



---

# New digital laboratories of experimental knowledge production: Artificial intelligence and education research

Ben Williamson 

## How to cite this article

Williamson, B. (2020) 'New digital laboratories of experimental knowledge production: Artificial intelligence and education research'. *London Review of Education*, 18 (2): 209–220. <https://doi.org/10.14324/LRE.18.2.05>

Submission date: 17 October 2019

Acceptance date: 16 March 2020

Publication date: 21 July 2020

## Peer review

This article has been peer-reviewed through the journal's standard double-blind peer review, where both the reviewers and authors are anonymized during review.

## Copyright

© 2020 Williamson. This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

## Open access

*London Review of Education* is a peer-reviewed Open Access journal.

---

# New digital laboratories of experimental knowledge production: Artificial intelligence and education research

Ben Williamson\* – *University of Edinburgh, UK*

## Abstract

Education data scientists, learning engineers and precision education specialists are new experts in knowledge production in educational research. By bringing together data science methodologies and advanced artificial intelligence (AI) systems with disciplinary expertise from the psychological, biological and brain sciences, they are building a new field of AI-based learning science. This article presents an examination of how education research is being remade as an experimental data-intensive science. AI is combining with learning science in new 'digital laboratories' where ownership over data, and power and authority over educational knowledge production, are being redistributed to research assemblages of computational machines and scientific expertise.

**Keywords:** artificial intelligence (AI), algorithms, data, experts, infrastructure

Since the 2010s, the expansion of data science has been accompanied by significant claims about the changing status of knowledge and expertise. Datafied knowledge production (Bonde Thylstrup *et al.*, 2019) has been made possible by technical advances in computational statistics, data analytics and machine learning algorithms – so-called artificial intelligence (AI) – along with new kinds of expert data practices in computer science, informatics, data science and software engineering (Mackenzie, 2017). These technologies and practices have become extremely valuable in a wide range of knowledge-producing fields, as data science has expanded from technical development to the media, business, finance, entertainment, government, the public sector and academia (Beer, 2019).

In educational research, the rise of learning analytics, education data mining, artificial intelligence in education (AIED) and education data science, along with a booming 'edtech' industry has brought advanced computational technologies into forms of analysis and knowledge production, empowering new actors who can claim expertise to generate novel understandings of educational processes through extensive and complex volumes of digital data (Perrotta and Selwyn, 2019). As in the data science discipline and data analytics industry more broadly, 'defining new objects of knowledge are very much entwined with the making of professions' (Ruppert, 2018: 17). The professionalization of education data science as an academic discipline is being established through research conferences, journal publications, special issues, institution building and increasing public exposure of evidence.

Education data science has also escaped the enclosures of academic research. Education data scientists inhabit education businesses, edtech companies, think tanks, consultancies, transnational governance organizations, independent research laboratories and government education departments, raising questions about the ownership of data and its capacities to be used for knowledge production and theory generation beyond the academy (Williamson, 2017). Even where education data science is practised in higher education institutions, its disciplinary home may be outside of education departments, located in the research infrastructures of the human sciences, psychology, cognitive science, neuroscience and genetics. The increasing uptake of data-centric experimental methods from the somatic, neural and psy-sciences in educational analyses is leading to a surge of interest in the science of learning, supported by foundations, government departments, think tanks, consultancies and transnational governance organizations as a fresh source for evidence-based education policy and practice (Kuhl *et al.*, 2019).

These shifts are changing the possibilities for educational research as a field, introducing data-intensive experimental methods and generating new forms of knowledge and theory that challenge existing frameworks in the disciplines of education. In particular, as AI has become normalized in education research and merged with approaches from learning science, new conceptualizations such as precision education (Brookman-Byrne, 2018) and learning engineering (Wagner and Lis, 2018) have emerged. Precision education specialists and learning engineers are new hybrid experts able to straddle fields of advanced computing, data and learning sciences. As in other data-intensive human sciences, the digital experts of learning science produce novel psychological, neural and somatic knowledge from big data using advanced computational and AI techniques. These arrangements of AI technologies, scientific expertise and data practices make it possible to see and know students in novel ways as digital traces of psychodata, brain data and biodata (Williamson, 2019). They are also reconfiguring and reworking power relations and epistemic expertise in education research itself. Although the datafied learning sciences, precision education and learning engineering are not all entirely symmetrical, in this article they are treated as a family of experimental approaches that combine AI and learning science, abbreviated as AILSci.

The purpose of this article is to open up the actors, technologies, experimental practices and power of AILSci research to analysis. Focusing on three key cases – the rise of learning engineering, molecular genetic analysis and neuroimaging – the analysis shows how education research is being remade as an experimental data-intensive science, where AI combines with learning sciences in new digital laboratories that exist inside computer machinery. As Bonde Thylstrup *et al.* (2019: 3) note, ‘datafied knowledge production’ is ‘quietly transforming how we see, read, organise, use and dispose of knowledge’. These changes demand analyses that pay close attention to: the specific contexts and actors involved in data generation; the resources, methods and procedures it includes; the digital devices and infrastructures that support its development; and the potentialities and limits of data-based forms of analysis and knowledge production (*ibid.*). While previous research has investigated the potential of AI to participate in teaching and educational management (Selwyn, 2019) or shape policymaking (Sellar and Gulson, 2019), this article examines how AILSci knowledge is being produced by specific research machinery, its claims to epistemic expertise and authority, its entanglement of human and artificial intelligence and the implications of data-intensive knowledge production for educational research.

## Research machines

Research across the sciences, humanities and social sciences is increasingly conducted through digital knowledge machines that are reconfiguring the ways knowledge is generated, circulated and used (Meyer and Schroeder, 2015). These knowledge machines are also subtly changing how human, social and natural worlds are known and understood, as objects, people and events have become traceable as digital data (Marres, 2017). Beyond individual devices, the generation of data depends on knowledge infrastructures, or networks of people, artefacts and institutions that generate, share and maintain specific knowledge (Edwards *et al.*, 2013). Knowledge infrastructures, such as those of statistical institutes or research-intensive universities, have undergone significant digital transformation with the arrival of data-intensive technologies, with knowledge production now enacted in myriad settings, from academic laboratories and research institutes to commercial research and development studios, think tanks and consultancies. Datafied knowledge infrastructures have become hubs of command and control over the creation, analysis and exchange of data (Bigo *et al.*, 2019).

How, then, should the analysis of these knowledge machines and infrastructures be approached? Fox and Alldred (2018: 194) refer to research assemblages as the heterogeneous elements that are bound together in the enactment of research, which include: the 'events' to be researched, such as bodies, things, settings or social formations; research tools or other apparatus; recording and analysis technologies; computer software and hardware; theoretical frameworks and hypotheses; research literatures and findings; the 'data' generated by these methods and techniques; and researchers and the physical spaces and establishments where the research takes place. Research assemblages are interconnected 'machines' that 'do specified tasks', such as data collection and data analysis, enable particular research capacities, and achieve specific methodological objectives. Additionally, any method or technique within a research assemblage does micropolitical work by opening up or closing down possibilities for action and knowledge production. This is a useful way of approaching the micropolitical work of computational knowledge machines in research processes. It recognizes how the events (bodies, things, settings) subjected to research are transformed into data or knowledge according to the methodological choices, theories and disciplinary styles of thinking embedded in the assemblage. The event in the computer program and the database is an artefact of the specific configuration of instruments, measures and methods mobilized to record, store and process it, which changes the very things the research machines and infrastructure were designed to capture.

As an emergent research assemblage, the research machines and knowledge infrastructure of AILSci are both data-intensive – rigorously concentrated on analysis of data – and life-sensitive – focused on detection, or sensing, of signals and traces from students' activities and bodies, and their translation into knowledge through the somatic, neural and psychological sciences. The AILSci research assemblage consists of: learning events and bodies to be researched; specific research machines and infrastructures for tracing those events and bodies; particular disciplinary theories and knowledge; data generated through AI techniques and methods; and the researchers undertaking research and producing knowledge in specific AILSci expert settings, spaces and institutions. By unpacking the AILSci research assemblage, it becomes possible to identify the micropolitical work of knowledge production in the new educational data sciences. The following three sections focus on some key examples to illustrate the key implications of the shift to AI-based knowledge production in

education: (1) a shift in expertise to the authority of data science exemplified by the elevation of the figure of the learning engineer; (2) the role of technical infrastructure in configuring educational knowledge, as shown by the rise of bioinformatics in new molecular genetics research in education; and (3) a shift in analytical focus from statistics to sensing, as illustrated by the uptake of neuroimaging as a source of educational knowledge production.

## Expert power

Scientific aspirations to measure education precisely through assessment of individuals' capacities have a long history. In the early twentieth century, the psychologist Alfred Binet argued that 'methods of scientific precision must be introduced into all educational work' (Binet quoted in Comfort, 2019: 168). Binet's quantitative measures of children's mental age relative to chronological age gave rise to the statistical discipline and industry of intelligence quotient (IQ) testing. IQ was progressively taken up in eugenicist attempts to measure children's innate intelligence, as 'IQ was converted from a measure of a given child's past performance to a predictor of any child's future performance' (Comfort, 2019: 168). A century later, the combination of AI with learning science has animated new objectives to measure and predict student performance with scientific precision, using very large digital data sets and powerful predictive analytics capacity. As the Organisation for Economic Co-operation and Development (OECD) director of education puts it, educational data science brings 'the rigours of scientific research to education policy' (Schleicher, 2018: 17), allowing 'digital exhaust' to be transformed into 'digital fuel', with 'data as a catalyst for educational practice' (ibid.: 262).

The combination of AI and learning science into an AILSci research assemblage consists of particular forms of scientific expertise embodied by knowledge actors – individuals and organizations – identified by categories including science of learning, AIED, precision education and learning engineering. These forms of scientific expertise are gaining authority in educational matters. The OECD, for example, has released a report on the science of learning, defining it as a combination of cognitive science, neuroscience and biomedicine with computer science, data analytics, machine learning, AI and engineering (Kuhl et al., 2019). The report claims 'significant insights have been achieved into the complex, dynamic processes and mechanisms that underlie how people learn', and advocates 'the use of Big Data, Artificial Intelligence algorithms, education data mining and learning analytics ... to improve learning and education' through 'science-based actions' in 'real-world education practice and policy' (ibid.: 13–14). As an influential transnational governance organization, the OECD is seeking to embed datafied AILSci approaches in educational policy and practice, empowering experts who are able to straddle fields of computer sciences and human sciences in the analysis and explanation of learning processes.

Precision education overtly uses psychological, neurological and genomic data to tailor or personalize learning around the unique needs of the individual (Williamson, 2019). Precision education approaches include cognitive tracking, behavioural monitoring, brain imaging and DNA analysis. Similarly, the figure of the learning engineer has emerged as an ideal type of AILSci practitioner. Learning engineering was first coined by computer scientist and AI pioneer Herb Simon at Carnegie Mellon University in the 1960s, resurfacing around 2013 in technology-focused US universities (Lieberman, 2018), especially Carnegie Mellon's LearnLab (<https://learnlab.org/>) and MIT's Learning Engineering Project Blueprint (<https://playful.mit.edu/projects/>

schools-of-tomorrow/). The learning engineer is not only a university expert, but is also a role promoted by education and technology consultancies, think tanks, foundations and venture capital funding firms. The Chan Zuckerberg Initiative (CZI), the for-profit philanthropy of Facebook founder Mark Zuckerberg and of Priscilla Chan, invests in learning science to shape school practices and policy through psychological, neuroscientific and biomedical evidence and expertise (<https://chanzuckerberg.com/education/learning-science/>). CZI's director of learning science actively advocates the figure of the learning engineer, who can work across 'computer science, linguistics, education, measurement science, cognitive science, motivational and social psychology, machine learning, cognitive neuroscience, among others', in order to make sense of evidence 'at multiple levels, from clickstreams, motion position data, speech streams, gaze data, biometric and brain sensing, to more abstracted feature sets from all this evidence' (Saxberg, 2018: viii).

Learning engineering is also the subject of significant inter-organizational effort to define professional standards and competencies. Led by the Institute of Electrical and Electronics Engineers (IEEE), the world's largest and most influential organization for standards-setting in computer technologies and engineering, the IEEE Industry Consortium on Learning Engineering (ICICLE) is a 'professional organization committed to the development of Learning Engineering as a profession and as an academic discipline': 'Learning Engineering is a process and practice that applies the learning sciences using human-centered engineering design methodologies and data-informed decision making to support learners and their development' ([www.ieeeicicle.org/](http://www.ieeeicicle.org/)). Its membership consists of over 60 organizations representing industry, academia and government, including Google, HP and IBM. ICICLE has established special interest groups to undertake projects in AI and adaptive learning technologies, learning analytics, competency frameworks, learning data standards, learning data governance and learning experience design, as well as teams exploring routes to growth of learning engineering both in industry and academia (Wagner and Lis, 2018). Its effort to construct the learning engineer as a new figure of expertise with defined professional competences cutting across computational and learning sciences will enable AILSci to scale up as an authoritative educational data science discipline.

The new experts of AILSci exemplify how data science expertise has proliferated across sectors to influence how social worlds, populations and individuals are seen, known and governed (Beer, 2019). Davies (2017) argues that data analysts represent an emergent form of elite power under advanced neoliberalism. In the context of declining public trust in institutions and official statistics, which are increasingly perceived by the public to be marked by producer bias, he claims, big data have attained privileged status as objective, authoritative sources of knowledge. Expert power is therefore claimed by those who can perform big data analyses, especially those able to translate and narrate the data for various audiences. Likewise, expert power in education is now claimed by those who can enact data-intensive science of learning, precision education and learning engineering research and development, and translate AILSci findings into knowledge for application in policy and practitioner settings. Educational data scientists claim ownership over data and over knowledge production, and in doing so confer on themselves new professional and interdisciplinary authority to intervene in educational debates. By invoking both AI and the life sciences as key elements in the research assemblage of educational data science, they also seek the historical and epistemic legitimacy of the statistical and natural sciences to know social worlds and living bodies intimately and objectively. These new professionals and experts of learning represent how a new hierarchy of educational research is emerging, with the



support of influential organizations such as the OECD, IEEE and CZI, whereby the experts who can mobilize data-scientific research machinery are elevated as dominant sources of epistemic authority in the disciplines of education.

## Thinking infrastructures

The elite experts of educational data science are also embedded in a knowledge infrastructure that consists of people, technologies, standards of practice and institutions, as well as values, norms, habits, epistemologies and desires. Knowledge infrastructures are:

robust internetworks of people, artifacts, and institutions which generate, share, and maintain specific knowledge about the human and natural worlds. Like all infrastructures, they are composed of many systems and networks, each with its own unique dynamics ... [and] embed social norms, relationships, and ways of thinking, acting, and working. As a corollary, when they change, authority, influence, and power are redistributed. (Edwards *et al.*, 2013: 23)

Advances in technology are now implicated in transformations to the knowledge infrastructures through which knowledge was historically produced. As knowledge infrastructures have become increasingly artificially intelligent through algorithms and machine learning, they have become thinking infrastructures of distributed human-machine cognition that can structure attention and orchestrate decision making, influence thought and action, and enable new forms of governance, power and control (Bowker *et al.*, 2019). As Bowker *et al.* (*ibid.*: 1) clarify, the thinking of a thinking infrastructure is not merely a conscious human cognitive process, but relationally performed across humans and socio-material strata, wherein interconnected technical devices and other forms 'organize thinking and thought and direct action'. Thinking infrastructures 'trace their objects, making them visible and available as objects of and for possible interventions', and they 'enable interventions into, and reorganization of, the governing of individuals, communities, organizations and entire markets' through 'establishing a distinct conception of the objects and objectives of government' (*ibid.*: 4–5). A thinking infrastructure, then, is a mode of human–non-human cognition that produces knowledge from digital data and thereby makes it possible to intervene in and govern the social world.

The thinking infrastructures of AILSci are concatenations of experts, technologies, practices, institutions and so on, internetworked to produce new kinds of datafied knowledge, in ways that are redistributing authority, influence and power in educational research. The AILSci infrastructure includes experts and practices as detailed above, but also specific computational research machines. As an infrastructure for AILSci analyses, these technologies at least partly structure how experts think: they generate new understandings and knowledge about processes of education and learning that are only thinkable and knowable due to the computational machinery of the research enterprise.

The analysis of genetic data through bioinformatics technologies exemplifies the emerging thinking infrastructures of AILSci. In recent years, huge volumes of molecular genetic data have become available for educational research along with data-intensive bioinformatic research infrastructure for studying it, opening up a fast-moving field of educational genomics that is concerned with producing new knowledge about the biological substrates of learning, intelligence, cognition and other educational

outcomes (Martschenko *et al.*, 2019). Building from this work, the behavioural geneticist Plomin (2018) specifically advocates a model of precision education based on tailoring education to students' DNA. A specific form of data-intensive biological analysis underpins this vision of precision education. Genome-wide association studies (GWAS) involve analysis of huge data sets of small genetic variations, captured on genotyping chips, that can be calculated together to produce polygenic scores, or statistical predictions of a certain phenotypical trait based on genotypical data. For Plomin (*ibid.*: 181), these 'polygenic scores are key for personalized learning, as they predict pupils' profiles of strengths and weaknesses, which offers the possibility to intervene early to prevent problems and promote promise'.

Big data-based molecular genetics studies are part of a bioinformatics-led transformation of biomedical sciences based on analysing exceptional volumes of data (Parry and Greenhough, 2018), which has transformed the biological sciences to focus on structured and computable data rather than embodied evidence itself. The infrastructural capacity for GWAS is provided by multinational biotechnology companies and the invention of DNA microarrays, genotyping chips and automated laboratory robots, which allow digital biodata to be data-mined and processed by machine learning algorithms in order to identify small but significant genotypical associations across large populations (Kragh-Furbo *et al.*, 2016). This transformation in the nature of evidence importantly has implications for how nature is evidenced. As Stevens (2017: 153) notes, data-inside-computers is created 'so that it can take on distinct forms and be manipulated in ways that are not possible without a computer'. Bioinformaticians are involved in different knowledge-making practices, and 'the quantities of data processed by computers, and the algorithms needed to deal with them, make a qualitatively different kind of knowledge' (*ibid.*: 172). A significant consequence is that biologists have learned to think of human genetics in cybernetic terms, as codes to be decoded, texts to be searched and as networks containing complex correlational patterns. The logic of big data biology powered by bioinformatics carries the same kinds of working knowledge as found in commercial web searching, indexing, sorting and detecting and matching patterns. The ways bioinformatics systems are constructed is therefore consequential to knowledge production, as vast biobanks, genotyping chips, automated laboratory robots and data analytics algorithms form into hybrid thinking infrastructures where the possibilities of scientific thought are shaped by the knowledge machines.

Educational data science projects based on bioinformatic molecular genetics are embedded in thinking infrastructures of human and machine cognition that produce knowledge. Hayles (2017) argues digital media are now crucial components of cognitive assemblages of distributed human and machine cognition. In cognitive assemblages, human and technical cognizers interact with each other, especially as technical systems are endowed with ever greater autonomy and agency to make decisions and choices. Drawing on actor-network theory, Hayles (*ibid.*: 116) argues that technical systems as well as human actors can be understood as important mediators that 'transform temporary and shifting configurations into durable, robust, and reproducible structures capable of creating, solidifying, and wielding power'. As a result, Hayles (*ibid.*: 116–17) views 'cognitive assemblages, and the mediators operating within them, as the means by which power is created, extended, modified, and exercised in technologically developed societies' where 'complex human systems are interpenetrated by technical cognition'.

From this perspective, the thinking infrastructures of ALLSci are cognitive assemblages of both human expertise and technical cognition, in which algorithms



play significant roles alongside human scientists as powerful mediators that can shape knowledge production. Bioinformatics technologies such as genotyping chips, laboratory robots and polygenic scoring algorithms are key mediators in precision education efforts, with the capacity to scan the human genome for patterns, generate associations between genotypes and phenotypes such as educational attainment and achievement, and contribute to new understandings of the genetic determinants of educational, social and economic outcomes. This kind of thinking infrastructure is transforming how education is known, understood and potentially acted upon, as technical and human cognizers combine into cognitive assemblages where each affects the knowledge making of the other, and especially as technical cognizers become mediators capable of shaping the practices and outcomes of knowledge production. New cognitive infrastructures consisting of AI techniques enable new forms of 'automated thinking' that shape human decision making, and that create new possibilities for educational policymaking and school practices (Sellar and Gulson, 2019). However, these processes of AI-infused 'automated knowledge discovery' risk producing highly contested understandings of 'learning' (Perrotta and Selwyn, 2019), such as bioinformatically produced insights into the molecular genetic substrates of intelligence and school achievement, which risk reintroducing biological discrimination into education under the guise of 'precision education' (Martschenko *et al.*, 2019). In the thinking infrastructure of molecular genomic analysis of education, bioinformatics technologies now play a key role in structuring how the biological bases of educational outcomes are known and understood. While attention to the work of human experts, learning engineers, education data scientists or precision education specialists, remains an important focus for further research on AI-based learning sciences, analysis should also focus on the technical cognizers with which they are cognitively entangled in thinking infrastructures, and on the specific role of those technologies in mediating and shaping educational knowledge production.

## Sensory power

Across the various programs of ALLSci, new ways of conceiving of learning and education are being produced. These new conceptualizations also generate novel proposals for how to intervene through policy, pedagogy or curriculum design to improve learning processes and outcomes, thereby creating new power dynamics over human subjects. Isin and Ruppert (2019) have recently conceptualized an emergent form of power that they characterize as sensory power. Building on Foucault, they note how sovereign power gradually metamorphosed into disciplinary power and biopolitical forms of statistical regulation over bodies and populations. The latter, biopower, significantly depends on practices of statistical stocktaking and on the expertise of the psychosciences, demography and official statistics in making human subjects knowable as populations with quantified characteristics that might be governed. Sensory power, they suggest, marks a subtle shift in biopower brought about by digital technologies and big data, as individuals have become traceable and knowable at high levels of granularity and in or near to real time. Sensory power marks a shift to practices of data-intensive sensing, and to the quantified tracking, recording and representing of living pulses, movements and sentiments through devices such as wearable fitness monitors, online natural-language processing and behaviour-tracking apps. Davies (2019: 515–20) designates these as 'techno-somatic real-time sensing' technologies that capture the 'rhythms' and 'metronomic vitality' of human bodies, and bring about 'new cyborg-type assemblages of bodies, codes, screens and machines' in a 'constant cybernetic loop of action, feedback and adaptation'.

AILSci is illustrative of how techno-somatic sensory technologies that can capture the metronomic vitality of students are being mobilized as modalities of power. GWAS research based on bioinformatics, for example, senses out polygenic patterns from deep in DNA and, through cross-analysing these data against social and economic data sets, calculates predictions of socio-economic outcomes. That is to say, it senses certain aspects of socio-economic progress from embodied life processes, leading to claims that precision education practices could then be developed as DNA-based personalized education interventions. Likewise, science of learning and learning engineering advocates pursue the potential of cognition tracking and affect tracing. The techno-somatic sensory possibilities for AILSci psychologists manifest most spectacularly in proposals to utilize wearable biometric monitors and facial emotion detection systems to monitor students' engagement and attention in real-time during learning tasks (Williamson, 2019).

Techno-somatic modes of neural sensing, using neurotechnologies for brain imaging and neural analysis, are the next frontier in AILSci. Real-time brainwave sensing is being developed and trialled in multiple expert settings. The Precision Learning Center at the University of California, San Francisco, deploys electroencephalogram (EEG) recording of student cognition and affects through partnerships with BrainLENS and Neuroscape brain-imaging labs ([www.precisionlearningcenter.org/](http://www.precisionlearningcenter.org/)), while the Brainwave Learning Center is a partnership between Stanford University and the Synapse School in Silicon Valley with its own onsite Brainwave Recording Studio ([www.synapseschool.org/about-us/blc](http://www.synapseschool.org/about-us/blc)). It uses a suite of wearable EEG headnets and wraparound neuroheadsets to perform real-time tracing of students' brainwave activity during learning exercises. The Harvard-incubated neurotechnology start-up company BrainCo, meanwhile, has developed the FocusEDU wearable neuro-sensing headband, and has partnered with a personalized learning edtech company to create a neuro-optimized education platform, whereby sensory data from the brain is processed by machine learning to determine how students individually receive learning micro-routines that are tailored to aspects of brain activity ([www.brainco.tech/](http://www.brainco.tech/)).

These cases exemplify how AILSci efforts are oriented to sensing somatic rhythms and the metronomic vitality of students' living bodies. Made possible by scientific advances at the intersection of AI and machine learning with the human sciences, they seek to make embodied processes and patterns visible as explanatory evidence of the cognitive, neural and affective substrates of academic learning and educational outcomes. In doing so, AILSci turns the student's body inside out, by body-mining for internally embodied traces and rhythms, and then extracting and revealing them for inclusion in computational data sets and analysis by advanced algorithms. In other words, AILSci is dedicated to identifying internally embodied substrates of learning located in DNA, bodies and brainwaves. Sensory power is exercised when these data are then used as the basis for intervening in the somatic substrates of learning, as knowledge produced about the student's body becomes the basis for improving, enhancing and governing the embodied and embrained processes that underlie individual educational development and longer-term social and economic outcomes.

## Conclusion

Education research is increasingly done in digital laboratories of precision engineering and experimentation. The digital laboratories of AILSci are located inside computer machinery as much as in physical settings, and they are inhabited by new kinds of professionalized experts and specialists in digital data analytics, neural analysis, digital

psychometrics and bioinformatics. The focus of experimentation and knowledge discovery in these digital laboratories is on digital trace data that have been sensed, mined and extracted from students' bodies, and are treated as techno-somatic samples of the embodied substrates of cognition, affects and learning. Specific research machines, including EEG sensors, platforms, biometrics and bioinformatic technologies, are key elements in the emerging infrastructure of AILSci, with the human cognition of AILSci experts merging with the cognitive capacity of research machines to produce new cyborg assemblages of knowledge production. Of course, unquestioning trust in these techniques and technologies to produce transformative understandings needs to be tempered with recognition of their partiality, epistemological limitations and ethical consequences. The aim of this article is to begin illuminating the ways that new forms of expertise, research machines, cognitive infrastructures and epistemologies of the learning body are now being configured as a new educational research assemblage. The production of this assemblage of knowledge production will itself demand further empirical scrutiny to more fully articulate its practical, epistemological and ethical implications. Two initial implications can be identified from this examination of the emergence of AILSci as a new mode of educational research.

First, the transformation of the knowledge infrastructures of education research potentially entails:

significant distributional consequences, advancing the interests of some and actively damaging the prospects of others in the production of knowledge ... As knowledge infrastructures shape, generate and distribute knowledge, they do so differentially, often in ways that encode and reinforce existing interests and relations of power. (Edwards *et al.*, 2013: 13–14)

Expert authority is increasingly concentrated among AILSci specialists, and these relations of power over other forms of disciplinary and methodological expertise are likely to grow as transnational policy influencers such as the OECD and the CZI continue to push science of learning knowledge as insights for policymaking and practice. The ownership over data claimed by AILSci also confers it power over knowledge production, displacing other forms of culturally situated and socially oriented enquiry. The assemblage of research machines that constitute AILSci is infused with methodological micropolitics, opening up certain forms of knowledge production and action while closing down others. It empowers the interests of neurotechnologists, bioinformaticians and digital psychologists to define student learning and educational outcomes in ways derived from the computational analysis of trace data (Williamson, 2019).

Second, however, this raises the issue of how the bodies and events that are the focus of AILSci are in fact artefacts of the research machinery designed to reveal and represent them:

Beyond which faction of professions within a field will exercise that authority at issue is the power 'to reveal things'. Choices about methods of producing data ... produce and reproduce the very objects that they ostensibly reflect. In this sense they are performative in that they do not involve the discovery of truths about objects but simultaneously represent and enact, that is, bring into being the very objects they are meant to describe and represent. (Ruppert, 2018: 19)

AILSci is in this sense deeply performative, conjuring new scientific truth about the embodied substrates of learning into being through the careful sociotechnical

calibration of its research machines and infrastructure. The brainwaves, bodies and polygenic scores that the epistemic authority of AILSci concentrates upon are in fact methodological accomplishments of assemblages of expert actors, sensing technologies, measurement instruments, epistemologies, theories, practices, choices and assumptions. AI learning science exemplifies the growing influence and power of datafied knowledge production to shape how social worlds, behaviours and even human bodies are known, understood, governed and enacted. The potential influence of AI in education extends beyond monitoring of students' activities on learning platforms, reaching into the somatic, neural and psychological substrates of learning.

Educational data sciences are becoming authoritative in educational research and policy debates, yet their knowledge-production accomplishments are highly shaped by specific configurations of research machines, epistemic expertise and thinking infrastructures that are owned by a privileged elite of research institutions and produce highly contested understandings of biological learning bodies. The new profession of the education data scientist, exemplified by AILSci specialists such as learning engineers, is creating new data objects as ways of conceptualizing learning and educational outcomes. Rather than merely revealing the hidden strata of learning in the body, education data science is bringing into being controversial conceptualizations of learning through the experimental apparatus of its new digital laboratories.

## Notes on the contributor

**Ben Williamson** is Chancellor's Fellow in the Centre for Research in Digital Education at the University of Edinburgh. He is an editor of *Learning, Media and Technology* and the author of *Big Data in Education: The digital future of learning, policy and practice* (SAGE, 2017).

## References

- Beer, D. (2019) *The Data Gaze: Capitalism, power and perception*. London: SAGE Publications. <https://dx.doi.org/10.4135/9781526463210>.
- Bigo, D., Isin, E. and Ruppert, E. (2019) 'Data politics'. In Bigo, D., Isin, E. and Ruppert, E. (eds) *Data Politics: Worlds, subjects, rights*. London: Routledge, 1–17. <https://doi.org/10.4324/9781315167305>.
- Bonde Thylstrup, N., Flyverbom, M. and Helles, R. (2019) 'Datafied knowledge production: Introduction to the special theme'. *Big Data and Society*, July–December, 1–5. <https://doi.org/10.1177/2053951719875985>.
- Bowker, G.C., Elyachar, J., Kornberger, M., Mennicken, A., Miller, P., Randa Nucho, J. and Pollock, N. (2019) 'Introduction to thinking infrastructures'. In Kornberger, M., Bowker, G.C., Elyachar, J., Mennicken, A., Miller, P., Randa Nucho, J. and Pollock, N. (eds) *Thinking Infrastructures*. Bingley: Emerald Publishing, 1–13.
- Brookman-Byrne, A. (2018) 'Precision education: What could the future of teaching and learning look like?'. *BoLD*, 26 March. Online. <https://bold.expert/precision-education/> (accessed 2 April 2020).
- Comfort, N. (2019) 'How science has shifted our sense of identity'. *Nature*, 574 (7777), 167–70. <https://doi.org/10.1038/d41586-019-03014-4>.
- Davies, W. (2017) 'Elite power under advanced neoliberalism'. *Theory, Culture and Society*, 34 (5–6), 227–50. <https://doi.org/10.1177%2F0263276417715072>.
- Davies, W. (2019) 'The political economy of pulse: Techno-somatic rhythm and real-time data'. *Ephemera*, 19 (3), 513–36. Online. <https://tinyurl.com/ybpxouq9> (accessed 22 April 2020).
- Edwards, P.N., Jackson, S.J., Chalmers, M.K., Bowker, G.C., Borgman, C.L., Ribes, D., Burton, M. and Calvert, S. (2013) *Knowledge Infrastructures: Intellectual frameworks and research challenges*. Ann Arbor, MI: Deep Blue. Online. <https://tinyurl.com/yaet72fg> (accessed 22 April 2020).

- Fox, N.J. and Alldred, P. (2018) 'Mixed methods, materialism and the micropolitics of the research-assemblage'. *International Journal of Social Research Methodology*, 21 (2), 191–204. <https://doi.org/10.1080/13645579.2017.1350015>.
- Hayles, N.K. (2017) *Unthought: The power of the cognitive nonconscious*. Chicago: University of Chicago Press.
- Isin, E. and Ruppert, E. (2019) 'Data's empire: Postcolonial data politics'. In Bigo, D., Isin, E. and Ruppert, E. (eds) *Data Politics: Worlds, subjects, rights*. London: Routledge, 207–27.
- Kragh-Furbo, M., Mackenzie, A., Mort, M. and Roberts, C. (2016) 'Do biosensors biomedicalize? Sites of negotiation in DNA-based biosensing data practices'. In Nafus, D. (ed.) *Quantified: Biosensing technologies in everyday life*. Cambridge, MA: MIT Press, 5–26.
- Kuhl, P.K., Lim, S.-S., Guerriero, S. and Van Damme, D. (2019) *Developing Minds in the Digital Age: Towards a science of learning for 21st century education*. Paris: Organisation for Economic Co-operation and Development.
- Lieberman, M. (2018) 'Learning engineers inch toward the spotlight'. *Inside Higher Ed*, 26 September. Online. <https://tinyurl.com/r6dqlzf> (accessed 2 April 2020).
- Mackenzie, A. (2017) *Machine Learners: Archaeology of a data practice*. Cambridge, MA: MIT Press.
- Marres, N. (2017) *Digital Sociology: The reinvention of social research*. Cambridge: Polity Press.
- Martschenko, D., Trejo, S. and Domingue, B.W. (2019) 'Genetics and education: Recent developments in the context of an ugly history and an uncertain future'. *AERA Open*, 5 (1), 1–15. <https://doi.org/10.1177%2F2332858418810516>.
- Meyer, E.T. and Schroeder, R. (2015) *Knowledge Machines: Digital transformations of the sciences and humanities*. Cambridge, MA: MIT Press.
- Parry, B. and Greenhough, B. (2018) *Bioinformation*. Cambridge: Polity Press.
- Perrotta, C. and Selwyn, N. (2019) 'Deep learning goes to school: Toward a relational understanding of AI in education'. *Learning, Media and Technology*, 1–19. <https://doi.org/10.1080/17439884.2020.1686017>.
- Plomin, R. (2018) *Blueprint: How DNA makes us who we are*. London: Allen Lane.
- Ruppert, E. (2018) *Sociotechnical Imaginaries of Different Data Futures: An experiment in citizen data*. Rotterdam: Erasmus University Rotterdam.
- Saxberg, B. (2018) 'Preface'. In Niemi, D., Pea, R.D., Saxberg, B. and Clark, R.E. (eds) *Learning Analytics in Education*. Charlotte, NC: Information Age Publishing, vii–x.
- Schleicher, A. (2018) *World Class: How to build a 21st-century school system*. Paris: Organisation for Economic Co-operation and Development.
- Sellar, S. and Gulson, K.N. (2019) 'Becoming information centric: The emergence of new cognitive infrastructures in education policy'. *Journal of Education Policy*, 1–18. <https://doi.org/10.1080/02680939.2019.1678766>.
- Selwyn, N. (2019) *Should Robots Replace Teachers? AI and the future of education*. Cambridge: Polity Press.
- Stevens, H. (2017) 'A feeling for the algorithm: Working knowledge and big data in biology'. *Osiris*, 32, 151–74. <https://doi.org/10.1086/693516>.
- Wagner, E. and Lis, J. (2018) 'The rise of learning engineering'. *eLearn Magazine*, August. Online. <https://elearnmag.acm.org/featured.cfm?aid=3242101> (accessed 3 April 2020).
- Williamson, B. (2017) 'Who owns educational theory? Big data, algorithms and the expert power of education data science'. *E-Learning and Digital Media*, 14 (3), 105–22. <https://doi.org/10.1177%2F2042753017731238>.
- Williamson, B. (2019) 'Digital policy sociology: Software and science in data-intensive precision education'. *Critical Studies in Education*, 1–17. <https://doi.org/10.1080/17508487.2019.1691030>.