Please cite the Published Version

Additional Information: This is an Open Access article published in Journal of Education Policy.
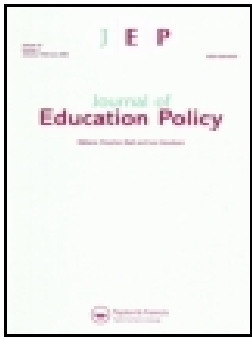
Becoming information centric: the emergence of new cognitive infrastructures in education policy

Sam Sellar & Kalervo N. Gulson

Routledge
Taylor & Francis Group

Becoming information centric: the emergence of new cognitive infrastructures in education policy

Sam Sellar [a] and Kalervo N. Gulson [b]

aDepartment of Childhood, Youth and Education Studies, Manchester Metropolitan University, Manchester, UK; bSchool of Education, University of New South Wales, Sydney, Australia

ABSTRACT

New cognitive infrastructures are emerging as digital platforms and artificial intelligence enable new forms of automated thinking that shape human decision-making. This paper (a) offers a new theoretical perspective on automated thinking in education policy and (b) illustrates how automated thinking is emerging in one specific policy context. We report on a case study of a policy analysis unit ('The Centre') in an Australian state education department that has been implementing a BI strategy since 2013. The Centre is now focused on using BI to support complex decision making and improve learning outcomes, and their strategy describes this focus as becoming 'information centric'. The theoretical framework for our analysis draws on infrastructure studies and philosophy of technology, particularly Luciana Parisi's recent work on automated thinking. We analyse technical documentation and semi-structured interview data to describe the enactment of a BI strategy in The Centre, with a focus on how new approaches to data analytics are shaping decision-making. Our analysis shows that The Centre is developing a cognitive infrastructure that is already creating new conditions for education policy making, and we conclude with a call for research designs that enable pragmatic exploration of what these infrastructures can do.

Introduction

The aim of this paper is to explore the emergence of new forms of data-driven decision making in the work of a policy analysis unit ('The Centre') within an Australian state education department. The Centre is using a cloud platform to implement their Business Intelligence (BI) strategy and is developing predictive analytics and modelling to pursue its goal of becoming 'information centric'.[1] BI solutions comprise multiple elements, from user interfaces that enable easy analysis and visualization of data, to data warehouses and cloud computing platforms with embedded artificial intelligence capabilities. The work of The Centre can be seen as part of an intensification of computational approaches to public policy work over the past decade (e.g. Gilbert et al. 2018), with an increasing role for data science in education policy (Williamson

CONTACT Sam Sellar ✉ s.sellar@mmu.ac.uk 🔲 Manchester Metropolitan University, Brooks Building, 53 Bonsall Street, Manchester M15 6GX, UK

2016). Data-driven evaluations and decision-making have been replacing embodied forms of professional knowledge since the 1970s, when new modes of accountability were introduced as part of the move towards New Public Management (NPM). Critical analyses of data-driven governance often focus on the constitutive power of statistics and the spread of instrumental rationality and performativity. We recognise that computational approaches to policy can delegitimise and replace other forms of knowledge and practice (Pasquale 2019), and there are significant concerns about how these approaches enable technology companies to become new players in education policy (Williamson 2018). However, in this paper we take a different view. Instead of using established values to evaluate changes in education policy processes produced by datafication and automation, we offer a speculative account of how data science and automation can create new values in education policy contexts.

We aim to contribute to the emerging literature that examines the use of data infrastructures in educational governance (e.g. Anagnostopoulos, Rutledge, and Jacobsen 2013; Hartong and Förschler 2019; Ratner and Gad 2018) by theorising and illustrating the new possibilities for thought created by these infrastructures. We take this perspective in response to the challenge presented by Deleuze (1995) in his writing on control societies. Deleuze argued that we have moved from disciplinary forms of power ('moulds') to more open, digital forms of control ('modulation') that shape subjectivities through new media and computational technologies. The production of a subject who seeks to continually increase its value through education and other experiences is a key technology of neoliberal governance (Feher 2009), and measurement and comparison have become powerful instruments of control. This perspective has provided useful insights into new modes of educational governance and the intensification of instrumental rationalities, including in some of our own previous work (Sellar 2015). However, less work has been undertaken to invent the new conceptual tools that Deleuze also argued were needed to engage pragmatically with the mechanisms of control societies. We do not employ a Deleuzian framework in this paper, but we do take up his challenge by employing a speculative theoretical framework to analyse the emergence of what Luciana Parisi (2017) calls *automated thinking*.

In her recent work, Parisi has pursued 'an alternative approach to reasoning that accounts for the inferential potential of automated computation' (Parisi 2016a, 480). Parisi (2016b, paragraph 9) argues that the conflation of automation and capitalist logics in critiques of instumental rationality blinds us to other possibilities of automation. She argues that '[t]he critique of the instrumentalisation of reason according to which automation and the logic of capital are equivalent needs to be re-visited in view of rapid transformations of automation today'. This critique has been influential in critical education studies and has lead to the view that data-driven technologies are part of a capitalist logic that is detrimental for public education, particularly due to their association with neoliberalisation and privatisation. There is a large body of scholarship that explicitly or implicitly rejects the production and use of quantitative data to drive policy and practice because it undermine the values and purposes of humanist education. This work includes critiques that highlight the connection of quantitative approaches and big data to the continuation of inequities in education policy, including in relation to race (Gillborn, Warmington, and Demack 2018). While recognising the importance of these critiques, we are also curious about new possibilities created by the the algorithms that enable automation. Parisi (2016b, paragraph 8) argues that it is

the transformation of the logic of the technical machine itself and thus of a philosophy of computation that needs to be unpacked, disarticulated and reconstructed so as to allow for a critique of capital that is not immediately a negation of automation and its possibilities for thinking.

In this paper, we examine new infrastructures that are being utilised for cognitive insight and, potentially, for cognitive engagement in education policy and governance (Davenport and Ronanki 2018), in order to identify new possibilties for thinking about, and with, these infrastructures.

We present an exploratory case study of 'The Centre', which is at the forefront of a set of developments in Australian schooling that includes the implementation of interoperability standards to enable the collection, analysis and use of data across multiple scales, jurisdictions and platforms (Gulson and Sellar 2019). Since 2013, The Centre has been implementing a BI strategy to support strategic and tactical decision making, and the improvement of student outcomes, through new modes of data analysis. The Centre's BI strategy outlines its aim to become 'information centric' through 'significant growth both in our usage of information and understanding of what we can do with it'. Moving from its previous 'data centric' approach to become 'information centric' involves moving from accessing and analysing data for specific purposes (e.g. generating a report on assessment outcomes) to embedding data analytics across all organisational decision making, in order to improve learning outcomes. A roadmap published in the BI strategy document outlines a plan to develop predictive analysis and modelling capabilities by 2019, before moving towards optimisation, at which point data analytics will ideally enable maximally efficient and effective education.

Becoming information centric requires new computational capacities. The Centre is using the Microsoft Azure cloud computing platform – which includes artificial intelligence capabilities – to create new data warehousing and analytics solutions.[2] Two data scientists have also been hired to undertake exploratory data analyses, along with other staff who have information science backgrounds. The use of algorithms to analyse big data sets, and the aim to embed near real-time data analytics in the day-to-day operations of the department, indicates a move towards automating aspects of decision-making. While The Centre is still in the early stages of developing these capacities, its ambition and progress to date arguably position it as a leader in this area among Australian education departments and potentially among education systems globally.

Our paper is divided into four remaining sections. In the next section, we provide a fuller discussion of our theoretical framework, drawing on literature from infrastructures studies, software studies and philosophy of technology. We then outline our methodology and introduce our larger study of educational data infrastructure from which we have drawn this case study. Following detailed analysis of the case, which includes data from interviews and technical documentation, we conclude by drawing out the distinctive insights pertaining to policy, values and uncertainty afforded by our theoretical framework.

Theoretical framework

Planetary-scale computation and infrastructure space

The introduction of cloud computing platforms like Microsoft Azure integrates education policy making into what Bratton (2014) calls 'planetary-scale computation', arguing

that '[i]nstead of seeing the various species of contemporary computational technologies as so many different genres of machines … we should instead see them as forming the body of an accidental megastructure' (1). Bratton (2015) describes this structure as a *stack* with multiple layers, including 'the computing and transmission hardware on which Stack software depends, such as data centers, transmission cables, geosynchronous satellites, and wireless technologies' (369), and 'platforms, such as Google and Amazon, which provide services to their federated *Users* through the applications they directly manage or those they support' (369). In education, cloud platforms are beginning to disrupt established practices and create new sources of value (Komljenovic 2019; Robertson 2019).

Planetary-scale computation is an increasingly integrated and multi-layered digital infrastructure space (Easterling 2014). Software and infrastructure studies have shown how algorithms are embedded in networks of technical devices and urban infrastructures that are now connected globally (e.g. Easterling 2014; Kitchin and Dodge 2011; Thrift 2014). Infrastructure is commonly perceived to be material – roads, rails, pipes and wires – but digital infrastructures are also built from interoperable networks of computer software, including abstract entities such as algorithms. An algorithm is a diagram or description of a procedure for performing an action, and it operates on information. Easterling argues that

> [i]nfrastructure space has become a medium of information. The information resides in invisible, powerful activities that determine how objects and content are organised and circulated. Infrastructure space, with the power and currency of software, is an operating system for shaping the city. (Easterling 2014, 13)

Infrastructure space has a *disposition* that emerges, in part, from the actions of algorithms. As Goffey (2008) observes, '[a]lgorithms act, but they do so as part of an ill-defined network of actions upon actions, part of complex power-knowledge relations, in which unintended consequences, like the side effects of a program's behaviour, can become critically important' (19). The layering and interaction of algorithms in infrastructure space produces distributed, emergent forms of cognition that can have powerful effects (e.g. surveillance of populations using facial recognition technology) by using, reproducing and transforming information (e.g. images of faces that become data points as 'facial signatures').

Nonconcious cognition and automated thinking

Cognition can be defined as an informational process that does not necessarily involve consciousness. Hayles (2014) distinguishes between thinking as a property of conscious entities and cognition as an informational process that occurs without consciousness, but, nonetheless, with intention. Cognition differs from material processes (e.g. geology) insofar as it involves emergence, adaptation or complexity, and it differs from thinking insofar as it does not require consciousness. As Hayles (2014) explains, 'we can say that all thinking is cognition, but not all cognition is thinking' (201). Nonconscious cognition 'operates across and within the full spectrum of cognitive agents: humans, animals, and technical devices' (Hayles 2014, 202). Moreover, when embedded in the networked technical devices that constitute infrastructure space, 'the cognitive nonconscious also

carries on complex acts of interpretation, which syncopate with conscious interpretations in a rich spectrum of possibilities' (215).[3] Nonconcious cognition inhabits different temporalities to human thought, creating new possibilities for integration with, and exploitation of, our thinking. As Hayles writes, '[o]ne of the ways in which the cognitive nonconscious is affecting human systems . . . is opening up temporal regimes in which the costs of consciousness become more apparent and more systemically exploitable' (212). Consider, for example, the advantages gained by automated algorithms in stock market trading in comparison with human traders. As we automate more and more of the cognitive load of modern life, we inescapably change the cognitive ecology in which human thinking evolves and functions (Bakker 2015).

In this context, Parisi (2017) argues that developments in the algorithms used for machine learning have created new conditions for automated thinking: '[s]ince 2006, with deep learning algorithms, a new focus on how to compute unknown data has become central to the infrastructural evolution of artificial neural networks' (7).[4] Historically, AI research has explored a number of different techniques, from Good Old-Fashioned AI, which is a symbolic, rule-based approach, to machine learning in which systems are trained on large data sets (Boden 2016). A recent meta-analysis of papers in the field of AI has showed substantial shifts toward machine learning, deep learning (machine learning using neural networks) and reinforcement learning (training neural networks using punishments and rewards) (Hao 2019). Reinforcement learning has received significant attention since the success of Google's AlphaGo and AlphaZero algorithms.[5] While deep learning approaches are likely to be replaced by another paradigm, just as previous approaches have been, they have taken over from knowledge-based approaches. Theories, logic and rules are being replaced by data, networks and learning.

We are still a long way from General AI, which is the imaginary popularised in much science fiction and recent debates about the existential risks of AI. But the other common view of AI as simply inductive data processing, which unthinkingly produces generalisations from big data sets, may also misrepresent the action of these algorithms. This view is exemplified, for example, in Lyotard's (1984) influential description of the emergence of performativity within information systems, which he defines as the optimisation of the relation between inputs and outputs. Lyotard's argument, along with work of other thinkers like Deleuze and Guattari (1994), offers a critical view about the 'automation of decision, where information processing, computational logic, and cybernetic feedbacks replace the very structure, language, and capacity of thinking beyond what is already known' (Parisi 2017, 13; see also Sellar and Thompson 2016). However, the introduction of new machine learning approaches raises new questions about the possibilities for thought.

The uncertainty of machine learning

Machines are more than efficient tools for instrumental reasoning. Parisi (2017, paragraph 13) argues for conceiving machine cognition as

> . . . not simply a cybernetic form aiming at steering decisions towards the most optimal goals. Instead, operating systems are computational structures defined by a capacity to calculate infinities through a finite set of instructions, changing the rules of the game and extending their capacities for incorporating as much data as possible. These systems are not simply tools of or for calculation.

Machine learning differs from earlier approaches to AI due to the increased volume and variety of data available to train algorithms today, but also because in artifical neural networks 'algorithms do not just learn from data, but also from other algorithms, establishing a sort of meta-learning from the hidden layers of the network' (Parisi 2017, 8). While much of the political work on bias and AI has been concerned with the idea that bias is built into both data sets and algorithms (Campalo et al. 2017), Parisi provides a more complicated, and more tenuous, theorisation of machine-based decision making.

Artificial neural networks do not necessarily calculate all possibilities within a system and then make the most optimal decisions. These algorithms can also make inferences about the most optimal decision in a given situation while remaining ignorant of a larger set of indeterminate possibilities. Indeed, building more powerful cognitive machines only increases the horizon of indeterminacy (Chatain 2006). Machine learning thus entails a move beyond deduction and induction towards abduction: 'the process of inferring certain facts and hypotheses to plausibly explain some situations and thereby also discover some new truths' (Parisi 2016a, 479). The possibilities of deduction and induction are already contained in premises or empirical cases, whereas abduction involves creative uncertainty in machine learning. In contrast to rule-based machine cognition, such as IBM's Deep Blue chess program, deep learning algorithms like AlphaZero are trained on the data they are fed and, in the process, can learn about this training.

> [D]eep-learning algorithms do not just learn from use but *learn to learn* about content- and context-specific data (extracting content usage across class, gender, race, geographical location, emotional responses, social activities, sexual preferences, music trends, etc.). This already demarcates the tendency of machines to not just represent this or that known content, or distinguish this result from another. Instead, *machine learning engenders its own form of knowing: namely, reasoning through and with uncertainty*. (Parisi 2017, 8; emphasis added)

The 'knowing' described here references the relationship between deep learning algorithms and data – in deep learning approaches the data is unstructured and the algorithms provide this structure. Automoted thinking of this kind can be hypothetical and creative.

This perspective provides a generative alternative to critiques of datafication and automation that focus on instrumental rationalities and the detrimental effects of these rationalities for teaching and learning. In the analysis that follows, we focus on the creative uncertainty of nonconscious cognition and its syncopation with human thinking. Through an examination of The Centre's BI strategy and its implementation, we aim to reopen 'the question of how to think in terms of the means through which error, indeterminacy, randomness, and unknowns in general have become part of technoscientific knowledge and the reasoning of machines' (Parisi 2017, 3–4). We hold open the possibility that becoming information centric can produce new knowledge, values, and decision-making processes in education policy contexts. We do not argue that these developments are desirable or beneficial, but we do argue that it is neceessary to understand how automated thinking, which is emerging with data science and machine

learning, is beginning to change the possibilities for education policy and the governance of school systems.

Methodology

Our case study of The Centre is part of a multi-national comparative research project examining data infrastructures, policy mobility and network governance in education, with fieldwork undertaken in Australia, Canada, Japan and the United States between 2015 and 2017.[1] The fieldwork for this project involved over 70 face-to-face interviews with policy makers, educators, and managers and technicians in schools, school boards, education departments, and education technology companies. We asked questions about what administrative and performance data are being produced and collected in schools and systems, and what programs and platforms are being used to store, analyse and use these data. We selected The Centre, both for the broader study and for this paper, in order to provide a descriptive account of new phenomenon in a real-life context and 'to develop pertinent hypotheses and propositions for further inquiry' (Yin 2009, 5). An exploratory case study was thus appropriate. We followed Flyvbjerg (2006) in using an information-oriented selection approach that 'aims to maximize the utility of information from small samples and single cases', with cases purposively 'selected on the basis of expectations about their information content' (230). Early fieldwork conducted as part of the broader study indicated that The Centre was considered to have one of the most highly developed BI strategies in Australian education. As such, we expected that this case study would generate substantial useful information about the development of data infrastructures and new data analytics in education.

Our fieldwork was informed by Star (1999), who pioneered an ethnographic approach to infrastructure studies based on the 'idea that people make meanings based on their circumstances, and that these meanings would be inscribed into their judgements about the built information environment' (383). Star also observed that studying infrastructure is 'a call to study boring things' because '[m]any aspects of infrastructure are singularly unexciting' (Star 1999, 377). However, attention to the mundane can produce unusual insights. Arguing for the value of an ethnographic approach for understanding technoscientific worlds, Fischer (2009) contends that:

> [t]he complexities of our times require ethnographic skills. This is a matter of opening up simplified accounts, making accountability possible at different granularities, signposting the labyrinths of possible inquiries for their relevance, their points of no return, their conceptual reruns, acknowledging in a politics of recognition pebbles of resistance that destabilize easy theories (xiii).

In a field like education policy studies, where power and politics are inexorably intertwined, ethnography affords an emphasis on opening up, disrupting, and destabilizing accepted interpretations. The focus on granularity is important and our attention to the banal and mundane work of experimenting with new data infrastructures suggested the

---

need for narratives beyond those that emphasise how new modes of big data analytics intensify control, calculation and surveillance.

The data sources for this paper include four semi-structured interviews with five senior policy makers, technical staff and data scientists in The Centre, in addition to analysis of publicly available documents. Interview participants were purposively sampled on the basis of their involvement with developing and implementing The Centre's BI strategy. We conducted a single group interview with three senior staff in 2016, then three individual follow-up interviews in 2017 after analysing the initial interview and material that we had gathered. In the analysis that follows, we draw directly on the interviews with the BI manager and one of the data scientists, because these interviews provided the most relevant technical information. The BI manager provided a strategic view of the development of the BI strategy, The Centre's data analytic capacities and the policy work of The Centre, while the data scientist provided detailed explanations of their technical work. Publicly available documents were also sourced from The Centre's website following thorough exploration of the site, and these include the current BI strategy. To protect anonymity and confidentiality of The Centre and its staff, we have not provided the source URL for these documents and have used pseudonyms for the organisation and our participants. However, like all insider research, there are no guarantees that people could not be identified by other members of their respective organisations. These participants are policy savvy experts in the topics that were discussed in interviews and played a role in shaping the foci of the discussions and the perspectives that emerged (Ball 1994).

Case study

We were surprised by the ambition of The Centre's BI strategy when we first visited in 2016. Phase 1 of the strategy was implemented from 2013, and in 2015 Phase 2 commenced with a new focus on moving 'from being data centric to being information centric', which is premised on '[t]he recognition that Business Intelligence is a tool to assist us in making complex decisions, and not just a data access point'. Phase 2 of the BI strategy aims to change decisions and actions to improve student learning outcomes. The strategy describes 30 initiatives that will be supported by BI from 2016 to 2020, and in what follows we focus on two of the most 'mature' projects. These projects demonstrate the emergence of new data-driven modes of reasoning in The Centre's work and we make no claims about the insights they have generated. Rather, we are interested in theorising the potential implications of strategic developments intended to occur from 2019 onwards: predictive analytics, predictive modelling and optimization (Figure 1).

The Centre established new data warehousing and data analytics capabilities during Phase 1 of the strategy, enabled by Microsoft's Power BI and Azure cloud platform. The Centre established a data warehouse using the data storage capabilities of Azure and Power BI provides the user interface. Azure's 'Infrastructure as a service' (IaaS) layer dramatically reduces the time needed to generate reports by scaling computational resources on demand. Staff in The Centre also mentioned Azure's embedded AI services. For example, the BI manager explained that the potential for machine learning is built into the platform:
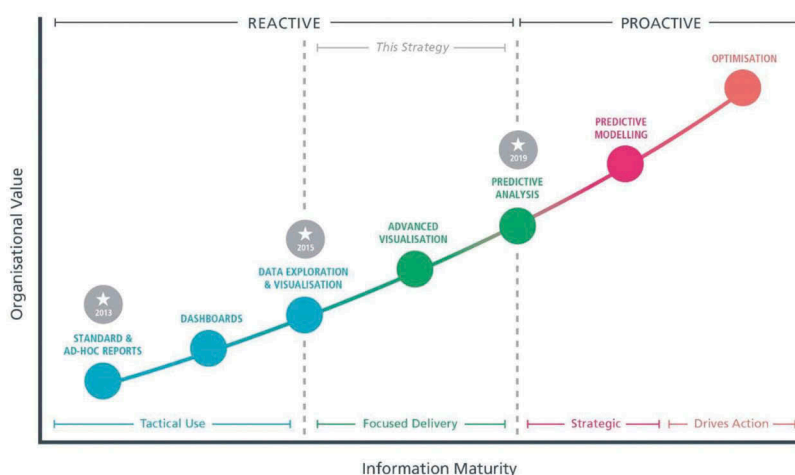
Figure 1. The centre's BI strategy – information maturity.

> Effectively, in our tool set we have access to machine learning, and we have access to create ... the correct term is clusters, and everything that goes with those. ... so basically we can do whatever we want, and then pull that data out of the data warehouse, and other sources push it back into the data warehouse. So effectively ... this is machine learning ... and we can pull data out of here into there, and push it back in again. (BI manager)

A follow-up visit in 2017 provided us with an update on implementation of the strategy and opportunities to explore early uses of data science and predictive analytics in aspects of The Centre's work. At this time, we asked one of the two data scientists in The Centre whether the machine learning capabilities described during the earlier visit had been implemented and he replied: 'Not really. Not at the moment, no. I can see in the future there could be more of that, but not just now'. This different view likely reflects the task differentiation within The Centre's work. The data scientist was not training algorithms on data sets in ways that would fit the definition of machine learning, while the BI manager was referring to machine learning capacities embedded in Azure.

The BI manager and data scientist identified two projects in which The Centre's BI strategy was most developed. We focus on these projects in the following sections to show that The Centre is undergoing a transition towards implementing AI-driven data analytics and is experimenting with data science approaches that involve abductive reasoning. We also show how the syncopation of human thinking and nonconcious cognition is emerging in The Centre's work across multiple methods and media, producing a new 'coevolutionary cognitive infrastructure' (Parisi 2017, 7) for education policymaking.

Project 1: predicting future demand for schools

The first project aimed to predict demographic change among the school-aged population in order to plan the future location of schools. These predictions will inform strategic decisions that are high-stakes, because the opening and closure of schools is often a zero-

sum game and was very controversial in this education jurisdiction in the early 2000s (Gulson 2007). When we asked the BI Manager about this project he explained that it is:

> ... fully up and running and in production right now, so we have – two floors below us is a team of about twenty people who are planning where to put schools. ... and they're doing it using this tool, that [shows] where the students will enrol in the schools. So basically we've taken information from our data warehouse ... so they can say, "I'm going to put a school here, and I'm going to double that school over there in size, and I'll close that school down." Run a simulation. And it takes about a minute for a simulation to run, and it allocates every single little [block] of students into their school, and it says, "OK, well, at the end of it, in fifteen years' time, in these patches over here, there'll be no schools for these kids." (BI manager)

Demographic data collected by the Australian census, and data on school location and size, are fed into simulations (a form of predictive modelling) that can be used to anticipate the infrastructure needs of the system. The data scientist explained that this is

> a really interesting tool that people are building, and we've done a bit of debugging of ... you sort of feed in shapefiles of New South Wales and populations and predicted populations, and it generates this map and says, "You need to put a school here", so that kind of thing I find really interesting, because that's very practical and direct. (Data scientist)

This is a clear example of becoming 'information centric' as The Centre's work shifts into predictive analytics and modelling, with an emphasis on strategic decision-making. Aspects of the policy process in The Centre are becoming automated, at least to the extent that the 'need' to locate schools in particular locations emerges from the use of algorithms to run simulations.

Project 2: predicting future assessment outcomes of students

The second area where the BI strategy is particularly advanced is an exploratory study of assessment data from the Australian National Assessment Program – Literacy and Numeracy (NAPLAN). NAPLAN is administered annually to students in Years 3, 5, 7 and 9. In the educational jurisdiction where The Centre is located, a matriculation certificate can be obtained upon completion of Year 12. The Data Scientist explained that this project is in a pilot phase and involves:

> ... trying to predict NAPLAN [Year] 9 scores from NAPLAN [Year] 7. So, a different body within the education sector decided that they would not allow students to write their Year 12 exams if they did not achieve a certain level of proficiency in NAPLAN [Year] 9. ... But what we're trying to do is give schools early warning that, actually, Little Johnny, in Year 7, based on what he's doing now, is probably not going to hit that grade in Year 9, so you need to pass some extra work there to get him up to that level. (Data Scientist)

This project is premised on two assumptions: (1) that performance in Year 12 can be predicted from Year 9 NAPLAN performance and, in turn, Year 7 performance; and (2) pedagogical decisions in Year 7 can be usefully informed by these predictions. Of course, across the five years between Years 7 and 12 there are myriad factors, both within and beyond schools, that will have an impact on student test achievement.

The experimental nature of data science

In these projects The Centre aims to use data science to provide diagnostic and formative feedback that is different from the analyses that schools currently perform with their own data and methods. Data science is 'the study of the generalizable extraction of knowledge from data' with a focus on 'the computational aspects of pragmatically carrying out data analysis, including acquisition, management, and analysis of a wide variety of data' Liu and Huang 2017, 2). There is overlap between AI and data science in the areas of data modelling, which includes computational statistics, and high-performance algorithms. Data science focuses on management, analysis and visualization of data, and it involves a degree of exploration and indeterminacy.

During our second visit to The Centre, in 2017, it became clear that data science approaches were becoming increasingly central to their work. The data scientist explained the nature of his work as follows:

> And so they say, "Here's some data." So sometimes you have to go off and get it yourself from the data store; other times someone will supply a file, a CSV file that they've made. Then you sit down with it and you kind of explore it, so you understand how it's structured and what the contents are, and how you're going to aggregate it, and summarise it, and then you spend a bit of time thinking about the explorations you're going to do, or the statistical methods you're going to do.

The BI manager also conveyed the exploratory nature of this work when describing how data science was applied in an analysis of non-student factors that may affect student outcomes. This work is beginning to drive policy:

> … we started doing some … data science investigation work on the impact of [school] principals on student outcomes, without really knowing what we were going to find, or what to expect. So, OK, here's the data warehouse, here's all this information, here's a data scientist, have a great time … So, data science is kind of like walking around a room in the dark, bumping into furniture. Oh, look! There's some gold! And oh, look! There's the mud. You never know what you're going to find. So we have two data scientists doing that at the moment. They're about to start a piece of work on the usage of … ICT in schools, and again, we have terabytes of data and no idea where we're going to go or what we're going to find. … That kind of thing will drive policy changes. (BI Manager)

As the examples of the school location simulation and NAPLAN prediction projects show, these techniques are being applied to a range of areas, from planning to performance data. This strategy is common in data science and machine learning, where substantive domain knowledge, while still used, is not a prerequisite for analysis. Furthermore, as we can see in these accounts, The Centre is implementing analytical techniques in multiple domains that are premised on indeterminacy (e.g. 'no idea where we're going to go or what we're going to find'). This indeterminancy, we suggest, is feeding into decision making and policy changes.

These aspects of the emergent 'automation' of the policy process seem to challenge a popular view of evidence-based policy making: that it is premised on pursuing the most efficient and instrumentally rational options, 'set against courses of action technically arrived at through the utilisation of explanatory-causal accounts' (Adams 2016, 293). Education data science has been similarly identified as an approach 'whose practices reflect a particular data scientific style of thinking that views learning in scientific terms as

quantifiable, measurable, actionable and therefore optimizable' (Williamson 2017, 120). However, Parisi provides another way of thinking about the introduction of data science and BI in education that is premised on the abductive reasoning of algorithms. Parisi (2016a) argues that

> abductive reasoning does not involve a probability calculation of the best explanation to the best hypotheses because, differently from inductive thinking, abduction does not rely on already-established hypotheses and observed discrete facts. ... abductive reasoning is ignorance preserving, which means that it entails the emergence of a new order, that is, new hypotheses and beliefs, or conceptual changes. (p. 478)

Indeterminacy is inherent in the description of The Centre's data science work (e.g. 'walking around in the dark, bumping into furniture'). As Parisi observes, 'indeterminacy is ... intrinsic to the algorithmic generation of hypothesis and as such the technoscientific articulation of truths and facts can no longer be confined to recurring functions and executions of the already known' (p. 10). This view of data science has affinities with critiques of policy science and the identification of policy making and implementation as ad hoc, contingent and uncertain (Ball 1994; Webb and Gulson 2015).

Aside from the machine learning services in Microsoft Azure, algorithms that enable new kinds of automated thinking are not yet evident in The Centre's work. Nonetheless, the simulations and explorations in the two projects described above are beginning to affect the 'cognitive infrastructure' of policy making and educational decion making in this context. For example, the NAPLAN project can be seen as exemplifying an embryonic form of automated thinking, and thus constituting very early example of 'the emergence of levels and meta-levels of inference that have radically changed our methods of reasoning' (Parisi 2016a, 480). NAPLAN is often criticised for only providing a snapshot of the student performance and it may be an ineffective predictor of the future academic performance of individuals. From this perspective, the project could be critiqued on the grounds that the predictions may well be incorrect, but such critiques focus on the validity of the analysis rather than its exploratory aspect. The analyses conducted in The Centre also illustrate how creative inferences can be made from multiple data sources, as part of a broader '"operating platform" for human thinking' (Parisi 2016a, 480), potentially opening new possibilities for decision-making and understanding policy problems.

Becoming information centric: the uncertain horizon of optimisation

The data scientist explained that the current work on on ' predictive analytics ... is ultimately going towards optimisation'. Optimisation is the goal located at the top end of the maturity curve in Figure 1 and was described to us as being the point at which data analysis was future looking and had direct impact on the improvement of outcomes. The BI manager described it as an aspirational goal that essentially captured the desire to use data to maximise outputs relative to inputs. When asked to elaborate on the organisational understanding of optimisation, the BI manager stressed the obstacles to reaching this point:

> [W]e have this ongoing argument with State Treasury ... because they try to have a cost benefit model, and we're like, OK, fine. How are you going to do it? ... They've gotten pretty close to

saying, if you complete Years 11 and 12, what that means is that society will be better off because you can then do the following things, which means you'll earn more in your career, which means you can pay more tax, which means more tunnels, or whatever. Alright? But there's so many things that feed into it that they can't control. It's like trying to be an economist, they have this great thing called *ceteris paribus* that, holding everything else constant, if I move this lever, that lever moves. Yes, but what you're holding constant is this enormous mountain of stuff that you have no control over at all. So you just try and have the best teachers and staff you can, in the right schools. And identify where it's not working, and fix it. (BI manager)

The Centre is using larger volumes of data and increasing the speed and impact of analysis, while acknowledging that these analyses are embedded in contexts that connot be controlled and are thus inherently inderterminate. The BI manager highlights that the nature of education as an unruly enterprise in which the abstract 'optimisations' suggested by data science and associated techniques like machine learning, if implemented as policy, will interact in complex and unpredictable ways with the multiplicity of forces at work in any policy context.

What is determined to be optimal in simulations and predictive analyses, for example, will never be optimal in practice given the messiness of the contexts in which data-driven policy will be implemented. Moreover, the operational definition of optimal depends on values, like what is considered to be 'good', as the BI manager explained: 'I once asked someone who'd been here for a long time how you know if you're getting better at education and they said, 'Well, first you have to define what "good education" is,' and no-one seems able to do that, because it's different for everybody". Here the BI manager expresses doubt about promises 'that the future is resolvable through the optimized output of algorithmic decision engines' (Amoore 2019, 19), which potentially creates another point where the possibility of things unfolding differently is kept open. The emergent automation of thinking in The Centre is both epistemologically indeterminate, as argued by Parisi, but also preserves uncertainty through embodied doubtfulness regarding the possibility of optimization.

Discussion and conclusion

Our case study shows that automated thinking is only beginning to emerge in The Centre's work, and AI is yet to be deployed outside of capabilities embedded in its BI platform. Nonetheless, the indeterminancy at the heart of The Centre's emergent coevolutionary cognitive infrastructure suggests the need for new theoretical and methodological approaches to critical policy analysis in order to keep pace with these developments, and, more importantly, to articulate with them in ways that open new possibilities for critical thought in education studies. As Deleuze (1995) emphasised, understanding and engaging with the mechanisms of control societies requires the creation of new concepts and strategies, and thus new values.

The role of values and creative uncertainty in automated thinking – *a syncopation of machine cognition and human thought* – are particularly important to consider as part of the study of data infrastructures and data science approaches in education policy. Rizvi and Lingard remind us that

[i]n education, policy processes have to juggle a range of values . . . often simultaneously, against a calculation about the conditions of possibility. . . . This requires privileging some

values ahead of others. In the process, the meaning and significance assigned to each value is re-articulated (Rizvi and Lingard 2011, 9).

The automated thinking that emerges in The Centre will shape the juggling, privileging and rearticulation of values. The developments that are contributing to the emergence of automated thinking can be critiqued based on present values by, for example, suggesting that automated thinking moves away from a school- and classroom-based professional knowledge about student progress. Such critiques highlight the partiality and instrumentality of performance data. Additionally, just as we have seen with the use of AI in search engines such as Google, biases in existing data sets and algorithms can produce biased and perverse outcomes (Noble 2018). There is a critical literature that raises concerns about the negative effects of datafication and associated technological change on education governance, as well as teaching and learning in classrooms. We acknowledge the importance of these critiques, however, focusing upon the instrumentality of machine cognition can also occlude other potentialities embedded in these developments.

Automated thinking will create new conditions for contestation over how education policy should be made and for what purposes. Rather than only conceiving of it as an instrument of optimisation, through inductive analyses of large data sets, we can also conceive of this work as a speculative method for creating new values and conditions for thought. As Parisi (2017) argues,

> '[i]f the antagonism between automation and philosophy is predicated on the instrumental use of thinking, techno-philosophy should instead suggest not an opposition, but a parallel articulation of philosophies of machines contributing to the reinvention of worlds, truths, and facts that exist and can change'.

The events surrounding DeepMind's AlphaGo and AlphaZero algorithms provide an instructive example of such reinvention. Most Go players initially took the view that their expertise could only be imitated by algorithms, just as many people tend to assume that algorithms can only offer reductive approaches to analysis and decision-making in education. However, champion Go players made three very interesting observations after playing against AlphaGo: (a) the algorithm made totally unexpected moves, indicating that it could create new play styles; (b) the algorithm made 'beautiful' moves, indicating that it could invent new values for judging play; and (c) champion players learnt how to play the game more effectively by playing against AlphaGo. These algorithms have opened up new possibilities for playing Go, and algorithms used for data analysis in education may also open up new possibilities for policy and practice. We note here that education is a field where new possibilites and values are needed, as extensive research on the pernicious and inequitable consequences of education policy for different groups has shown (e.g. Dumas, Dixson, and Mayorga 2016).

We cannot predict what new policy values might emerge from the rise of automated thinking, just as it was not possible to anticipate AlphaGo's strange new moves. But, we can predict that we will learn how to make policy differently, in the sense that our values and expectations change, as a result of integrating AI into policy analysis. These changes are already occurring in our daily lives as Netflix and Amazon algorithms that shape decisions about what to read or watch, and Apple watches shape decisions about walking or sleeping. The syncopation of machine cognition and human thought is reconfiguring meaning and significance, which is a dangerous reconfiguration in the

sense that Foucault outlined when he explained that 'not that everything is bad, but . . . everything is dangerous, which is not exactly the same as bad. If everything is dangerous, then we always have something to do' (Foucault 1983, 231–232). The theoretical perspective that we have adopted points to the need for more *doing* in critical education policy studies. Our case study has only begun to discern some of the possibilities of automated thinking and we see a need for methodological developments and new research designs that enable more pragmatic and collaborative explorations of what new cognitive infrastructures can do. This experimentation must of course be accompanied by critical questions about the values that are being replaced or changed, the interests driving the development and use of new platforms and data science methodologies, and the risks of these approaches for teachers, students and societies.

The optimizations pursued by The Centre will always find their way into connection with the messy and contested worlds in which policy is implemented. All forms of analysis create arbitrary neatness from uncertainty and this is their purpose, as Gregory Bateson's famous line about maps and territories conveyed so clearly. Policy analysis and policy texts are no different; both are forms of abstract representation that collapse into ad hocery when implemented. In this paper, our theoretical frameork offered conceptual tools that can help us to see how new data analytics can both provide reductive abstractions *and* actively add to the creative uncertainty of policy making. The point is not simply that automated thinking involves uncertainty, but that this uncertainty may increasingly enable education policy makers to create new values and possibilities for acting upon the world. This creativity is not only due to developments in the algorithms that cognize, but also in the coevolutionary cognitive infrastructure in which these algorithms are embedded. As Parisi argues (2016b), '[t]oday it is crucial to develop a philosophy of computation accounting for the transformations of logic, and the emergence of a social artificial intelligence, in which the nonconscious function of automated algorithms is just one dimension'. In this paper, we have taken a first tentative step in this direction.

Notes

1. *Cloud computing* enables users to access computer power and data storage over the internet. *Business intelligence* describes a range of techniques for gathering and analyzing data in order to inform business decisions and improve outcomes. *Predictive analytics* and *modelling* involves the analysis of data to predict future events. *Data* are anything that is given to our senses, while *information* describes the form that data take when captured by a particular system (e.g. bits of information processed by a computer). However, these terms are often used interchangeably in our discussion and the distinction between 'data centricity' and 'information centricity' depends on the embedding of data analytics in everyday business decision making and action, rather than a difference between data and information.
2. Microsoft Azure is a cloud computing platform. Azure enables users to run applications in using Virtual Machines and store data in Microsoft's servers. Azure provides an alternative to setting up and running an on-site data centre. Cloud computing can reduce costs while providing greater performance and reliability through economies of scale.
3. The syncopation of nonconscious cognition and thinking can be understood in terms of subtle shifts between the major points of conscious thought that interrupt expectations. For

example, algorithms that analyse our previous behavior to present us with targeted choices subtly change the context of our thought.

4. The key year for the establishing machine learning as the preeminent contemporary AI approach was 2006, when Geoff Hinton and colleagues published 'A fast learning algorithm for deep belief nets' in *Neural Computation, 18*(7): 1527–1554.

5. AlphaGo is a computer programme that plays the game of Go. It was developed by Google's DeepMind. The programme is based on a neural network that uses deep learning algorithms. Since 2015, AlphaGo has beaten a series of professional Go players in high profile matches and has surprised the Go community with strange new moves. AlphaZero is a more powerful and generalized variant of AlphaGo and is the world's top Go player.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Notes on contributors

Sam Sellar is Reader in Education Studies at Manchester Metropolitan University. Sam's research focuses on education policy, educational theory, large-scale assessments and datafication. His current projects explore the introduction of AI into education policy making and the role of digital data in new modes of educational governance. Sam co-lead's the Education and Global Futures research group in the Education and Social Research Institute at MMU. Sam's most recent book is the *World Yearbook of Education 2019: Comparative Methodology in an Era of Big Data and Global Networks* (2019, Routledge), co-edited with Radhika Gorur and Gita Steiner-Khamsi

Kalervo N. Gulson is a Professor in Education at the University of Sydney, Australia. Kalervo's research is located across social, political and cultural geography, education policy studies, and science and technology studies. His current research programme is focused on education governance and policy futures and the life and computing sciences. This research investigates whether new knowledge, methods and technologies from life and computing sciences, with a specific focus on Artificial Intelligence, will substantively alter education policy and governance. Kalervo's most recent book is *Education policy and racial biopolitics in multicultural cities* (2017, Policy Press) with P. Taylor Webb.

ORCID

Sam Sellar 🆔 http://orcid.org/0000-0002-2840-5021
Kalervo N. Gulson 🆔 http://orcid.org/0000-0002-3581-9935

References

Adams, P. 2016. "Education Policy: Explaining, Framing and Forming." *Journal of Education Policy* 31 (3): 290–307. doi:10.1080/02680939.2015.1084387.

Amoore, L. 2019. "Doubt and the Algorithm: On the Partial Accounts of Machine Learning." *Theory, Culture & Society*. doi:10.1177/0263276419851846.

Anagnostopoulos, D., S. A. Rutledge, and R. Jacobsen, Eds. 2013. *The Infrastructure of Accountability: Data Use and the Transformation of American Education*. Cambridge, MA: Harvard Education Press.

Bakker, R. S. 2015. "Crash Space." *Midwest Studies in Philosophy* 39 (1): 186–204. doi:10.1111/misp.2015.39.issue-1.

Ball, S. J. 1994. *Education Reform: A Critical and Post-structural Approach*. Philadelphia: Open University Press.

Boden, M. A. 2016. *AI: Its Nature and Future*. Oxford: Oxford University Press.

Bratton, B. H. 2014. "The Black Stack." *e-flux journal* 53: 1-12. https://www.e-flux.com/journal/53/59883/the-black-stack/.

Bratton, B. H. 2015. *The Stack: On Software and Sovereignty*. Cambridge, MA & London: MIT Press.

Campalo, A., M. Sanfilippo, M. Whittiker, and K. Crawford. 2017. *AI Now 2017 Report*. New York: AI Now.

Chaitin, G. 2006. ""The Limits Of Reason"." *Scientific American* 294 (3): 74–81.

Davenport, T. H., and R. Ronanki. 2018. "Artificial Intelligence for the Real World." *Harvard Business Review*. January–February. https://hbr.org/2018/01/artificial-intelligence-for-the-real-world

Deleuze, G. 1995. "Postscript on Control Societies." In *Negotiations*, edited by G. Deleuze, 1972–1990. New York: Columbia University Press.

Deleuze, G., and F. Guattari. 1994. *What Is Philosophy?* New York: Columbia University Press.

Dumas, M. J., A. D. Dixson, and E. Mayorga. 2016. "Educational Policy and the Cultural Politics of Race: Introduction to the Special Issue." *Educational Policy* 30 (1): 3–12. doi:10.1177/0895904815616488.

Easterling, K. 2014. *Extrastatecraft: The Power of Infrastructure Space*. London: Verso Books.

Easterling, K. 2014. *Extrastatecraft: The Power Of Infrastructure Space*. London & New York: Verso.

Feher, M. 2009. "Self-appreciation; Or, the Aspirations of Human Capital." *Public Culture* 21 (1): 21–41. doi:10.1215/08992363-2008-019.

Fischer, M. M. 2009. "Forward: Renewable Ethnography." In *Fieldwork Is Not What It Used to Be: Learning Anthropology's Method in a Time of Transition*, edited by J. D. Faubion and G. E. Marcus, vii–xiv. Ithaca and London: Cornell University Press.

Flyvbjerg, B. 2006. "Five Misunderstandings about Case-study Research." *Qualitative Inquiry* 12 (2): 219–245. doi:10.1177/1077800405284363.

Foucault, M. 1983. "On the Genealogy of Ethics: An Overview of a Work in Progress." In *Michel Foucault: Beyond Structuralism and Hermeneutics*, edited by H. Dreyfus and P. Rabinow, 229–252. 2nd ed. Chicago: University of Chicago Press.

Gilbert, N., P. Ahrweiler, P. Barbrook-Johnson, K. Narasimhan, and H. Wilkinson. 2018. "Computational Modelling of Public Policy: Reflections on Practice." *Journal of Artificial Societies and Social Simulation* 21 (1): 1–14. doi:10.18564/jasss.3669.

Gillborn, D., P. Warmington, and S. Demack. 2018. "QuantCrit: Education, Policy, 'big Data' and Principles for a Critical Race Theory of Statistics." *Race Ethnicity and Education* 21 (2): 158–179. doi:10.1080/13613324.2017.1377417.

Goffey, A. 2008. "Algorithm." In *Software Studies: A Lexicon*, edited by M. Fuller, 15–20. Cambridge, MA & London: MIT Press.

Gulson, K. N. 2007. "Repositioning Schooling in Inner Sydney: Urban Renewal, an Education Market and The 'Absent Presence' Of The 'Middle Classes'." *Urban Studies* 44 (7): 1377-1391.

Hao, K. 2019. "We analyzed 16,625 papers to figure out where AI is headed next". *MIT Technology Review*, January 25. https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/

Hartong, S., and A. Förschler. 2019. "Opening the Black Box of Data-based School Monitoring: Data Infrastructures, Flows and Practices in State Education Agencies." *Big Data & Society* 6 (1): 1–12. doi:10.1177/2053951719853311.

Hayles, N. K. 2014. "Cognition Everywhere: The Rise of the Cognitive Nonconscious and the Costs of Consciousness." *New Literary History* 45 (2): 199–220. doi:10.1353/nlh.2014.0011.

Kitchin, R., and M. Dodge. 2011. *Code/space: Software and Everyday Life*. Cambridge, MA & London: MIT Press.

Komljenovic, J. 2019. "Big Data and New Social Relations in Higher Education: Academia.edu, Google Scholar and ResearchGate." In *World Yearbook of Education 2019: Comparative Methodology in the Era of Big Data and Global Networks*, edited by R. Gorur, S. Sellar, and G. Steiner-Khamsi, 148–164. London & New York: Routledge.

Liu, M.-C., and Y.-M. Huang. 2017. "The Use of Data Science for Education: The Case of Social-Emotional Learning." *Smart Learning Environments* 4 (1): 1. doi:10.1186/s40561-016-0040-4.

Lyotard, J. F. 1984. *The Postmodern Condition: A Report on Knowledge*. Minneapolis: University of Minnesota Press.

Noble, S. U. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Parisi, L. 2016a. "Automated Thinking and the Limits of Reason." *Cultural Studies ↔ Critical Methodologies* 16 (5): 471–481. doi:10.1177/1532708616655765.

Parisi, L. 2016b. "Automation and Critique." Reinventing horizons symposium, Prague, 18-19 March. http://www.reinventinghorizons.org/?p=355#more-355

Parisi, L. 2017. "Reprogramming Decisionism." *e-flux journal* 85: 1–12. https://www.e-flux.com/journal/85/155472/reprogramming-decisionism/.

Pasquale, F. 2019. "Professional Judgment in an Era of Artificial Intelligence and Machine Learning." *Boundary 2: an International Journal of Literature and Culture* 46 (1): 73–101. doi:10.1215/01903659-7271351.

Ratner, H., and C. Gad. 2018. "Data Warehousing Organization: Infrastructural Experimentation with Educational Governance." *Organization* 26 (4): 537–552. doi:10.1177/1350508418808233.

Rizvi, F., and B. Lingard. 2011. "Social Equity and the Assemblage of Values in Australian Higher Education." *Cambridge Journal of Education* 41 (1): 5–22. doi:10.1080/0305764X.2010.549459.

Robertson, S. 2019. "Comparing Platforms and the New Value Economy in the Academy." In *World Yearbook of Education 2019: Comparative Methodology in the Era of Big Data and Global Networks*, edited by R. Gorur, S. Sellar, and G. Steiner-Khamsi, 169–186. London & New York: Routledge.

Sellar, S. 2015. "A Strange Craving to Be Motivated: Schizoanalysis, Human Capital and Education." *Deleuze Studies* 9 (3): 424-436.

Star, S. L. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–391. doi:10.1177/00027649921955326.

Thrift, N. 2014. "The 'sentient' City and What It May Portend." *Big Data & Society* 1: 205395171453224. doi:10.1177/2053951714532241.

Williamson, B. 2016. "Coding the Biodigital Child: The Biopolitics and Pedagogic Strategies of Educational Data Science." *Pedagogy, Culture & Society* 24 (3): 401–416. doi:10.1080/14681366.2016.1175499.

Williamson, B. 2017. "Who Owns Educational Theory? Big Data, Algorithms and the Expert Power of Education Data Science." *E-Learning and Digital Media* 2042753017731238. doi:10.1177/2042753017731238.

Williamson, B. 2018. "Silicon Startup Schools: Technocracy, Algorithmic Imaginaries and Venture Philanthropy in Corporate Education Reform." *Critical Studies in Education* 59 (2): 218–236. doi:10.1080/17508487.2016.1186710.

Yin, R. 1994. *Case Study Research: Design and Methods*. 2nd ed. Thousand Oaks: Sage.

Yin, R. 2009. Case Study Research: Design and Methods. 4th ed. Thousand Oaks: Sage