Article

new media & society
1–20

DOI: 10.1177/14614448211012794
journals.sagepub.com/home/nms
SAGE

What are the risks of Virtual Reality data? Learning Analytics, Algorithmic Bias and a Fantasy of Perfect Data

Marcus Carter
The University of Sydney, Australia

Ben Egliston
Queensland University of Technology, Australia

Abstract

Virtual reality (VR) is an emerging technology with the potential to extract significantly more data about learners and the learning process. In this article, we present an analysis of how VR education technology companies frame, use and analyse this data. We found both an expansion and acceleration of what data are being collected about learners and how these data are being mobilised in potentially discriminatory and problematic ways. Beyond providing evidence for how VR represents an intensification of the datafication of education, we discuss three interrelated critical issues that are specific to VR: the fantasy that VR data is 'perfect', the datafication of soft-skills training, and the commercialisation and commodification of VR data. In the context of the issues identified, we caution the unregulated and uncritical application of learning analytics to the data that are collected from VR training.

Keywords

Algorithmic bias, automated decision-making, data ethics, datafication, learning analytics, virtual reality

Introduction

The digitisation of education and learning has been accompanied by acceleration in the collection and analysis of educational data. Learning analytics – 'the practice of

Corresponding author:
Marcus Carter, Department of Media and Communication, The University of Sydney, Camperdown, NSW 2006, Australia.
Email: Marcus.carter@sydney.edu.au

developing actionable insights through the collection, analysis and reporting of data about learners and their contexts' (Society for Learning Analytics Research (SoLAR), nd) – promises to leverage these data to enhance learning experiences, maximise student success, improve the performance of teachers and optimise institutional outcomes. Yet the rapid deployment of these technologies outpaces the research verifying their efficacy (Corrin et al., 2019), and many scholars are growing increasingly critical of the 'datafication' of education and its surveillant and neoliberal dimensions, the quantification and automation of education, and the potential for algorithmic bias and discrimination (see Buckingham Shum and Luckin, 2019; Selwyn, 2020; Slade and Prinsloo, 2013; Williamson et al., 2020).

Virtual reality (VR) is an emerging technology that presents a myriad of political and ethical issues in the context of the increasing datafication of society. Nascent discussions around VR – in education and more broadly – typically focus on the promising potential of VR for immersion and simulation (see Daniela, 2020; Liu et al., 2017), but in the context of what we observe as an emerging tradition of 'critical VR studies' (Bollmer, 2017; Egliston and Carter, 2020; Evans, 2019; Golding, 2019; Harley, 2019; LaRocco, 2020; Saker and Frith, 2019, 2020; Wallis and Ross, 2020), here we focus our attention on the data extraction capabilities of VR. As we have recently argued elsewhere, VR devices represent one of 'the most data-extractive digital sensors we're likely to invite into our homes in the next decade' (Carter and Egliston, 2020), yet we are not aware of any critical discussions of the application of learning analytics to VR in higher education or elsewhere, or the necessary speculation (or anticipation, see Stilgoe et al., 2013) regarding the dangers that VR may present in further accelerating the datafication of education and learning.

Consequently, we set out to study how emerging VR education and learning technology companies discursively frame, use and analyse the data that are generated via the use of a VR device. Drawing on approaches established in studies of the political economy of communication (see Corrigan, 2018), and critical technocultural discourse analysis (Brock, 2018), we analysed four VR tech start-ups. Cognisant of the issues previously identified with learning analytics in the context of online learning management systems, we found both an expansion and acceleration of what data are being collected about learners and how these data are being mobilised to compare learners and assess teachers. In our results, we identify the immediate potential for data-related harms in the use of VR education tools in enterprise for hiring and promotion, and problematic claims about the predictive potential for VR data. Beyond providing evidence for how VR represents an intensification of the datafication of education, we discuss three interrelated critical issues specific to VR that are of broader of concern to researchers, practitioners and policymakers:

1. *A fantasy of perfect data*: the way in which VR data is framed as capturing the total experience of learning is – in line with what Mark Andrejevic (2019) calls the 'fantasy of automation' (p. 114) – a fantasy of perfect data, exacerbating issues with use and application of learning analytics.
2. *Quantifying the qualitative*: the expansion of VR learning analytics into previously unquantified soft skills like empathy relies on this fantasy, and highlights

how VR education technologies are being shaped to fit into neoliberal ‘metric power’ (via Beer, 2016) – the impulse to measure, compare and classify.

3. *Virtual reality, real data empires*: in line with other domains of big data, VR for education will increasingly become a new frontier for data capitalism (following Sadowski, 2020; Zuboff, 2019) via the application of machine learning and automation to commodified VR data.

Taken together, our findings caution the unregulated and uncritical adoption of learning analytics to VR and point towards an emerging frontier in the (mis)use of new media technologies. As VR – along with the similarly data-extractive Augmented Reality (AR) – is anticipated to proliferate widely over the next decade, it is crucial that we pay sustained attention to how all VR data is collected and used, and resist the extraction and commodification of VR data. Separate from the exciting potential for VR and AR as an educational tool for visualising and immersing learners, we ask – as Andrejevic and Selwyn (2020) ask of facial recognition technology in education – if the ‘added value’ or ‘gained efficiencies’ of VR learning analytics outweigh the potential for harm through algorithmic bias, discrimination and surveillance. Where we anticipate VR expanding from learning into a platform or the performance of labour, the potential impacts on society are far-reaching.

Related work: datafication, learning analytics and automated decision-making

Present scholarly work on the ethics of learning data and analytics mostly focuses on the higher education sector, as one of the few institutions large enough – and with enough data – motivated to pursue the proposed benefits of learning analytics. Drawing on prior extensive critiques of datafication and learning analytics by other authors (see Buckingham Shum and Luckin, 2019; Selwyn, 2020; Slade and Prinsloo, 2013; Williamson et al., 2020), here we describe some of the key issues and concepts that informed our research and analysis.

First, and perhaps foremostly, data are not a neutral thing. Although often framed as being ‘objective, neutral and free of bias’, data – as Rob Kitchin (2014a) argues – ‘are not simply natural and essential elements that are abstracted from the world in neutral and objective ways’ (p. 135). Data are something that are collected for a specific purpose, making visible something that was previously concealed, constructing a new view about the world (Jasanoff, 2017). What is measured, and what isn’t measured, are mechanisms for constructing – and forgetting – knowledge (Bowker and Star, 2000). As Buckingham and Ferguson (2012) summarise, ‘if a phenomenon is not visible within a classification scheme, it is systematically erased’ (p. 18). Similarly, the process of making sense of data also casts how it is interpreted (Kitchin, 2014a: 136).

Within the context of learners and learning, this has the potential for negative impacts. Consider, for example, the data collected about students through a Blackboard Learning Management System: clicks and page views. Blackboard’s Data & Analytics website claims that these data can identify motivation and persistence to learn, and help ‘identify and overcome barriers to student success’ while also ‘optimizing student outcomes’.

Students with irregular and infrequent page views may – based on historical data of previous years – be categorised as 'at risk', creating a new view about members of a student cohort. The positive promise of this kind of technology is that this student might subsequently receive the support and motivation they require. Yet, a student with caring responsibilities, or no Internet access at home, is inevitably going to have patterns of engagement that differ from the typical student without (see also Gilliard and Culik, 2016 on how acceptable use policies work as a form of 'digital redlining'). Categorising this student at risk is based on an often unacknowledged political claim about what 'good learning' might look like, and the design of learning analytics systems that omit such realities are a similarly political claim about what 'good educators' should know and consider in their teaching. Data-driven learning analytics inherently work to classify students, which can define how they are viewed by teachers and how students understand themselves as learners (Brown, 2020; Williamson et al., 2020). The danger is that this student – who may learn efficiently, even if erratically – being classified as 'at risk' may affect how the teacher engages with them, demotivate the student and encourage them to drop the class.

Anna Lauren Hoffman (2018) coins the term 'data violence' to describe these types of (often, but not always, inadvertent) harm done by data systems as the result of implicit and explicit choices in data and engineering. Engaging with Hoffman, Os Keyes (2019) points out that the solution to these shortcomings in data science is often problematically to expand the scope of data collection. Keyes (2019) argues that data science is 'the inhumane reduction of humanity down to what can be counted' and, as currently constituted, 'responds to critique only by expanding the degree to which it surveils us'. Data systems do more than just reflect problematic social attitudes, but 'reinforce and amplify them' (Hoffman, 2018), typically causing most harm to marginalised populations that fall outside of the universalised and historically couched data-view of what people are, or at least, the ways in which people are machine readable. In a related critique, Shea Swauger (2020) argues that algorithmic test proctoring engages in the 'Eugenic Gaze' (via Davis, 1995) in the ways that it draws on data collection and automated decision-making (ADM) to commit data violence by 'measuring student's bodies and behavior [machine learning and facial recognition software], defining what bodies and behaviors are associated with the ideal student [cisgender, White, able-bodied, neurotypical, male, non-parent, non-caretaker, etc.], attempts to reform students who deviate from the ideal student [flagging them as suspicious], or exclude them from the community [academic misconduct investigations which can lead to expulsion]'. As Slade & Prinsloo (2013) points out, there are 'inherently unequal power relations in the use of data generated by students', which exacerbates these issues. More than just being 'not neutral' then, data and its inherent classifying, are actively political.

All of this is occurring in the context of what Ben Williamson, Sian Bayne and Suellen Shay call the datafication of education – the rendering of the social and natural world of education and learning in a machine-readable format – which they contextualise as something integrally connected to processes of neoliberalisation that is impacting the entire higher education sector (Williamson et al., 2020). This critique is particularly relevant for our study of how VR learning analytics – developed by venture funded technology start-ups – are being are being discursively framed, adopted and employed in the neoliberal contexts of enterprise To very briefly summarise, Williamson et al. (2020)

draw on David Beer's (2016) concept of 'metric power' – that datafication sets limits on what is known, and knowable, which influences how people and institutions behave – to argue that the datafication of higher education is integrally part of the ongoing marketisation of the sector into 'a market in which institutions, staff and students are all positioned competitively, with measurement techniques required to assess, compare and rank their various performances' (Williamson et al., 2020: 354). In this way, neoliberal metric power changes the way that higher education operates, quantifying and rendering comparable and competitive every aspect of teaching and research (producing the contemporary 'toxic university, Smyth, 2017). The relevant point here for our analysis is that data have both micro and macro impacts: the decision to centre data impacts how people, and organisations, act and come to be.

The issues with data in learning analytic systems are exacerbated by the use of algorithms, because 'the algorithms used to process the data are reimbued with particular values and contextualised within a particular scientific approach' (Kitchin, 2014a: 135). Approaches that employ 'Machine Learning' and 'Artificial Intelligence' are using sets of algorithms to draw on big data sets to draw connections and correlations otherwise unseen, or unsought, by human users. Sun-ha Hong (2020) describe these predictions as *fabrications*, a concept that helps point out how processes of datafication and analysis solidify approximations into something endowed with authority and credibility. Nonetheless, algorithms are increasingly becoming part of decision-making, despite many notable examples of algorithms introducing bias, such as racially biased risk assessments used in criminal sentencing (Angwin et al., 2016) and socioeconomically biased exam result predictions used in England as a COVID-19 response (Katwala, 2020). One reason why this happens is because training algorithms on historical data reinforces long-standing structural inequalities in societies (Crawford and Paglen, 2019). This is also the case in learning analytics, where the process of rendering students readable for algorithmic classification 'reinforces and reproduces historical inequalities' (Williamson et al., 2020). Corrin et al. (2019: 8) also conclude that 'vulnerable students, such as those from a low SES background, might be stereotyped as a result of algorithms and learning analytics applications'. One of the underlying and integral issues of algorithms – in education and society more broadly – is that 'they rely on correlational patterns rather than casual narratives' to make meaning from data (Andrejevic, 2019: 38), meaning that algorithms do not explain, just predict, and with increasingly complex data systems (like VR, for instance) the basis for these predictions is often inaccessible and incomprehensible (Hong, 2020).

These issues with algorithmic bias are compounded when combined with automated decision-making (ADM): the use of algorithms not just for prediction and classification, but to make decisions without human involvement. The example of ADM that we are perhaps most familiar with is Facebook's social media news feed, that automatically determines – based on the view Facebook's platform has of us based on our clicks, views, location data and Internet habits – what news we encounter on a daily basis. In his book *Automated Media*, Mark Andrejevic (2019) makes the point that personalised news is not necessarily bad, but in practice – where it has 'arrived on the back of commercially owned and controlled platforms to service their advertising and marketing imperatives' (p. 28) – it has served to exacerbate political divides (Statt, 2020), provide a feeding ground for conspiracy theories and QAnon alternate realities (Warzel, 2020), and degrade democratic systems in the service of more time spent on Facebook's platform. Andrejevic's (2019) point here is that 'the choice to implement automation within the existing socio-economic context carries with it a set of built-in tendencies that have important societal consequences' (p. 26). To apply the argument presented by Keyes (2019) on the inhumanity of data science, ADM is inhumane because it implies the deduction of the human from parts of the decision-making process, with very real consequence. In the context of learning analytics, ADM presents the deeply problematic incorporation of algorithmic bias – based on inaccessible, incomprehensible, and thus unchallengeable fabrications – into learning decisions.

Our approach

Consequently, our research applied a discourse analysis approach (informed by Brock's, 2018 critical technocultural analytic techniques) to understand how four VR focused educational technology companies used, framed and discursively constructed the potential for VR data and learning analytics. Our focus in this study was the traces left by these emerging companies online – their websites, publications, webinars, press-releases, media interviews, podcasts, LinkedIn posts, patents, job advertisements, privacy policies and so on. Our review identified between 100 and 250 text sources for each company, and 12 hours of video and audio content. This follows our earlier work (Egliston and Carter, 2020), where we similarly analysed material from the *Oculus Developer Conferences*, inspired by studies of technology developer conferences (Dalton, 2015; Liao, 2018), and the more recent focused analysis of the *Oculus Best Practices Guide* by LaRocco (2020), and of the education company Pearson by Williamson (2020).

Our approach to handling the empirical material draws methods established in political economy of communication research that utilise 'trade publications' to examine media industries (see Corrigan, 2018; Wasko, 2004; Wilken et al., 2019), and the complementary techniques, via Corrigan (2018), of 'burrowing down' into the fine-print and 'listening in' on the discourses about these practices to immerse ourselves in the data. Corrigan (2018) draws the comparison between this approach and the immersive work that ethnographers do, arguing that 'with sufficient immersion, literal and interpretive meanings can crystalize, while inauthentic, unrepresentative, and untrustworthy documents are more easily spotted and scrutinized' (p. 2764). Rowan Wilken et al. (2019) – who recently fruitfully applied a political economy of communication approach to dating apps – demonstrates that this kind of approach can make significant contributions to software/media studies via its capacity to consider the role of economic logics, different stakeholders, and the role of data in influencing the developments of new technologies (see also Wilken and Sinclair, 2009). Consequently, our collection and review of the empirical material also involved a process of note-taking and drafting of analytic memos familiar to ethnographic work. These memos then underwent discourse analysis through a process of open and axial coding (Brock, 2018), alongside the text, video and audio sources that concerned VR and data (broadly conceived). Initial themes were identified by Carter, and a process of review and discussion with Egliston – sensitised by a review of the critical cultural and social issues previously identified in our literature review – identified the three primary themes that structure the results of the present article.

Our discourse analysis included as wide a range of materials as we could find, and was attentive to the varied imagined audiences of enterprise customers, venture capitalist investors and prospective employees of each text. Importantly, and drawing from existing critical discourse analyses of 'data imaginaries' (Beer, 2018), the present work focuses on how these companies discursively framed themselves and their products as necessary solutions to current and future problems, interrogating the assumptions implicit in these discursive framings. An initial review of VR focused on education technology start-ups identified 21 VR education companies. Our study focused on the following four companies:

> STRIVR is a US 'Immersive Learning' company founded by Derek Belch in 2015, based on research conducted with the Stanford University Virtual Human Interaction Lab [led by prominent VR researcher and STRIVR co-founder Jeremy Bailenson]. Initially focused on using VR to train athletes, STRIVR is now one of the largest VR education technology companies with 22,000 headsets and over 6 million VR training sessions deployed, principally via its partnership with Walmart. STRIVR has ~140 employees and has raised $53m in venture capital funding following a $30m Series B round in March 2020.
>
> TaleSpin is a US/Netherlands 'XR-Based Learning' company founded by Stephen Fromkin and Kyle Jackson in 2015 with a dual focus on training for the insurance industry via a partnership with Farmers Insurance, and highly real simulations and soft skills training for managers and business leaders. TaleSpin has ~90 employees and has raised $20.2m in venture capital, following a $15m Series B round in March 2020.
>
> Mursion is a US-based 'provider of VR simulations for workplace training' founded by Arjun Nagendran in 2014. Using both 2D screen and head-mounted VR, Mursion uses 'digital puppetry' with real human actors controlling virtual avatars for soft-skills training in management, sales, customer service and higher education contexts. Mursion has ~85 employees and has raised $15m, following an $8m series A round in May 2019.
>
> Immerse.io is a UK-based 'Enterprise VR' company founded by Justin Parry, with a focus on its Virtual Enterprise Platform for VR training at scale within an organisation. Immerse.io is partnered with GP Strategies, a global training consultancy, and has prominent partners including Shell, DHL and BP. Immerse has ~35 employees and has raised an undisclosed amount.

Of the 21 identified companies, 9 were small, newly conceived and had little material online available for analysis. STRIVR and TaleSpin were chosen for focused analysis because they are the largest VR education companies in the market, with successful partnerships with established companies. Mursion was included in our analysis because they utilised a different technological approach – 'digital puppetry' rather than AI characters – and Immersion.io was selected because of its explicit focus on enterprise. Together, these four companies showcase the range of underlying technologies, approaches and the breadth of different market focuses in the contemporary VR education market. In the subsequent three sections, we present our findings via a focus on the three primary issues that we believe are specifically pertinent to VR in education and the application of learning analytics to VR.

Results

A fantasy of perfect data

A material and consequential difference between VR learning data and web-based learning data is the fantasy of perfect data. By this, we refer to the tendency of the VR companies we studied to describe the data being collected about VR users as being complete, having no gaps. For instance, TaleSpin's Kyle Jackson praises the fact that 'we can measure anything, from your sentiment to your gaze to what you said and how you said it' (Press Interview, 2019). Immerse.io co-founder Justin Parry describes VR as 'fundamentally different' from other learning mediums, because they can 'record absolutely everything that user did' with '30 data points per second' (Webinar, 2020), and STRIVR describes VR as providing the 'next generation of data' that will provide 'insights about proficiency never before captured by traditional learning methods' (STRIVR Homepage, 2020). To support these claims, VR providers show how VR training sessions can be recalled and replayed, and point to the breadth of possible sensors that can further enhance the 'data picture'; insights into cognition via eye-gaze, physiological data such as heart rate and stress, and analyses of emotive state using facial recognition and brain-computer interfaces.

The way that VR data is framed as capturing the total experience of the learning is – in line with what Mark Andrejevic (2019) calls the 'fantasy of automation' (p. 114) – a fantasy of perfect data, exacerbating issues with use and application of learning analytics (also see Beer, 2018). In discussing this in the context of automation, Andrejevic offers the concept of framelessness to describe the contemporary drive of big tech to create a digital 1:1 map of the entire world (see Kelly, 2019). This global platform, powered by cameras and sensors *everywhere*, is intended to serve the purpose of 'reflecting the world back to itself in machine readable form' (Andrejevic, 2019: 114). The totalising ambition of automation relies on this 'post-subjective perspective of the view from everywhere' (Andrejevic, 2019: 115), but (despite pervasive public understandings to the contrary) such an apparatus is both flawed – objectiveness is inaccessible – and practically impossible, while also describing a deeply surveillant dystopia.

VR education technology providers rely on this same fantasy, with concordant consequences. Claiming, as Immerse.io does, that VR data 'capture every detail' (Immerse.io Homepage, 2020) presupposes that the representation that they capture of the learning experience is frameless – that it is a 'purely objective representation that leaves nothing out' (Andrejevic, 2019: 115). This is of course enabled by the wider data imaginary and perception of data's veracity, but also the idea of VR as being psychologically similar to 'real' experiences. Mursion claims that its 'simulations achieve the realism needed to deliver measurable, high-impact results' (Mursion Website, 2020). STRIVR's 'science resources' webpage claims that VR simulations 'activate the same neural pathways in the brain' as real scenarios, and in a webinar on the Future of Work their Chief Science Officer describes how VR is a:

> more valid reflection of their world behaviors. So, sure, it's all choice data, but it's more meaningful choice data when it's done in an immersive training environment like VR because it's more reflective of their choice the choices they would make in the real world and then this is just an example of some of the ways that we can look at performance data.

This comment reflects what Rob Kitchin (2014b) has identified as a new kind of empiricist epistemology, that truth can be inducted via big data projects. This provides a rhetorical force behind the fabrications that companies – both the developers and the clients – can make about the quality, veracity, accuracy and predictive power of their analytics. In doing so, VR learning analytics appears to be increasingly incorporated into forms of ADM.

The most prominent existing example of ADM that we found is in STRIVRs partnership with Walmart. With 45 training modules and 17,000 headsets in 4700 locations, STRIVR VR provides Walmart the opportunity to simulate events that would be difficult to run as physical training scenarios (like a Black Friday shopping crowd), learn how to use new technology before it is installed, and for soft-skills training such customer service, empathy and dealing with difficult conversations. These examples make sense as an application of VR as an educational technology in enterprise. There are clear logistical benefits – an infrastructure of VR headsets means training can be quickly created and distributed across its stores at scale, circumventing the labour and travel costs and limitations of human teaching staff – in addition to the capacity for VR to create novel learning experiences based on its realistic simulations and videos.

Our concern though is with the application of learning analytics to these training scenarios and the introduction of new forms of algorithmic bias via their use to automate parts of decision-making. The datafication of workplace decision-making (particularly to do with hiring, promotion) is an area that has been fraught with debate and critique (Kim, 2017; Rhagavan et al., 2019; Rosenblat, 2018; Sanchez- Mondero et al., 2019). Walmart specifically has a long and well-documented history of discrimination surrounding promotion of its service workers (see Dukes vs. Walmart Stores, 2011), union-busting, and a predatory culture of worker control and coercion based on exploiting precarity and insecurity (see Haiven, 2013). In a blog post, Senior Vice President for Associate Experience Drew Holler describes how Walmart is using VR in the hiring process:

> Rachel and a team of technology and business leaders developed a skills-based assessment that uses virtual reality to simulate everyday obstacles. Once a candidate completes a 15-minute assessment, leaders use the results to help them remove subjectivity and unconscious bias from the selection process. This solution enables a people-led, tech-empowered way of working. (Holler, 2019)

In related news coverage (Tuchscherer, 2019), Holler is quoted as saying that the VR training is as expansive as the knowledge of store departments, decision-making, leadership capacities and soft skills. They describe the promotion – and 10% pay rise – for a 12-year employee based on performance in the VR training. While Holler emphasises that VR assessment is only one of the 'data points' used in hiring decisions, the VR companies we examined enthusiastically frame the potential for the complete automation of these decisions.

One of STRIVR's key patents (US10,586,469 B2, granted on 10 March 2020) is over an 'algorithm to predict how performance in a virtual environment will map to performance in that same situation or task in real life. This method automatically clusters learners into groups based on sensing data, which can include head, hand, and eye movements, as well as physiological data' (via a STRIVR press release). In a webinar, Michael Casale (Chief Science Officer) says that the data they collect (in this case, decision making, performance, attention and engagement) predict 'in almost eighty percent or more than 80 percent of the cases how people would actually perform in the real world', subsequently suggesting that companies – based on as little as '20 minutes of VR' – can therefore 'actually start to make predictions of real-world performance just based on what's going on in the headset, and that's of course incredibly efficient'. STRIVR is not alone in this; Immerse.io also suggest in a press release that VR has the potential to be used in recruitment, and in a press interview Mursion claims that the data gathered through their existing (non-AI based) VR training will be able automatically 'measure human behavioural change', and they describe a project that aims to design performance tasks that 'will be used as an alternative teacher certification assessment' in higher education. TaleSpin also makes a similar proposal for the future of their products.

Our point here isn't that performance in VR can't sometimes predict success in the real world, but reiterate that 'the choice to implement automation within the existing socio-economic context carries with it a set of built-in tendencies that have important societal consequences' (Andrejevic, 2019: 26). VR's fantasy of perfect data – that it captures for objective analysis a mirror-like reflection of the learning experience – is likely based on normative and exclusionary assumptions (which have in the past been gendered, classed, and raced in their outlook, see Crawford & Joler, 2018; Hoffman, 2018; Noble, 2018). While we found no discussion of the training datasets in our study, commercial machine-learning products are typically trained on biased datasets of neurotypical, male and able-bodied engineers (Mehrabi et al., 2019). In the context of VR learning analytics, this has the potential to be a form of Eugenic Gaze (via Swauger, 2020) as they may codify xenophobia, ableism and White supremacy behind the black-box of algorithmic bias, 'while avoiding equity-based critiques because of our belief in the neutrality of data and technology' (Swauger, 2020). STRIVR, for instance, describe in a webinar how they use 'verbal analytics' to measure 'verbal fluency', which they suggest provides 'objective' and 'automated' predictions of a trainee's capability to deal with an emotional customer. Our concern is that the implementation of these technologies may overlook the limitations of speech recognition which 'work best for white, highly educated, upper-middle class Americans' (Harwell, 2018), with recent research suggesting error rates almost twice as high for African Americans speakers versus White speakers (35% vs 19%, Koenecke et al., 2020). Presumably, age, physical fitness and experience with VR technologies also all play a role – hidden in the 'objectivity' of VR's embodiment data – in the predictive power of VR learning analytics. On this basis, we argue that the growing use of VR learning analytics without critically interrogating the VR fantasy of 'perfect' data has the potential to exacerbate, rather than solve, issues of bias and discrimination.

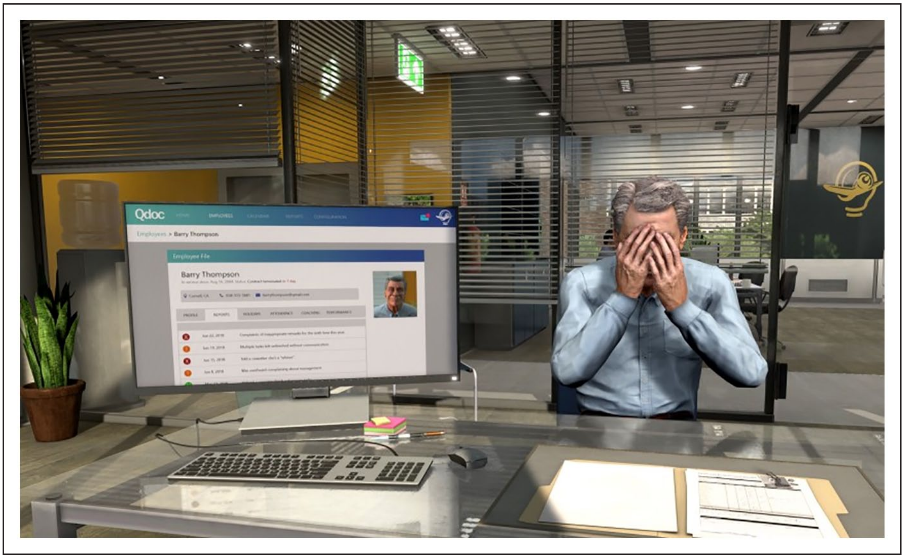

Figure 1. Barry getting fired in TaleSpin's virtual reality training module (L.A. Times/TaleSpin).

Quantifying the qualitative: the datafication of soft-skills training

In 2019, TaleSpin exhibited a demo of their Virtual Human technology that puts 'the user in the shoes of an HR manager tasked with terminating a fellow employee named "Barry"'. The press release goes on to describe how:

> The scenario captures the real stress and emotion typically associated with this situation and presents trainees with common wrongful termination pitfalls such as demonstrating bias. The AI-enabled software provides real-time feedback so a trainee can gain virtual experience that feels real enough to create emotional muscle memory and get real-time guidance on how to empathetically and effectively terminate an employee.

The highly realistic Barry – an older adult in his sixties (see Figure 1) – uses artificial intelligence to respond to the behaviour of the user. If your responses are deemed too aggressive, Barry will be dismayed at his chances of finding employment elsewhere at his age. Too soft, Barry might get angry and pound his fists on the table. In interviews, TaleSpin co-founder Kyle Jackson is quoted as saying 'trainees hands sweat, some start crying, others take off the headset, and he says that is the whole point'. Following the negative media attention that this application received (the LA Times headline was 'Barry sobbed as he begged for his job. VR is getting heavy, man'), TaleSpin have claimed that the Barry demo is just a demo, and no clients have requested using this software to train the employee termination. However, TaleSpin co-founder Stephen Fromkin still suggests that the Barry demo demonstrates 'the emotional resonance of a virtual human for learning' (via Bergman, 2020).

Barry is an evocative example of how VR is allowing an expansion of datafication and learning analytics into the previously qualitative domain of training soft skills such as empathy, communication and de-escalation (for critiques of the datafication of workplaces, see Kim, 2017; Rhagavan et al., 2019; Rosenblat, 2018; Sanchez-Mondero et al., 2019). TaleSpin's website advertises the underlying technology as being able to 'make training for 'soft skills' measurable', and the potential implication of this is the part, or full, automation of these kinds of learning assessments. Beyond Barry, we found a wide variety of examples: TaleSpin's Virtual Human provides effective feedback and sales training to executives, Mursion's puppetry-based live-actors are used in universities for pre-service elementary school teachers to train classroom teaching skills, and STRIVR partnered with Verizon to train employees how to respond to an armed robbery. Among other benefits such as cost savings, these companies frame the benefit of training these 'soft skills' with VR tools like Barry in their capacity to for the quantification of performance and outcomes. TaleSpin's press release describes how Barry is part of a platform that provides 'gamified scoring, dynamic feedback and enterprise integrations allow[ing] Talespin customers to track employee engagement and development'. STRIVR's interpersonal skills training with Walmart 'monitors employees' quantitative progress' (STRIVR website, 2020), Mursion's website advertises how its 'virtual reality training software includes machine learning to quantify the results of each training experience', and one of the three key benefits of Immerse.io's VR learning management system is that it can 'capture every detail with customised reporting' for 'data-driven insights' such as demonstrating the return of investment on training programmes (Immerse.io Homepage, 2020). We argue this is a key and important divergence in the use of VR as an educational technology, between VR's use as a simulator of rare or expensive situations (such as fire at a gas refinery) on one hand, and VR's capability to quantify the learning experience on the other.

This quantification of the qualitative relies on the fantasy of VR offering perfect data and highlights how VR is being shaped by neoliberal 'metric power' (via Beer, 2016). The desire that neoliberal enterprise has to be able to measure, quantify, compare and rank every aspect of how their institutions operate and behave may lead to the datafication of the qualitative aspects of their operation. Viewed through this lens, the adoption of VR training for soft skills is not just attractive to enterprise because it is cheaper, or provides a better form of training, but because it materialises the competitive neoliberal impulse to quantify and make decisions based on data. VR companies are attuned to this and responding to it. In an announcement of a significant contract with BP, Immerse.io claims that the key benefit of their Virtual Enterprise Platform is the ability to 'generate detailed data and reports on learner progress' (Immerse.io Press Release, 2020). In a *Training Industry* magazine article, STRIVR CEO Derek Belch similarly describes how, 'thanks to the data-collecting ability inherent in VR, businesses can better map their L&D efforts to quantifiable results and benefits'. Mursion also proposes that their analytics 'can inform future training plans' (LinkedIn Profile, 2020). In a video interview with Josh Bersin (CEO of an Industry Research and Advisory Firm) uploaded to STRIVR's YouTube Channel, Bersin puts it plainly: 'immersive learning institutionalizes that in a way I have never seen before, so this is a huge ROI investment'.

Understanding VR training and learning analytics through this lens helps us understand and anticipate its upcoming adoption, and potential issues. Based on the flawed fantasy that the data collected via VR are both objective and complete, neoliberal institutional desire for quantification (following Beer's, 2016 concept of 'metric power') may push VR into domains that were previously dominated by qualitative measurements. This is one of the reasons why we are seeing VR adopted in hiring and promotion decisions: Walmart claims, in its press release about its partnership with STRIVR, that using VR can 'remove subjectivity and unconscious bias from the selection process'. Again, this is not to say that VR does not have a potential in this area. TaleSpin CEO discussed at an Oculus Connect DevCon panel how their sales training can help start a conversation about bias, as they found that users would significantly change their behaviour when the avatar of the sales target was changed (e.g. from an old White male CEO to a young Black female CEO). Our concern is with the potential extraction and analysis of data from this training that obfuscates the subjectiveness of data and relies on incomprehensible machine learning models. Any new form of measurement and classification calls into question issues of power – who defines what is measured, what is not measured and what values are prescribed to these measurements. Simplifying the messy landscape of soft skills, empathy and bias into what is machine readable has the potential to disenfranchise and discriminate.

VR, real-data empires

The third key issue we identified in our analysis was around the speculative value of data. In an episode of *The Georgian Impact Podcast* – a product of Georgian Partners, a Canadian Venture Capital firm that led STIVR's US$30 m Series B funding round in March 2020 – STRIVR founder and CEO Derek Belch discussed the expansiveness, and the value, of the data that they have collected from their users so far:

> Then when you think about machine learning and AI and all of this, well data begets data, it becomes this infinite loop . . . the more people that are using this training, the more data that we're collecting, the more data we have to build the models, the more data the models have, we have more people. It just circles itself. Interesting parallel for STRIVR is self-driving cars . . . we have probably a hundred to a thousand times more data than anybody else. And so our models will be that much further along when they start to become more refined and more specific.

Since STRIVR owns the aggregate data associated with usage of their VR training tools and collects data about how this VR data maps to the real-world performance, he further suggests the potential for comparing employees across companies for the development of learning models and adaptive exams. A Data Platform Engineer role advertised at STRIVR claims 'tens of millions' of in-headset sessions that will feed into 'a streaming analytics platform that will allow us to process, join, aggregate, reform and query these very large structured and unstructured datasets to produce immersive analytics with deep insights on learning sessions'. Interconnected with the issues we previously discussed around the fantasy of VR data, Belch goes on to envision a future in which these adaptive and autonomous simulations send users 'down certain paths based on how you perform, and that's going to impact your score'. Both the training and outcomes may become subject to new forms of analysis, segmentation and bias (as discussed by Amoore, 2020).

Yet beyond the issues with automation, what Belch is indicating in this interview is a new educational data gold rush. His claim is that whichever VR company can acquire the most data first will get an unassailable lead over their competitors, as the models and automation that they can create will provide an exponential lead. In the interview above, Belch compares the situation to the race to release a self-driving car, often measured by which technology company has collected the most data (O'Kane, 2018). Immerse.io echoes this speculation. In a webinar on data capture, Immerse.io's COO Justin Parry describes how VR experiences can be replayed, and therefore re-analysed, if analysis methods improved. All of this, of course, hinges on speculation (Hong, 2020). That the data collected now will be of any use: that machine learning, data science and artificial intelligence will determine some way to refine these lakes of data and incorporate them into the development and sale of new products that can similarly extract more data.

What this indicates is – perhaps unsurprisingly, but importantly worth forewarning – the further expansion of data capitalism (following Sadowski, 2020; Zuboff, 2019) into VR education technologies (see also Komljenovic, 2020; 2021 for recent work in the specific context of higher education). For Srnicek (2016), in his account of platform capitalism, one of the key economic tendencies of big tech platforms is their extraction of data from large userbases. Yet, as the above shows, data are not necessarily always neatly converted into monetary value or used to any particular end. Rather, as Sadowski (2019) writes:

> The capitalist is not concerned with the immediate use of a data point or with any single collection, but rather the unceasing flow of data-creating . . . The imperative, then, is to constantly collect and circulate data by producing commodities that create more data and building infrastructure to manage data. The stream of data must keep flowing and growing. (p. 4)

The issue then, for higher education, businesses, unions, researchers and policymakers and, so forth, is understanding how these data are being owned, traded and commodified. The pressing issue here is not just one of privacy, but in the abuse of these data against the interest of the user or their communities. Greller and Drachsler (2012) suggest that the default position for data ownership in learning analytics is that the data 'belongs to the owner of the data collection tool, typically also the data client and beneficiary', such as the university in the context of higher education, but in the emerging landscape of VR education technologies we see that this is not the case. Both STRIVR and Immerse.io retain the data that account for the users' actions within the VR application (i.e. their entire learning experience). Indeed, the retention of these data is one of their key value offerings to enterprise, and the future promise of analytics that may be applied to it.

One reason that enterprise and higher education institutions should be militant about the ownership and trade or VR data is the potential for biometric identification of users based on their behaviour in VR. Both Pfeuffer et al. (2019) and Miller et al. (2020) have recently demonstrated the potential to identify specific individual users based on unique behavioural biometric markers via movements of the hand, head and eye – data that can be captured by current VR devices. The implications of this are profound, as it suggests that the de-anonymisation of individual training sessions may always be possible.

Where one of the key principles that typically underscores questions of data ethics and privacy is one of informed consent (Hildebrandt, 2006), we can easily speculate about where VR data analytics might be used beyond expected use. For instance, a collection of multinational companies or governments to rank universities based on the comparison of aggregated VR training data of graduate students? Or the later sale to of this valuable data in the bankruptcy proceedings of a failed VR start-up? Williamson et al. (2020) warns us about organisations that have already suggested that it 'may be possible to quantify the value of every university module, course or career choice' via online learning data, and we further point to companies like Learning Economy that seek to integrate the results of these analyses in permanent blockchain-based records for the benefit and analysis of employers (see Gent, 2020; TaleSpin proposes something similar). Considering the potential for behavioural biometrics to identify unique users from VR data sets, the use of additional technologies may not even be necessary for this purpose.

This is not purely the domain of Silicon Valley, but we note some academic learning analytics research that describes – as 'very promising future trends' – the potential for learning analytics to take 'into account all students' personal data through their whole life' and in 'analyzing and mining data directly gathered from students' brain for a better understanding of the learning' (purposefully not cited).[1] As Corrin et al. (2019) conclude in their discussion article on the ethics of learning analytics in higher education, 'a more nuanced approach to the ethics of data and analytics is needed, one that accounts for the multiple levels and uses of analytics' (p. 4). Where the ethical implications of data analytics – exacerbated by the expansiveness and framelessness of VR data – are currently unclear, data ownership and control need to be at the forefront of discussions about VR as an education technology, and the adoption of it in society more broadly.

Conclusion

VR undoubtedly offers significant potential as an education technology. Our intent in this article is not to dismiss the enormous potential for engaging, powerful and transformative learning experiences that can be delivered in VR. Indeed, Carter has developed and researched promising VR experiences for conservation education (Carter et al., 2020), and a wide variety of research projects are showing the potential for VR beyond technological novelty. The capability to realistically simulate dangerous, rare or expensive scenarios will have a transformative impact in education and entertainment more broadly.

However, the interrogation of emerging VR education companies we have presented calls for caution in the unregulated and uncritical application of learning analytics to the data that are collected from VR training, particularly in higher education where there is a greater moral responsibility to serve the interests of learners. It is crucial that scholars and policymakers pay sustained attention to how VR learning data is collected and used and resist the extraction and commodification of VR learning data, the use of data analysis to compare users in aggregate and automating decision-making based on VR performance. We specifically recommend that higher education institutions recognise and treat any data collected in VR as individually identifiable information, afforded the same protection as other private data. We are not alone in this recommendation: the recent XRSI privacy framework (XR Safety Initiative, 2020) argues that VR data falls under the GDPR's (General Data Protection Regulation) classification of biometric data. We also advocate for, at a bare minimum, greater degrees of transparency in the use of VR learning data in processes such as hiring, promotion and so on (such as through independent audits to assess validity) – as advocated for in wider critique of algorithmic hiring technologies (see Rhagavan et al., 2019), technologies which generally have been identified as instituting new forms of bias and discrimination (Kim, 2017). Such are the risks that we ask – as Andrejevic and Selwyn (2020) ask of facial recognition technology in education – if the 'added value' or 'gained efficiencies' of VR learning analytics outweigh the harmful potential for algorithmic bias, discrimination and surveillance.

In this article, our analysis has – at times – been *necessarily speculative*, an approach inspired by Carah and Angus's (2018) work on algorithmic brand culture, which acknowledges how many actual and emerging technology practices are hidden behind non-disclosure agreements and within technological black boxes, and Stilgoe et al'.s (2013) call for improved anticipation in the context of their *responsible innovation* framework. Critical speculation about the potential consequences and designs of technologies, grounded in prior technological movements and critical research, are necessary because of the *Collingridge Dilemma* (via Kudina and Verbeek, 2019) – that is, the methodological problem where, if a technology is not yet widely adopted, we run the risk of being too speculative in anticipating its impacts. But, if we wait until it is widely adopted in society, it becomes difficult to challenge its entrenchment and power. In this way, speculation about emerging technologies is a necessary component in developing critical approaches towards them.

Mark Zuckerberg, among others, has claimed that VR will emerge as the next dominant platform for computing, meaning that VR will expand beyond learning and potentially become a platform where labour is performed, and thus subjected to similar levels of data extraction, analysis and surveillance. Facebook's recently announced *Infinite Office* – a VR office space – and increased surveillance of people working at home during COVID-19 anticipates such creeping surveillance (Hern, 2020). We have also focused solely on VR in this article, but we believe similar issues will also be relevant in the context of AR devices such as the Microsoft Hololens, and forthcoming AR or 'Mixed Reality' headsets from Apple, Google and Facebook. While the AR experience is not fully virtual like VR, it functions – as Mark Pesce (2020) argues – via surveillance, requiring enormous amounts of data about the users' environment to function. Such data extractive capacities may lead to similar fantasies about perfect and complete data pictures, further exacerbating issues with datafication and overconfidence in the veracity of the data collected. Future work will be needed as these technologies become widely available in educational and workplace contexts.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by funding from the Faculty of Arts and Social Sciences at The University of Sydney.

ORCID iDs

Marcus Carter https://orcid.org/0000-0002-4866-4928
Ben Egliston https://orcid.org/0000-0002-7878-7208

Note

1. We have not cited the source of this quote, as we do not want to implicitly endorse it via the function of citation metrics. It is findable if you search the quote online.

References

Amoore L (2020) *Cloud Ethics*. Durham, NC: Duke University Press.

Andrejevic M (2019) *Automated Media*. New York: Routledge.

Andrejevic M and Selwyn N (2020) Facial recognition technology in schools: critical questions and concerns. *Learning, Media and Technology* 45(2): 115–128.

Angwin J, Larson J, Mattu S, et al. (2016) Machine bias. *Propublica*. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Beer D (2016) *Metric Power*. London: Palgrave Macmillan.

Beer D (2018) *The Data Gaze: Capitalism, Power and Perception*. Thousand Oaks, CA: SAGE.

Bergman B (2020) Talespin worker training platform raises $15 million. *Dot.la*, 3 March. Available at: https://dot.la/talespin-virtual-reality-training-2645361687.html

Bollmer G (2017) Empathy machines. *Media International Australia* 165(1): 63–76.

Bowker G and Star S (2000) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

Brock A (2018) Critical technocultural discourse analysis. *New Media & Society* 20(3): 1012–1030.

Brown M (2020) Seeing students at scale: how faculty in large lecture courses act upon learning analytics dashboard data. *Teaching in Higher Education* 25(4): 384–400.

Buckingham S and Ferguson R (2012) Social learning analytics. *Educational Technology and Society* 15(3): 3–26.

Buckingham Shum SJ and Luckin R (2019) Learning analytics and AI: politics, pedagogy and practices. *British Journal of Educational Technology* 50(6): 2785–2793.

Carah N and Angus D (2018) Algorithmic brand culture: participatory labour, machine learning and branding on social media. *Media, Culture & Society* 40(2): 178–194.

Carter M and Egliston B (2020) Facebook's virtual reality push is about data, not gaming. *The Conversation*. Available at: https://theconversation.com/facebooks-virtual-reality-push-is-about-data-not-gaming-145730

Carter M, Webber S, Rawson S, et al. (2020) Virtual reality in the zoo: a qualitative evaluation of a stereoscopic virtual reality video encounter with little penguins Eudyptula minor. *Journal of Zoo and Aquarium Research* 8(4): 239–245.

Corrigan T (2018) Making implicit methods explicit: trade press analysis in the political economy of communication. *International Journal of Communication* 12: 2751–2772.

Corrin L, Kennedy G, French S, et al. (2019) The ethics of learning analytics in Australian higher education: a discussion paper. Available at: https://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0004/3035047/LA_Ethics_Discussion_Paper.pdf

Crawford K and Joler V (2018) Anatomy of an AI System: the Amazon Echo as an anatomical map of human labor, data and planetary resources. *AI Now Institute and Share Lab*, 7 September. Available at: https://anatomyof.ai

Crawford K and Paglen T (2019) Excavating AI: the politics of training sets for machine learning. 19 September. Available at: https://excavating.ai

Dalton CM (2015) For fun and profit: the limits and possibilities of Google-Maps-based geoweb applications. *Environment and Planning A: Economy and Space* 47(5): 1029–1046.

Daniela L (2020) New perspectives on virtual and augmented reality. In: Daniela L (ed.) *New Perspectives on Virtual and Augmented Reality*. Available at: https://www.taylorfrancis.com/books/edit/10.4324/9781003001874/new-perspectives-virtual-augmented-reality-linda-daniela

Davis L (1995) *Enforcing Normalcy: Disability, Deafness and the Body*. Verso: London.

Egliston B and Carter M (2020) Oculus imaginaries: the promises and perils of Facebook's virtual reality. *New Media & Society*. Epub ahead of print 24 September. DOI: 10.1177/1461444820960411.

Evans L (2019) *The Re-Emergence of Virtual Reality*. Routledge: London.

Gent E (2020) Is education the new currency? *BBC Worklife*. Available at: https://www.bbc.com/worklife/article/20200210-is-education-the-new-currency

Gilliard C and Hugh C (2016) Digital redlining, access, and privacy. *Common Sense Education*, 24 May. Available at: https://www.commonsense.org/education/articles/digital-redlining-access-and-privacy

Golding D (2019) Far from paradise: the body, the apparatus and the image of contemporary virtual reality. *Convergence* 25(2): 340–353.

Greller W and Drachsler H (2012) Translating learning into numbers: a generic framework for learning analytics. *Educational Technology and Society* 15(3): 42–58.

Haiven M (2013) Walmart, finance, and the cultural politics of securitization. *Cultural Politics* 9(2): 239–262.

Harley D (2019) Palmer Luckey and the rise of contemporary virtual reality. *Convergence* 26(5–6): 1144–1158.

Harwell D (2018) The Accent Gap. *The Washington Post*, 19 July. Available at: https://www.washingtonpost.com/graphics/2018/business/alexa-does-not-understand-your-accent/

Hern A (2020) Shirking from home? Staff feel the heat as bosses ramp up remote surveillance. *The Guardian*, 27 September. Available at: https://www.theguardian.com/world/2020/sep/27/shirking-from-home-staff-feel-the-heat-as-bosses-ramp-up-remote-surveillance

Hildebrandt M (2006) Privacy and identity. In: Claes E, Duff A and Gutwirth S (eds) *Privacy and the Criminal Law*. Oxford, UK: Intersentia, pp. 61–104.

Hoffman AL (2018) Data violence and how bad engineering choices can damage society. Available at: https://medium.com/s/story/data-violence-and-how-bad-engineering-choices-can-damage-society-39e44150e1d4

Holler D (2019) Customers are changing: jobs are changing: at Walmart, the future of work is bright. *Walmart Newsroom*, 30 October. Available at: https://corporate.walmart.com/newsroom/2019/10/30/customers-are-changing-jobs-are-changing-at-walmart-the-future-of-work-is-bright

Hong S-H (2020) *Technologies of Speculation: The Limits of Knowledge in a Data-driven Society*. New York: New York University Press.

Jasanoff S (2017) Virtual, visible, and actionable: data assemblages and the sightlines of justice. *Big Data & Society* 4(2): 1–15.

Katwala A (2020) An algorithm determined UK students' grades: chaos ensued. *Wired*, 15 August. Available at: https://www.wired.com/story/an-algorithm-determined-uk-students-grades-chaos-ensued/

Kelly K (2019) AR will spark the next big tech platform – call it mirrorworld. *WIRED*, December. Retrieved from https://www.wired.com/story/mirrorworld-ar-next-big-tech-platform/

Keyes O (2019) Counting the countless why data science is a profound threat for queer people. *Real Life Mag*, 2 December. Available at: https://reallifemag.com/counting-the-countless/

Kim P (2017) Data-driven discrimination at work. *William & Mary Law Review* 48: 857–936.

Kitchin R (2014a) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Thousand Oaks, CA: SAGE.

Kitchin R (2014b) Big Data, new epistemologies and paradigm shifts. *Big Data & Society*: 1(1): 1–12.

Koenecke A, Nam A, Lake E, et al. (2020) Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America* 117(14): 7684–7689.

Komljenovic J (2020) The future of value in digitalised higher education: why data privacy should not be our biggest concern. *Higher Education*. Epub ahead of print 19 November. DOI: 10.1007/s10734-020-00639-7.

Komljenovic J (2021) The rise of education rentiers: digital platforms, digital data and rents. *Learning, Media and Technology*. Epub ahead of print 22 February. DOI: 10.1080/17439884.2021.1891422.

Kudina O and Verbeek PP (2019) Ethics from within: Google glass, the Collingridge dilemma, and the mediated value of privacy. *Science Technology and Human Values* 44(2): 291–314.

LaRocco M (2020) Developing the 'best practices' of virtual reality design: industry standards at the frontier of emerging media. *Journal of Visual Culture* 19(1): 96–111.

Liao T (2018) Mobile versus headworn augmented reality: how visions of the future shape, contest, and stabilize an emerging technology. *New Media & Society* 20(2): 796–814.

Liu D, Dede C, Huang R, et al. (2017) *Virtual, Augmented, and Mixed Realities in Education*. Cham: Springer.

Mehrabi N, Morstatter F, Saxena N, et al. (2019) A survey on bias and fairness in machine learning. Arxiv Preprint Arxiv:1908.*09635*. Available at: https://arxiv.org/abs/1908.09635

Miller M, Herrera F, Jun H, et al. (2020) Personal identifiability of user tracking data during observation of 360-degree VR video. *Scientific Reports* 10: 17404.

Noble S (2018) *Algorithms of Oppresion*. New York: New York University Press.

O'Kane S (2018) Tesla vs. Waymo: who's winning the race for self-driving cars. *The Verge*, 18 April. Available at: https://www.theverge.com/transportation/2018/4/19/17204044/tesla-waymo-self-driving-car-data-simulation

Pesce M (2020) *Augmented Reality*. London: Polity Press.

Pfeuffer K, Geiger MJ, Prange S, et al. (2019) Behavioural biometrics in VR. In: *Conference on human factors in computing systems – Proceedings 12*. Available at: https://dl.acm.org/doi/10.1145/3290605.3300340

Rhagavan M, Barocas S, Kleinberg J, et al. (2020) Mitigating bias in algorithmic hiring: evaluating claims and practices. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency (FAT*'20)*, Barcelona, January, pp. 469–481. New York: ACM Press.

Rosenblat A (2018) *Uberland*. Berkeley, CA: UC Press.

Sadowski J (2019) When data is capital: datafication, accumulation, and extraction. *Big Data & Society* 6(1): 882054.

Sadowski J (2020) *Too Smart*. Cambridge, MA: MIT Press.

Saker M and Frith J (2019) From hybrid space to dislocated space: mobile virtual reality and a third stage of mobile media theory. *New Media & Society* 21(1): 214–228.

Saker M and Frith J (2020) Coextensive space: virtual reality and the developing relationship between the body, the digital and physical space. *Media, Culture & Society* 42: 1427–1442.

Sanchez- Mondero J, Dencik L and Edwards L (2019) What does it mean to solve the problem of discrimination in hiring? *Social, technical and legal perspectives from the UK on automated hiring systems*. Available at: https://arxiv.org/abs/1910.06144

Selwyn N (2020) Re-imagining 'Learning Analytics' . . . a case for starting again? *The Internet and Higher Education* 46: 100475.

Slade S and Prinsloo P (2013) Learning analytics: ethical issues and dilemmas. *American Behavioral Scientist* 57(10): 1510–1529.

Smyth J (2017) *The Toxic University Zombie Leadership, Academic Rock Stars and Neoliberal Ideology*. London: Palgrave Macmillan.

Society for Learning Analytics Research (SoLAR) (nd) What is learning analytics? Available at: https://www.solaresearch.org/about/what-is-learning-analytics/

Srnicek N (2016) *Platform Capitalism*. New York: Wiley.

Statt N (2020) Facebook reportedly ignored its own research showing algorithms divided users. *The Verge*, 26 May. Available at: https://www.theverge.com/2020/5/26/21270659/facebook-division-news-feed-algorithms

Stilgoe J, Owen R and Macnaghten P (2013) Developing a framework for responsible innovation. *Research Policy* 42(9): 1568–1580.

Swauger S (2020) Our bodies encoded: algorithmic test proctoring in higher education. *Hybrid Pedagogy*. Available at: https://hybridpedagogy.org/our-bodies-encoded-algorithmic-test-proctoring-in-higher-education/

Tuchscherer R (2019) Walmart uses virtual reality to test new store managers. *USA TODAY*, 8 July. Available at: https://www.usatoday.com/story/tech/2019/07/08/walmart-uses-virtual-reality-hire-new-managers/1635311001/

Wallis K and Ross M (2020) Fourth VR: indigenous virtual reality practice. *Convergence* 27(2): 313–329.

Warzel C (2020) Is QAnon the most dangerous conspiracy theory of the 21st century? *The New York Times*, 4 August. Available at: https://www.nytimes.com/2020/08/04/opinion/qanon-conspiracy-theory-arg.html

Wasko J (2004) The political economy of communications. In: Downing J, McQuail D, Schlesinger P, et al. (eds) *The SAGE Handbook of Media Studies*. Thousand Oaks, CA: SAGE, pp. 309–330.

Wilken R, Burgess J and Albury K (2019) Dating apps and data markets: a political economy of communication approach. *Computational Culture a Journal of Software Studies* 7(7): 1–26.

Wilken R and Sinclair J (2009) 'Waiting for the kiss of life' mobile media and advertising. *Convergence* 15(4): 427–445.

Williamson B (2020) Making markets through digital platforms: Pearson, edu-business, and the (e)valuation of higher education. *Critical Studies in Education* 62: 50–66.

Williamson B, Bayne S and Shay S (2020) The datafication of teaching in Higher Education: critical issues and perspectives. *Teaching in Higher Education* 25(4): 351–365.

XR Safety Initiative (2020) The XRSI privacy framework (version 1.0). Available at: https://xrsi.org/publication/the-xrsi-privacy-framework

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

Author biographies

Marcus Carter is a Senior Lecturer in Digital Cultures at The University of Sydney, and co-director of the Socio-Tech Futures Lab. He is also an honorary research fellow in the Interaction Design Lab at The University of Melbourne. An overview of his research interests and publications can be found on his website, www.marcuscarter.com.

Ben Egliston is a postdoctoral research fellow in the Digital Media Research Centre at Queensland University of Technology. He currently researches the practices and politics of emerging technologies such as data analytics and Virtual and Augmented Reality.