

Refocusing Zuboff's 'division of learning' on Education

Jeremy Knox

University of Edinburgh

Email: jeremy.knox@ed.ac.uk

Abstract

This paper examines the concept of the 'division of learning' (Zuboff 2019), and the broader thesis of 'surveillance capitalism' within which it is situated, in terms of its relevance to education. It begins with defining the term, before suggesting two key ways in which aligning the 'division of learning' with perspectives from educational research provides productive insights. The first considers the impact of increasing 'datafication' in education, where platform technologies are proliferating as powerful actors that both mediate and shape educational activity. Here the 'division of learning' offers useful insights concerning the disparities resulting from learning in and learning from educational platforms. The second explores the extent to which education theory might offer ways to develop the concept of the 'division of learning', through critique of the term 'learning' itself, as well as the foregrounding of questions about educational 'purpose'. Here the 'division of learning' is suggested to maintain, rather than challenge, the dominant practices of data exploitation, for which further engagement with a purposive, political, and emancipatory form of 'data science' is suggested.

Keywords: 'division of learning'; 'surveillance capitalism'; datafication; platformisation; learnification; education

Introduction

This paper begins with a definition of the concept of the 'division of learning' (Zuboff 2019), before suggesting two key ways in which the concept might be relevant to the discussion of education and its increasing relations with a data-fuelled technology industry. Concerning the first, the 'division of learning' will be suggested to highlight important issues of privacy, inequality, and authority where data-driven platform software is increasingly subjecting educational activity to the same kinds of datafication practices that are more usually identified in wider society. An important separation, and contrast, will be highlighted here between the 'learning' undertaken by students within platform software, and the 'learning' undertaken by machines and their private owners, deriving from the resulting student data. This student data can be understood as behavioural data, in the sense that it is predominantly derived from the ways students interact with platform software, and analysed using theories of behavioural learning (Knox et al. 2020). The concept of the 'division of learning' will be proposed, therefore, as a valuable means of identifying potentially significant inequalities emerging directly from the activities of public education, particularly where private companies have intensified their interests in educational activity as a result of the recent COVID-19 pandemic and the subsequent lockdown of schools, colleges, and universities in some countries.

In the second, ideas from educational theory will be suggested to offer critical perspectives on the concept of the 'division of learning'. While Zuboff's (2019) work is critical of the power and influence accrued by private companies, it doesn't necessarily challenge some of the problematic ways in which 'learning' itself is defined. Here, the 'division of learning' will be examined for the extent to which it focuses exclusively on the narrow idea of 'training' in the technical skills required to benefit from datafication, rather than on the broader socially-oriented arrangement of formal education. Drawing on the work of Biesta (2009; 2013; 2015), this section draws a distinction between discursive representations of 'learning' as individualised, self-directed and 'life-long' in duration, and 'education' as public, institutionalised, and fundamentally concerned with collective purposes. Making connections with work in 'critical data literacy' (Sander 2020), this analysis will suggest, not only more critical ways of understanding 'learning' in the datafied society, but also the need for better recognition of the role that formal education can play in addressing the associated issues of power and inequality.

This paper therefore aims to emphasise educational concerns within the broader discourses around data-driven technologies. Firstly, by highlighting how prominent ideas, such 'surveillance capitalism', are specifically relevant to the practice of education. And secondly, how perspectives from education can critically develop wider public discourse about the relationships between technology and learning.

Defining the division of learning

Zuboff's (2019) concept of 'the division of learning' is grounded in a wider thesis of 'surveillance capitalism', which has emerged as a key term in a field of critique focused on the relationships between powerful 'big tech' companies and wider society, and the ways such associations are increasingly mediated through data-driven systems and their attendant technologies of machine learning and analytics. Zuboff suggests that 'just about everything we now do is mediated by computers that record and codify the details of our daily lives' (p182), and it is this codification that has become central to a new form of data-fuelled capitalism. The concept of the 'division of learning' is the notion that there is a barrier between the ability to access, understand, and benefit from this data, and the practice of simply producing it. In this way Zuboff posits a significant rupture between 'big tech' companies and the wider population, where the former are able to gain unprecedented advantages through fine-grained analyses of societal behaviours, accelerating their ability to dominate markets and accumulate wealth, in ways that manifest a stark contrast between those *with* knowledge of data-driven systems and those *without*. Google in particular are singled out for their aggressive corporate approach to the exploitation of data, through which Zuboff suggests that they have 'amassed world-historic concentrations of knowledge and power' (2019, p179).

The phrase 'division of learning' is proposed as a revision of the concept of the 'division of labour'¹, and as such is suggested to form a foundational theory and definitive way of understanding the ordering principles of data-capitalism. The 'division of learning' is, for Zuboff, 'the axial principle of social order in an information civilisation' (2019, p179), and the 'ascendant principle of social ordering in our time' (2019, p181). Here Zuboff (2019) reveals the centrality of this concept within her broader critique, but also a primary interest in the notion of an 'information age' and its accessibility, a position which will be examined further in subsequent sections. As a route to resolving the suggested dilemmas of 'knowledge, authority, and power' surfaced by the 'division of learning', Zuboff poses three key questions: 'who knows?'; 'who decides?'; and 'who decides who decides?' (2019, p180). The first question concerns the accessibility of knowledge, and 'whether one is included or excluded from the opportunity to learn', for which Zuboff's answer is an 'elite cadre able to wield the analytic tools to trouble shoot and extract value from information' (2019, p180). This, as will be elaborated below, appears to be Zuboff's (2019) primary concern, directed towards the privatised realm of the 'big tech' companies. The second and third questions concern authority and power respectively, through which Zuboff suggests both an exclusory system that maintains control of who is allowed to engage in 'learning'

¹ Zuboff (2019) draws specifically from Durkheim's work on the 'division of labour' (1964), and the proposition that such organisation and specialisation constituted a kind of 'natural law' of developing societies (and organic life more generally, although Zuboff focuses specifically on society). The division of learning thus 'follows the same migratory path from the economic to the social domain' (Zuboff 2019, p184)

from data, as well as an underlying driving force of 'financial capital bound to the disciplines of shareholder-value maximisation' (2019, p181). These questions appear to add a critical depth to the concept of the 'division of learning', however the final section of this paper will return to the themes of authority and power in order to assess whether they are carried through in Zuboff's (2019) work.

One of the central ideas in the concept of the 'division of learning' is what Zuboff defines as the 'shadow text'. This is developed from her earlier work on the notion of a 'new electronic text' (Zuboff 2019, p182), produced primarily in workplace settings and taking the form of data about manufacturing processes and associated labour. Given the great proliferation of data collection practices in more recent times, Zuboff suggests that 'the dilemmas of knowledge, authority, and power' that were intrinsic to the production and use of this electronic text 'have burst through the walls of the workplace to overwhelm our daily lives' (2019, p181). For this reason, Zuboff extends this idea into *two* texts: a 'public-facing text', comprised of the participatory and information-rich spaces of the web, much of which is produced by 'us', and a second 'shadow text', for which the first text 'functions as the supply operation' (2019, p185). In essence, the 'shadow text' is composed of the surplus extracted from the first text, rendered into 'raw material to be accumulated and analysed as means to others' market ends' (Zuboff 2019, p185). For Zuboff, 'surveillance capitalism' has ultimate control over both texts, the 'shadow text' hidden from view and used to extract value from our behavioural data, as well as the public text of the web, which is manipulated into forms that best serve the 'shadow text' beneath it. In this sense 'surveillance capitalists' hold 'asymmetrical power that enables them to bend the division of learning to their interests' (Zuboff 2019, p187). In this sense, the distinction between the 'public-facing' and 'shadow' texts constitutes the 'division of learning', where the ultimate benefit comes from the ability to access, comprehend, and derive value from the hidden side of the partition.

The 'division of learning' and education

As noted by Williamson (2019), Zuboff's work on 'surveillance capitalism', including the specific proposition of the 'division of learning', makes no explicit reference to education. Instead, as we have seen above, Zuboff (2019) centres on the term 'learning', which is framed as the practice of extracting, understanding, and deriving value from data, and suggested to be undertaken by a small and elite group of knowledgeable experts and their sophisticated learning machines. In this sense, we might understand concepts such as the 'division of learning', despite the critique of technology, as fundamentally aligned with a sector that has long called for the disruption of traditional education through technological intervention (Christensen et al. 2008). Without any consideration of formal education, two central questions become relevant. Firstly, what useful insights can educationalists derive from the concept, given that Zuboff (2019) provides no examples of 'surveillance

capitalism' in the context of education? And secondly, how might the concept of the 'division of learning' itself be (re)considered in the light of perspectives from education?

In answer to the first question, given the increasing prevalence of data-driven platforms in education (Decuyper *et al.* 2021), the 'division of learning' may provide a useful conceptual framework through which to examine the impact on educational activity. In this sense, research from outside of education, and specifically that concerning the relationships between technology, capitalism, and society, can strengthen the ways we might develop understanding about the changing structures, pedagogies, and subjectivities of increasingly datafied institutions and classrooms. Even before the COVID-19 pandemic drastically altered the ability of institutions around the world to conduct educational activities, a commercial 'ed tech' industry was a significant and growing sector that had been promising to disrupt education for decades (Selwyn 2013), as well as exerting considerable influence over teaching practices and student experiences through increasingly powerful data-driven software. However, as emerging scholarship is highlighting (see Williamson and Hogan 2020), the COVID-19 pandemic is ushering in unprecedented relations between public education systems and private education companies, often with little apparent government oversight or in-depth engagement with the far-reaching consequences. The 'big tech' companies examined by Zuboff (2019), such as Google, have now moved, with all their infrastructural might, into mainstream educational spaces, situating their convenient software tools at the very core of teaching and learning functions. Zuboff's (2019) critique of 'big tech' is therefore now as relevant to education as it has been to wider debates about social media. Given that the pandemic has intensified concerted efforts by a range of commercial 'ed tech' actors to further privatise the future of education systems (Williamson and Hogan 2020), education research would do well to engage with ideas such as the 'division of learning', which emphasise the key issues of privacy and inequality that characterise such developments.

In answer to the second question, foregrounding educational theory offers an important route to developing the ways 'learning' itself is theorised in relation to 'surveillance capitalism'. Broader considerations of education can suggest ways, to not only move beyond the rather narrow visions of technical training and data analysis assumed by Zuboff (2019), but also to problematise the underlying assumptions about behaviour that appear to ground both sides of the 'division of learning'. As Knox *et al.* (2020) suggest, particular behaviouristic understandings of 'learning' are becoming entrenched in the era of data-driven technologies, both in terms of the human learning supposedly enhanced through educational analytics, as well as in the sense of the learning undertaken by machines, often derived from psychological theories such as 'reinforcement learning' (see Sutton & Barto 2018). In other words, the idea of 'learning' itself needs better theorisation in order to contend with the significant issues of privacy and inequality that the concept of the 'division of learning' suggests, for which educational theory offers productive insights. Zuboff suggests of the 'division of learning' that it 'reflects the primacy of learning,

information, and knowledge in today's quest for effective life' (2019, p184-5), seeming to underscore a general conviction in the autodidacticism of the internet age. However, it is precisely this idea of the 'primacy of learning' that has undergone considerable critique in the field education, encapsulated most directly in Biesta's concept of 'learnification' (2005; 2009; 2013), which surfaces the inherent politics of 'self-direction' and an all-too-easy alignment with neoliberal ideology. Further, better recognition of specific theories of learning helps to identify a range of contested assumptions about the human condition embedded within the data-driven technologies that Zuboff (2019) examines in such detail. This understanding, as will be suggested below, supports a much deeper form of critique and resistance to 'big tech' corporate practices than merely 'upskilling' in the very same techniques of data exploitation that remain inaccessible through the 'division of learning'.

The 'division of learning' in education platforms

The 'datafication' of education has emerged as a central concern in the field of educational research, as powerful data-driven systems are increasingly being integrated in policy and governance (Williamson 2017), as well as deployed in everyday teaching and learning practices (see Brown 2020; Bulger 2016). This deployment usually occurs through the establishment of a software 'platform', designed with functions to supposedly facilitate teaching and learning activities, as well as a range of administrative functions. The concept of 'platformisation' (van Dijck et al. 2018; Poell et al. 2019) - which concerns the processes through which platform software and its associated practices become involved in the transformation of markets, institutions, governance, and cultural practices - is therefore now highly relevant to the study of education (Decuyper et al. (2021). Writing before the COVID-19 pandemic, Selwyn et al. predict a movement towards '*classrooms on platforms* rather than *platforms in classrooms*' (2020, p2), which appears now to be a prescient encapsulation of the 'the expansion and embedding of digital technologies and media in education systems, institutions and practices across the world' (Williamson et al. 2020, p107) as a consequence of school, college, and university closures in some countries. However, 'platformisation' has been a longer trend in education, where software has been developed for:

the provision of online learning, the management of learning in blended settings, home learning and early childhood education, the promotion of classroom engagement, language learning, test-preparation, the design of innovative schools, school administration and learning analytics (Grimaldi and Ball 2021, p115)

Given that for every aspect of education 'there seems to exist a corresponding digital platform used within (and often especially made for) the educational field' (Decuyper et al. 2021, p1), there is a pressing need for conceptual tools with which to analyse the impact of this platformisation. However, as Decuyper et al. stress, the specificities of education platforms need to be acknowledged:

the educational sphere is at once characterized by its adoption of a huge variety of global platforms entering education 'from the outside' and by the circulation and adoption of various locally produced platforms that have been programmed and designed with the educational sector exclusively in mind. (Decuyper *et al.* 2021, p3)

Education is therefore impacted by the wider reorganisational logics of platformisation at the societal level, while at the same time rearticulating the form of the platform itself according to particular visions of productive teaching, learning, and educational management. This is one of the key tensions into which the concept of the 'division of learning' might be applied.

Motivated by the question of what kind of learning happens in such educational platforms, three interrelated forms might be identified: firstly, the learning undertaken by students through their involvement with the software platform as part of their formal educational activity; secondly, the learning *about students* undertaken by teachers and institutions through their use of the analytic tools provided by the platform; and thirdly, the learning *about 'users'* (in this case students, teachers, and institutional administrators) undertaken by the owners of the platform software. As Perrotta *et al.* (2021) suggest in the case of 'big tech' companies such as Google, this third form of learning might also extend into a broader ecosystem of applications on the web, either directly owned by Google or linked to its extensive infrastructure. All of these kinds of learning are co-constitutive of one another, in the sense that teachers are supposed to make pedagogical decisions based on the insights from data analytics in order to intervene in student behaviours (van Barneveld *et al.* 2012), while the platform software in turn often adapts to individual student responses, particularly in the case of so-called 'personalised' systems (Bulger 2016).

First and foremost, the concept of the 'division of learning' helps to articulate concerns related to the form of learning *about users*, where private companies appear to extract value from the conduct of teachers and students engaged in public education. One of the central functions of the platform is to extract value from the data traces that are produced from those using the software (Srnicek 2017), and this, of course, underpins the educational use of such systems, where the activities of teachers and learners are 'rendered as observable and measurable behavior' and 'turned into transactional data' (Decuyper *et al.* 2021, p6). It is important to see this, not as a by-product of the platform's involvement in educational activity, but a core function. Perrotta suggests 'digital platforms turn data into intangible assets, channelling them along financial circuits and subjecting them to capitalisation' (Perrotta *et al.* 2021, p107), identifying value-extraction as a guiding rationale for the ways such systems work. This doesn't mean that the other forms of learning in educational platforms are necessarily superficial or insincere, but rather that the driving force of the platform design appears to be for the benefit of the learning

undertaken by platform owners. As Perrotta et al. further contend in an examination of the functioning of the Google classroom platform, 'Google's extractive tendencies in education acquire a mission-critical purpose' (Perrotta et al. 2021, p107). Thus the 'division of learning' might be seen as an important indicator of which form of learning is most privileged in relation to educational platforms, and where concerns about inequality might be situated. This precise division is identified in the functioning of Google classroom, with, on the one hand a 'façade' comprised of a 'readable and actionable interface engineered to facilitate a particular form of coded pedagogy', and on the other, an 'underlying text of algorithmic surveillance, which is hidden from view and accessible only to Google's epistemic elites' (Perrotta et al. 2021, p107).

The 'division of learning' (Zuboff 2019) suggests that the benefits of this hidden form of privileged learning are considerable. However, the question remains whether students also benefit from this relationship, for which there is ample empirical research attempting to link platform use to 'learning outcomes' (see Passey & Higgins 2011). However, it is worth emphasising that any sense of an 'enhanced learning' facilitated by data-driven systems is highly contested, precisely because of the ways platforms work to restructure education in profound ways. Educational platforms 'make specific forms of education visible, knowable, thinkable, and, ultimately, actionable, in ways that reformat, redo, restructure, and reconceive what education is or could be about' (Decuyper *et al.* 2021, p7). One example of this tendency is the rise in a culture of surveillance and performativity in education (for example, see Manolev *et al.* 2019), where platform software, rather than serving as a neutral conduit for connecting students, teachers, and resources, acts to discipline and control users towards particular kinds of behaviours. Referring to the same educational platform as Manolev et al. (2019), Williamson suggests 'ClassDojo amply illustrates the combination of behavioural data extraction, behaviourist psychology and monetization strategies that underpin surveillance capitalism' (2019). In reference to the concept of the 'division of learning' specifically, Williamson further contends that '[t]he prospect of smart machines as educational engines of social reproduction should be the subject of serious future interrogation.' (Williamson 2019). The supposed neutrality of platforms, therefore, needs to be questioned:

[r]ather than simply hosting content, activity and interactions, platforms set the stage for how all this can come about, and which specific sorts of content, activities and interactions are possible and permissible (and which not). (Decuyper *et al.* 2021, p5)

Perrotta et al. identify this influence in Google classroom specifically as a 'bidirectional algorithmic pedagogy' (2021, p109), where the technology is not only structured in ways that enable machines to learn from extracted behavioural data, but designed to manage and automate the conduct of teachers and students. That this managed behaviour is then itself datafied and used to extract value, warrants, for Perrotta et al., the acknowledgement

of a 'shadow labour' (2021, p106) undertaken by students and teachers for the benefit of platform owners.

Focusing usefully on the student experience in educational software platforms, Grimaldi and Ball (2021) suggest paradoxical notions of value. On the one hand 'the value of achievement, of competence and realisation, and of the capacity to address a task and an assignment' becomes possible, but on the other 'a relative and analytical value that depends on the positioning of datafied performance within a hierarchical order' conditions the experience (Grimaldi and Ball 2021, p123). Thus, '[o]n one side, the learner is invited to be active and enterprising' while [o]n the other side, this involves continuous work on and the policing and solicitation of an online self through the datafied visualisation of its value' (Grimaldi and Ball 2021, p123-4). The result of this is a paradoxical state of learner subjectivity, where students are 'invited to be different' at the same time as 'continuously "normalised" through performance data' (Grimaldi and Ball 2021, p125).

In these ways one might see the 'division of learning' as resonating particularly well with the rise of datafication in education, where, not only are platform owners imbued with substantial advantages through the extraction of data, but students and teachers are left with impoverished, laborious, and contradictory educational experiences, as teaching and learning becomes increasingly managed by platform software. However, while there may well be an extreme 'division of learning' between the epistemic elites who own the platforms and the users who merely undertake activities teaching and learning through them, understanding the specific character of this learning would be essential to any attempt to address the supposed inequality.

Refocusing the 'division of learning' on education

Examining Zuboff's (2019) discussion of the 'division of learning' in detail reveals a tendency to focus on the perceived inequality resulting from the division, rather than on the nature of the learning presumed to be taking place. As appears to be *de rigueur* for discussions of the influence of the internet on society and its knowledge, Zuboff draws comparison with the Gutenberg press, suggesting that through the 'division of learning' society 'reverts to a pre-Gutenberg order' where learning is 'captured by a narrow priesthood of privately employed computational specialists, their privately owned machines, and the economic interests for whose sake they learn' (p189). Following the analogy that moveable type printing supposedly dismantled barriers to elite and exclusionist centres of power, Zuboff's (2019) argument here appears to be that the 'shadow text', and the abilities to understand, process, and benefit from its contents, need to be distributed more fairly in society. The solution is therefore not one which seems necessarily concerned with the practices of data exploitation itself, rather only with the ways it is excluded from the wider population.

This appears to overlook, not only a more thorough interrogation of the forms of 'learning' undertaken by the 'surveillance capitalists', but also the question of whether such practices are indeed desirable for wider society. Indeed, it is the lack of public access to the techniques of 'data science' that seems to be Zuboff's primary concern, particularly in the suggestion that an:

unprecedented concentration of knowledge produces an equally unprecedented concentration of power: asymmetries that must be understood as the *unauthorized privatisation of the division of learning in society*. (Zuboff 2019, p191, emphasis original)

Ultimately, Zuboff's (2019) discussion of the 'division of learning' appears to distil a critical perspective into a struggle for privacy and the ownership of data, rather than a more profound consideration of how data-driven approaches are entrenching problematic views of society and how individuals learn within it. This ends up endorsing the power of data-fuelled behaviourism by assuming that it is indeed a new and authoritative way of knowing and shaping society.

In terms of proffering solutions for the overcoming of the 'division of learning', at times Zuboff appears fatalistic: 'every attempt at understanding, predicting, regulating or prohibiting the activities of surveillance capitalists will fall short' (2019 p193). However, such an impasse may result, at least in part, from a lack of attention to the forms of learning assumed to take place across the 'division of learning'. Zuboff labels the knowledge elites 'the new priesthood' (2019, p186): 'the 10,000 or so professionals on the planet who know how to wield the technologies of machine intelligence to coax knowledge from an otherwise cacophonous data continent' (Zuboff 2019, p188). However, rather than probe the substance or efficacy of this knowledge, the main contention seems to be that there are not enough of experts: '[t]he huge salaries of the tech firms have lured so many professionals that there is no one left to teach the next generation of students' (Zuboff 2019, p188). At the same time as condemning the inequalities created through the 'division of learning', this appears to end up valorising the specialist expertise employed by the 'surveillance capitalists', rather than seeking to challenge it. Zuboff certainly poses pertinent critical questions, for example: 'which people, institutions, or processes determine who is included in learning, what they are able to learn, and how they are able to act on their knowledge' (Zuboff 2019, p180). However, democratising the ability to gain skills in 'data science', and therefore to understand and extract value from the 'shadow text' of behavioural data, would seem likely to create new frontiers and barriers, complicating the 'division of learning' rather than dispelling it.

Turning to education theory may be one way to better theorise the forms of learning assumed in concepts such as the 'division of learning', and suggest more robust ways of challenging the assumptions about human behaviour that underpin the rise of data-driven capitalism. Firstly, the concept of 'learnification' (Biesta 2005; 2009; 2013; 2015) provides

a useful critique of the dominance of 'the language of learning' (Biesta 2005, p54), particularly over notions of formal teaching and expertise in general. Primarily directed at the field of education, 'learnification' refers to a wholesale reorientation of institutional and pedagogical arrangements around the idea of a centralised, self-directing, and individual 'learner' (Biesta 2005; 2009; 2013; 2015). However, to borrow a phrase from Zuboff (2019) quoted earlier, we might say that the language of learning has 'burst through the walls of the classroom to overwhelm our daily lives'. Not only are there calls for a 'learning society', in which the figure of the 'lifelong learner' (Yang and Valdés-Cotera 2001) is required to constantly update their expertise in order to remain relevant and employable in a swiftly shifting marketplace (and the ability to analyse data clearly figures centrally in this vision), but machines are now also 'learning' in supposedly powerful new ways that society needs to pay increasing attention to (Domingos 2017; Alpaydin 2016). The 'division of learning' might be seen as part of this wider trend in the way that it positions expert data-analytic 'learning' as the privileged practice of our times, and a desirable condition to which wider populations should aspire. The problem here is that such a focus on learning, as Biesta (2005; 2009; 2013; 2015) warns, tends to be blind to questions about how aspirations are negotiated and established, and what kind of power underlies such processes. Further, if we consider a straightforward application of the 'division of learning' in educational contexts, as explored above, one might see 'learning' as constituting both sides of the divide; students 'learning' (often in 'personalised' and self-directed ways) through platforms on one side, and the owners of the platforms 'learning' from user-data on the other. Despite the notion that the latter is a much more privileged and beneficial form of learning, it is also worth recognising the entrenchment of particular behaviouristic ideas about learning across both sides of the divide. Radical behaviourist ideas are being reintroduced into educational practices through the use of data-driven platforms (also see Knox et al. 2020), while at the same time underpinning the techniques of data extraction undertaken by the platform owners.

Secondly, in order to move beyond the limited purview of 'learning', Biesta differentiates 'education' as a practice that is defined by *purpose*: 'when we are engaged in decision making about the direction of education we are always and necessarily engaged in value judgements – judgements about what is educationally desirable' (Biesta 2009, p35). Learning is further distinguished from education by suggesting that the former is 'basically a process term' denoting 'processes and activities but is open – if not empty – with regard to content and direction' (Biesta 2009, p39), while the latter 'always raises the question of its purpose' (Biesta 2015, p84). It is this idea of education that might usefully develop the concept of the 'division of learning', by suggesting a route to overcoming data-driven inequality, not through simply widening access to the behaviourist techniques of machine learning, but rather through concerted ongoing efforts to define collective *purposes* for working with data. Biesta distinguishes between three domains of educational purpose: *qualification*, the domain of skills acquisition; *socialisation*, the ways individuals become

part of society; and *subjectification*, which concerns how individuals become autonomous (2013; 2015). While one might see the 'division of learning' as straightforwardly concerned with the domain of qualification, and the important questions of how accessible those qualifications are, and who gets to define them, there is little to discern in the way of engagement with a social or subjective purpose to becoming skilled in the techniques of data analytics and machine learning. For all of Zuboff's detailed examination of the inequality, discrimination, and injustice produced through datafication, there is little to suggest a clear social objective or rationale for democratising access to the 'shadow text', other than to reclaim 'the larger promise of an information capitalism aligned with a flourishing third modernity' (2019, p194)².

Engaging with the domain of socialisation would require considerations of collective purposes for the mass deployment of data-driven technologies. As has already been noted, recent years have seen the proliferation of opportunities to undergo technical training in machine learning and other analytic techniques, often in the form of open-access, self-directed online courses, justified through discourses of impending technological unemployment, and accompanied by criticisms of formal schooling as antiquated in comparison (Knox 2019). However, there appears to be little in this drive for technical training beyond the idea of future-proofed employment. Addressing the social purpose of datafication might be achieved through more recognition of the inherent politics of data (Bigo et al. 2019) and the agonistic and contested ways in which its attendant technologies are developed (Crawford 2016), rather than its assumed objectivity or neutrality. An *educational* approach, as opposed to straightforward technical training, might therefore frame the emerging discipline of 'data science', for example, as 'a form of political action' (Green 2019, p7). Green further suggests:

'[d]ata scientists must recognize themselves as political actors engaged in normative constructions of society and, as befits political work, evaluate their work according to its downstream material impacts on people's lives' (Green 2019, p7).

Such an approach would seem to direct attention towards forging tangible connections with society, as opposed to merely extracting data from it. In a similar way, calls for greater 'big data literacy' (D'Ignazio and Bhargava 2015), 'personal data literacies' (Pangrazio & Selwyn 2019), or 'critical data literacy' (Sander 2020) have made the case for widening public understanding of the social impacts of datafication. Related to this work, a range of accessible tools and strategies have been developed to promote broader civic awareness of issues such as privacy, surveillance, inequality, and discrimination as they relate to the use of data-driven technologies (for examples see Yates et al. 2020 and Sander 2020). While the suggestions in this paper align with much of this work in terms of its interest in

² This commitment to capitalism, albeit a more inclusion form, is suggested elsewhere too: '[I]et it be an insistence that raw surveillance capitalism is as much a threat to society as it is to capitalism itself' (Zuboff 2019, p193), and criticised in detail by Morozov (2019).

developing the critical understanding of datafication, there remain important questions about how collective purposes for the development of 'data science' might be achieved. In developing a useful typology for the plethora of online tools and resources that have been developed in recent years, Sander (2020) found not only a diverse range of forms, but also categorical differences to the kinds of criticality such materials presented. Further, and surprisingly for Sander, 'the sample did not include any resources from governmental or public service institutions, nor from traditional educational avenues and only very few efforts on the part of academia' (2020, p9), most being produced by non-government organisations. Given the central role of formal public education in society, such an institution would seem to be the ideal place to consider collective (social) purposes for the development of 'data science'. However, as Raffaghelli et al. (2020) discuss in the context of higher education, there are often underlying tensions between the institutional desire to produce functional benefits through the deployment of data-driven technologies, and a concern for 'ethical' and 'fair' data cultures.

Biesta's final domain of subjectification, perhaps being the most profound, emphasises transformation and liberation as a fundamental purpose of education:

subjectification has an orientation towards emancipation – that is, towards ways of doing and being that do not simply accept the given order, but have an orientation towards the change of the existing order so that different ways of doing and being become possible. (Biesta 2013, p6)

While the concept of the 'division of learning' appears committed to democratising, and therefore disseminating and replicating the existing techniques of 'big data' analysis, there appears to be little opportunity for the kind of weighty individual emancipation referred to here, or the prospect of overcoming the established orders of data exploitation. As discussed above, the public pedagogy of 'critical data literacy' (Sander 2020) suggests tangible routes through which individuals might re-think their relationships with data, and perhaps emerge as critically-minded citizens in the era of datafication. However, significant questions remain about, not only the diversity and fragmentation of the online resources produced for this purpose, but also the capacity for formal educational institutions to take up greater responsibility in providing the kind of education that can provide an alternative to the dominant rationales of 'data-skills' acquisition.

Conclusions

While Zuboff's (2019) 'surveillance capitalism' is undoubtedly a seminal text in the emerging understanding of the societal impacts of datafication, the question of how educationalists might relate to this work is less clear. This paper has attempted to suggest two possible responses. Firstly, where data-driven platforms 'are in the process of significantly transforming the educational sector' (Decuyper *et al.* 2021, p2), this paper

has suggested the 'division of learning' to be a productive concept through which to examine concerns about increasing privatisation and inequality in education. Here the methods through which private companies derive commercial benefits from platform data, and the labour of the students, teachers, and administrators that produce it, appears to far outweigh the benefits students experience from 'learning' with the use of data-driven platforms. Critical attention to issues of accountability are therefore needed, echoing Perrotta et al's recent concerns in relation to Google's expansion as a provider of mainstream education:

the role of Google (and other platform owners operating in a similar capacity) as a mere data operator needs to be challenged and, ideally, replaced by its unequivocal categorisation as an educational actor with considerably higher responsibilities. (Perrotta et al. 2021, p109)

Secondly, this paper has argued that, for all of Zuboff's criticism of 'surveillance capitalism' as operating 'by predicting, influencing, and controlling human behaviour' (2019, p189), the emphasis often appears to be on widening access to such methods, rather than questioning data-driven behaviour-manipulation itself. In order to challenge this view, this paper has suggested the need for perspectives from educational research that emphasise collective purposes to the project of education, as opposed to a narrow and individualistic idea of learning as 'training' in data skills. Drawing from an already flourishing domain of public pedagogy around 'critical data literacy', formal institutional education might therefore adopt a more central role in developing a 'data science' with social and emancipatory *purpose*; one that is foundationally engaged in collective justificatory practices about how and why particular technologies and techniques should be developed, as well as attentive to the kind of individual citizen-subjects produced through such methods.

References

- Alpaydin, E. 2016. *Machine Learning: the new AI*. Cambridge: MIT Press. ^[1]_{SEP}
- Biesta, G.J.J. (2005). Against Learning. Reclaiming a Language for Education in an age of Learning. *Nordisk Pedagogik* 25(1): 54–66. <https://orbilu.uni.lu/handle/10993/7178>
- Biesta, G.J.J. (2009). Good education in an age of measurement: on the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability* 21:33–46 <https://doi.org/10.1007/s11092-008-9064-9>
- Biesta, G.J.J. (2013). Interrupting the Politics of Learning. *Power and Education* 5(1): 4–15. <https://doi.org/10.2304/power.2013.5.1.4>

- Biesta, G.J.J. (2015). What is Education For? On Good Education, Teacher Judgement, and Educational Professionalism. *European Journal of Education*, 50(1): 75-87
<https://doi.org/10.1111/ejed.12109>
- Bigo, D., Isin, E., & Ruppert, E., 2019. *Data Politics: Worlds, Subjects, Rights*. London; Routledge.
- 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.
<https://doi.org/10.1080/13562517.2019.1698540>
- Bulger, M. (2016). Personalized Learning: The Conversations We're Not Having. *Data & Society working paper*.
https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf
- Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. New York: McGraw-Hill.
- Decuyper, M., Grimaldi, E. & Landri, P. (2021). Introduction: Critical studies of digital education platforms, *Critical Studies in Education*, 62:1, 1-16.
<https://doi.org/10.1080/17508487.2020.1866050>
- D'Ignazio, C., & Bhargava, R. (2015). Approaches to Building Big Data Literacy. In Bloomberg Data for Good Exchange 2015. New York, NY, USA. https://dam-prod.media.mit.edu/x/2016/10/20/Edu_D'Ignazio_52.pdf
- Domingos, P. (2017). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. London; Penguin^[L]_{SEP}.
- Crawford, K. 2016. Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics. *Science, Technology, & Human Values* 41(1): 77-92.
<https://doi.org/10.1177%2F0162243915589635>
- Green, B. 2019. Data Science as Political Action: Grounding Data Science in a Politics of Justice. [arXiv:1811.03435](https://arxiv.org/abs/1811.03435).
- Grimaldi, E. & Ball, S.J. (2021). Paradoxes of freedom. An archaeological analysis of educational online platform interfaces, *Critical Studies in Education*, 62:1, 114-129.
<https://doi.org/10.1080/17508487.2020.1861043>
- Knox, J., Williamson, B. & Bayne, S. (2020). Machine Behaviourism: future visions of 'learning' across humans and machines. *Special Issue of Learning Media and Technology: Education and technology into the 2020s: speculative futures*, 45:1, 31-45. <https://doi.org/10.1080/17439884.2019.1623251>

- Knox, J. (2019). The 'Creative, Problem-Solving Entrepreneur': alternative futures for education in the age of machine learning? In M. Peters, P. Jandrić, & A. Means (eds.) *Education and Technological Unemployment*, Springer, 297-312.
https://doi.org/10.1007/978-981-13-6225-5_19
- Manolev, J., Sullivan, A. & Slee, R. (2019). The datafication of discipline: ClassDojo, surveillance and a performative classroom culture, *Learning, Media and Technology*, 44:1, 36-51, <https://doi.org/10.1080/17439884.2018.1558237>
- Morozov, E. (2019). Capitalism's New Clothes. *The Baffler*.
<https://thebaffler.com/latest/capitalisms-new-clothes-morozov>
- Pangrazio, L., & Selwyn, N. (2019). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, 21(2), 419–437. <https://doi.org/10.1177/1461444818799523>
- Passey, D. & Higgins, S. (2011). Learning platforms and learning outcomes – insights from research, *Learning, Media and Technology*, 36:4, 329-333,
<https://doi.org/10.1080/17439884.2011.626783>
- Poell, T. and Nieborg, D. and van Dijck, J. (2019). Platformisation. *Internet Policy Review*, [online] 8(4). <https://doi.org/10.14763/2019.4.1425>
- Perrotta, C., Gulson, K.N., Williamson, B. & Witzemberger, K. (2021). Automation, APIs and the distributed labour of platform pedagogies in Google Classroom, *Critical Studies in Education*, 62:1, 97-113, <https://doi.org/10.1080/17508487.2020.1855597>
- Raffaghelli, J.E., Manca, S., Stewart, B., Prinsloo, P. & Sangrà, A. (2020). Supporting the development of critical data literacies in higher education: building blocks for fair data cultures in society. *International Journal of Educational Technology Higher Education*, 17, 58. <https://doi.org/10.1186/s41239-020-00235-w>
- Sander, I. (2020). What is critical big data literacy and how can it be implemented?. *Internet Policy Review*, 9(2). <https://doi.org/10.14763/2020.2.1479>
- Selwyn, N. (2013). Discourses of digital 'disruption' in education: a critical analysis. Paper presented to Fifth International Roundtable on Discourse Analysis, City University, Hong Kong, May 23-25, 2013
- Selwyn, N., Hillman, T., Eynon, R., Ferreira, G., Knox, J., Macgilchrist, F. & Sancho-Gil, J.M. (2020). What's next for Ed-Tech? Critical hopes and concerns for the 2020s, *Learning, Media and Technology*, 45:1, 1-6,
<https://doi.org/10.1080/17439884.2020.1694945>

- Sutton and Barto. (2018). *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge: MIT Press. <https://drive.google.com/file/d/1xeUDVGWGUUv1-ccUMAZHJLej2C7aAFWY/view>
- van Barneveld, Angela, Kimberly E. Arnold, and John P. Campbell. 2012. Analytics in Higher Education: Establishing a Common Language. *ELI Paper 2012 (1)*: 1–11.
- van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connectiveworld*. Oxford University Press
- Williamson, B. (2019). Learning from Surveillance Capitalism. *Code Acts in Education*. <https://codeactsineducation.wordpress.com/2019/04/30/learning-from-surveillance-capitalism/>
- Williamson, B. & Hogan, A. (2020). Commercialisation and privatisation in/of education in the context of COVID-19. Education International. Available https://issuu.com/educationinternational/docs/2020_eiresearch_gr_commercialisation_privatisation?fr=sZDJkYjE1ODA2MTQ
- Williamson, B, Eynon, R & Potter, J 2020, 'Pandemic politics, pedagogies and practices: Digital technologies and distance education during the coronavirus emergency', *Learning, Media and Technology*, 45 (2): 107-114. <https://doi.org/10.1080/17439884.2020.1761641>
- Yang, J. & Valdés-Cotera, R. (Eds) (2001) *Conceptual Evolution and Policy Developments in Lifelong Learning*. Hamburg: UNESCO Institute for Lifelong Learning.
- Yates, S., Carmi, E., Pawluczuk, A., Lockley, E., Wessels, B. & Gangneux, J. 2020. *Me and My Big Data Report 2020: Understanding citizens' data literacies: thinking, doing & participating with our data*. <https://www.liverpool.ac.uk/media/livacuk/research/heroimages/Me-and-My-Big-Data-Report-1.pdf>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The fight for a human future at the new frontier of power*. London: Profile Books