarly, Maria evaluated the fictions in terms of their applicability from a professional service perspective. Many of her comments connected the fictions to what was already happening, so credibility arose from what was being described as being an extension of a known trend. This points to a different form of credibility based not on technical possibility but on institutional realities and needs. Indeed, Dave liked the fiction as instantiating key societal issues around automation, but thought it was technically challenging.

Richard thought most of the fictions plausible both from his experiences with the technology and from a political or institutional perspective. He thought fiction 7 was the most remote because bringing together all the data sources would be so challenging particularly data sharing between schools and universities. So it was not the technical possibility that determined credibility, so much as the cultural barriers to sharing data judged by how limited such data transfer there is currently.

Thus credibility is multi-layered. It could imply that the technology described is possible now or feasible, or that technologies fulfilled known needs or fitted into current trends, that people behave in plausible ways or that recognisable institutional drivers are at work.

Resonance

There was a sense that the stories did resonate with all of this group of readers and that this was an important criterion for them. This seemed to relate to the way that the fictions imagined what living with a technology might be like on a day-to-day level. Maria enjoyed that the stories made the issues more real, that they 'focused on the people side of AI rather than the technology'. Stories allow us to connect emotionally in a way we would not necessarily have with academic writing. She said she got 'quite carried away thinking' about how the AI in fiction 1 might be extended to library-specific tasks. So she was prompted to think about how to apply a technological possibility to her work. This is an important affordance that fiction has. At the same time she connected to the sense of what the cultural and practical barriers to actually putting AI into effect would be.

The multivocality achieved by having a collection of stories written from different points of view seemed to be central for Tim. Dave, in contrast, was more struck by the thought that most of the stories could almost belong to the same storyworld. But all the participants made comments connecting fictions together seeing them as contrasting and connecting. This supports the value of a collection rather than a single fiction.

The participants were rather divided on fiction 6. One felt it rang 'so true' for its portrayal of the disciplinary divisions behind institutional politics that might be a barrier to AI. However, two others felt it was guilty of exploiting simplistic stereotypes about academic disciplines. Alessandro thought it was stereotyped to have engineers thinking in somehow mechanistic, even inhuman, ways, when they are actually creative and sociable. Equally, Richard thought the fiction stereotyped humanities scholars as resistant to technology and ignorant about data analysis, when many, such as linguists and economic historians, are quite the opposite. In both cases the troubling lack of nuance to the stereotype got in the way of their engaging with the fiction, though perhaps the issue of disciplinary difference in using AI is an important consideration. The discussion reinforced Leavy's (2016) point that empathy is promoted when fictions move beyond stereotypes.

Fiction 2 was not greatly commented on. There was a sense that it failed to resonate in a way that others did seem to do better. This might have been because it was more futuristic. It was less closely related to HE, at least from a staff perspective, because it was about a student's relationship with a team of robotic football players. Possibly the presentation through the eyes of a young sportsman made it hard to relate to this audience.

Ambiguity

The fictions did exploit ambiguity in a number of ways. Firstly, they were all short fragments. For example, the conversation with CriticalBot in fiction 3 is self-evidently part of

a longer conversation. Fiction 6 is explicitly an extract of meeting minutes. Ambiguity was a strategy to enhance engagement because it left room for imaginative completion by the reader. Secondly, the fictions were often quite ambiguous about the technologies actually in use. Thirdly, and most fundamentally, the fictions essentially posed dilemmas without seeking to resolve them.

> The fictions prompt some interesting questions there for people to do their own homework on. The fictions don't necessarily have a stance. I think they are well written in that regard. It enables people to take stances. (Dave)

Aesthetics

Aesthetics was another aspect that seemed important to how all the participants responded to the fictions. The group simply enjoyed reading them. As Maria pointed out: people respond to storytelling as a form. 'I really enjoyed reading them. It made me stop and think.'

> The stories are neat. Some are really fun. I like reading them. (Dave)

Dave was taken with this approach to writing where 'being right' is not the aim, as he put it. He thought that needing to be right made most scientific writing conservative. He commented that it

> Creates a levity and lightness that you can use to say important things with. And in that more open frame of mind, start thinking.

Alessandro also found them stimulating a different way of thinking because they were not focused on a narrowly defined problem, as was more usual in the research process. 'It's a lot easier to read and is engaging … humorous or almost poetic and resonant. It makes it easier to think and connect to' compared to if the same material were presented in a literature review.

Richard particularly liked the fictions in conversational form, especially fiction 3, because it gave you a direct sense of how the AI might actually interact with students. Alessandro commented 'I really loved the humorous tone and the back and forth' in this fiction. But for Richard, fiction 1 was a bit more like a student quote from a brochure. He also suggested that there could be more stylistic variation, e.g., there was no story written in a conventional third person narrative and while fiction 8 closed in on something akin to poetry, this could have been taken further.

Thus, the aesthetics of the rich variety of styles within fiction seemed to be important to how the texts were appreciated, and dialogue with humour was being particularly valued. The collection was successful in these terms, but could have drawn on an even wider range of genres.

Other criteria

Substantive contribution, credibility, resonance, ambiguity and aesthetics seemed central to the discussions around this collection. It is worth commenting briefly on some of the other criteria that were developed above.

The stories were not intended to be coherent in the simple sense of fitting into one narrative or even one storyworld, rather they intentionally had a complex relation to each other, with themes recurring in different contexts. For most participants this seemed to work well. For example, Richard described the fictions as 'wide ranging' in terms of the different contexts investigated (teaching, campus, research), different types of AI or robots and different seeming relations between humans. This led you to reflect 'on which a university might choose', he said. This is important because it reflected the intention to show that there is not one pre-determined future for AI and robots in HE. Maria also felt any one of the worlds described could exist and that they were 'not necessarily "interdependent"' yet that they were 'not massively far apart' either. However, perhaps because he was expert on the technologies and had a strong feel for what was technologically possible and what was not, Alessandro found it a problem that the AI was too different in the different fictions.

Interestingly participants did not really show any curiosity about how the fictions were produced – their 'rich rigour' – perhaps because the production of fiction is understood to be 'creative' in a way that academic texts are not expected to be. 'Sincerity' did not really arise in the discussions, either, perhaps because the fictions intentionally do not themselves have a particular stance. As is apparent from earlier discussions the main response was to think about the ethical dilemmas around AI and robot technologies. So, in this sense 'ethics' was central to the fictions but not in the sense of the ethics of the process of producing them.

Conclusion

Fiction, particularly this collection of short, inter-related pieces proved to be a successful medium to engage a group of university staff, with very different backgrounds and different levels of expertise, in the debates around AI and robots in HE. It enabled them to imagine what living in a world with the technology in daily use might be like. Being open-ended to interpretation about whether the future imagined was good or bad it enabled them to think in an open way. The fictions achieved a breadth of context and perspective hard to accomplish in academic writing. They were simply an engaging form of reading. There were some missing aspects of the issues (some of which were addressed in a larger collection of fictions available at https://doi.org/10.15131/shef.data.12826076.v1).

In doing this the fictions drew on multiple levels of *credibility*, as much by reflecting organisational realities and current trends, as through how they relate to technical possibilities. The problem of embodiment raised by Avin (2019) did not seem to arise, because the readers' responses showed that it was possible to represent the issues around datafication quite successfully, even though they are potentially rather abstract. Most of the fictions *resonated* with this set of readers, especially through the dialogue and humour, but less successfully where the application was less relevant to them as staff, as in fiction 2, or where some behaviour was seen as stereotypical, as in fiction 6. Much of their value arose from how they helped envisage what it might be like to be 'enmeshed' in a world where the technologies were in wide use (Selwyn et al., 2020, p. 100). *Ambiguity* was also essential to the quality of the fictions in engaging the readers with the educational dilemmas, without themselves taking a position (Spindler,

2008). *Aesthetics* were important in accomplishing this engagement. The fictions were found to be enjoyable to read without this undercutting their ability to raise complex and challenging issues or their credibility, in a contrast to Avin's (2019) claim that entertainment is achieved at the cost of truth. The light tone helped to prompt different types of thought.

The *rich rigour* of the process of writing the fictions and the *sincerity* of intention did not seem a central issue in the discussions but they must be important in response to Avin's (2019) concern about the lack of accountability in popular AI fictions. His point prompts us to think about ways to explain how the fictions were developed in rigorous ways. The process discussed in this article supplies this.

Thus, the argument of this article is that fiction can make a distinctive and *substantive contribution* to understanding the issues around AI and robots in HE. The approach suggested of writing a carefully balanced collection of fictions set in different storyworlds is a particular contribution of this research. The article has also contributed to discussion about what criteria should be applied to appreciating quality in research presented in this form. The synthesised set of criteria is useful to structure thought about how to evaluate such fictions. There is no intent to produce a prescriptive checklist, however.

The article has shown that the fictions produced are of value with staff working in academia in a range of disciplines and roles. As the article is limited to focusing on a small collection of fictions read by a small number of readers, the natural next step for the research is to extend the discussion to a wider set of groups with a stake in the nature of education, particularly students. One of Avin's (2019) critiques of fiction is that it lacks diversity. Indeed, the authors are conscious of the need to work harder to escape the WEIRD bias identified by Blanchard (2015) in AIEd literature. Developing the approach of getting students to write their own stories (Rapp, 2020) would seem to both expand the audience and address this concern with diversity. Rapp (2020) has some suggestions about scaffolding the process of authoring fictions for those unused to writing fiction and which would help make the approach scalable to a wide audience of HE students and staff. Building fiction writing into design and implementation processes would help to engage students and staff more deeply and probably help developers produce AI applications that better fitted the cultural context of education.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The project was funded by the Society for Research into Higher Education – Research Scoping Award -SA1906.

ORCID

Andrew Cox http://orcid.org/0000-0002-2587-245X
Alessandro Checco http://orcid.org/0000-0002-0981-3409
Tim Herrick http://orcid.org/0000-0002-4586-6559

References

Avin, S. (2019). Exploring artificial intelligence futures. *Journal of AI Humanities*, *2*, 170–194. https://doi.org/10.17863/CAM.35812

Bayne, S. (2015). Teacherbot: Interventions in automated teaching. *Teaching in Higher Education*, *20*(*4*), 455–467. https://doi.org/10.1080/13562517.2015.1020783

Blanchard, E. G. (2015). Socio-cultural imbalances in AIED research: Investigations, implications and opportunities. *International Journal of Artificial Intelligence in Education*, *25*(*2*), 204–228. https://doi.org/10.1007/s40593-014-0027-7

Blythe, M. (2017). Research fiction: Storytelling, plot and design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5400–5411).

Caulley, D. N. (2008). Making qualitative research reports less boring: The techniques of writing creative nonfiction. *Qualitative Inquiry*, *14*(*3*), 424–449. https://doi.org/10.1177/2F1077800407311961

Cox, A. M. (2021). Exploring the impact of Artificial Intelligence and robots on Higher Education through literature-based design fictions. *International Journal of Educational Technology in Higher Education*, *18*(3), https://doi.org/10.1186/s41239-020-00237-8

Crawford, K., & Joler, V. (2018). *Anatomy of an AI system*. https://anatomyof.ai/

Dalton, N. S., Moreau, R., & Adams, R. K. (2016). Resistance is fertile: Design fictions in dystopian worlds. In Proceedings of the 2016 *CHI conference extended abstracts on human factors in computing systems* (pp. 365–374).

Dourish, P., & Bell, G. (2014). 'Resistance is futile': Reading science fiction alongside ubiquitous computing. *Personal and Ubiquitous Computing*, *18*(*4*), 769–778. https://doi.org/10.1007/s00779-013-0678-7

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education. *Boston, MA: The Center for Curriculum redesign*.

JISC. (2019). *The intelligent campus guide*. https://www.jisc.ac.uk/rd/projects/intelligent-campus

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(*9*), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jones, E., Kalantery, N., & Glover, B. (2019). *Research 4.0 Interim Report*. Demos. https://demos.co.uk/project/research-4-0-research-in-the-age-of-automation/

Jones, K. (2019). 'Just because you can doesn't mean you should': Practitioner perceptions of learning analytics ethics. *Portal*, *19*(*3*), 407–428. https://doi.org/10.1353/pla.2019.0025

Leavy, P. (2016). *Fiction as research practice: Short stories, novellas, and novels*. Routledge.

Luckin, R., & Holmes, W. (2017). A.I. is the new T.A. in the Classroom. https://howwegettonext.com/a-i-is-the-new-t-a-in-the-classroom-dedbe5b99e9e

Luckin, R., Holmes, W., Griffiths, M., & Pearson, L. (2016). Intelligence unleashed an argument for AI in education.

Nathan, L. P., Klasnja, P. V., & Friedman, B. (2007). Value scenarios: A technique for envisioning systemic effects of new technologies. In *CHI'07 extended abstracts on Human factors in computing systems* (pp. 2585–2590).

Rapp, A. (2020). Design fictions for learning: A method for supporting students in reflecting on technology in human-computer interaction courses. *ComputerS & Education*, *145*, https://doi.org/10.1016/j.compedu.2019.103725

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. John Wiley.

Selwyn, N., Pangrazio, L., Nemorin, S., & Perrotta, C. (2020). What might the school of 2030 be like? An exercise in social science fiction. *Learning, Media and Technology*, *45*(1), 90–106. https://doi.org/10.1080/17439884.2020.1694944

Spindler, J. (2008). Fictional writing, educational research and professional learning. *International Journal of Research & Method in Education*, *31*(*1*), 19–30. https://doi.org/10.1080/17437270801919867

Tracy, S. J. (2010). Qualitative quality: Eight 'big-tent' criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(*10*), 837–851. https://doi.org/10.1177/2F1077800410383121

Tsai, Y. S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges and opportunities. *British Journal of Educational Technology*, *50*(*6*), 2839–2854. https://doi.org/10.1111/bjet.12846

Watson, A. (2022). Writing sociological fiction. *Qualitative Research*, *22*(3), 337–352. https://doi.org/10.1177/1468794120985677

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(*3*), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Zawacki-Richter, O., Marín, V., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0