Journal of Educational Controversy

Volume 11
Number 1 *Is "Best Practices" Research in Education Insufficient or even Misdirected?*
AN INVITED ISSUE DEDICATED TO JOHN G. RICHARDSON

Article 6

2016

Big Data and Technologies of Self

Bernadette Baker
bbaker@education.wisc.edu

Big Data and Technologies of Self

Bernadette Baker

Abstract

The entry of Big Data into the educational field has generated noticeable binary reactions and a recycling of criticisms already directed at the quantification of reality, datafication in the social sciences, standardization in education, and neoliberalism in the West. This paper reapproaches Big Data's entry into education from a curriculum studies perspective, which deploys interdisciplinary approaches from philosophy, history, sociology. and politics of knowledge and wisdom. The analyses of key definitional debates, binary reactions, and systematization are considered from the point of view of historically shifting technologies of self as core conditions of possibility for the controversies that emerge when two fields intersect. Specifically, the alliance presumed between self and knowledge, and of both with reality, have long and provincial heritages that contemporary movements such as Big Data seem to reanimate and reconfigure. The paper concludes with consideration of whether Big Data can be understood as a game changer in the educational and curriculum fields and if so, on what basis.

Introduction

> The invention of the special classes, reducing a diverse range of physical and mental differences to a few number of categories, was a societal movement to be sure. But more accurately, it was a movement that projected the idea of the nation as a bounded geopolitical entity. As census counts presumed to delineate the size and composition of real and distinguishable groups, the visibility of the nation relied increasingly on this presumption... Of equal importance to census categories was the parallel movement that reduced diverse types of knowledge into a delimited number of educational categories... The point where these counts intersected was the curriculum of the common school… (Richardson, 2009, p. 142).

At a time when both the "death of globalization" and the "age of hyper-globalization" are heralded with equal confidence, assembling an accurate view

> of the future is less a matter of binary choices – a rosy versus a gloomy scenario – than of constructing a mélange of several visions. Today we don't get to choose between a world of great power competition, globalized interdependence, and powerful private networks; we have all three at the same time (Khanna, 2016, xviii).

Current efforts to mimic educational endeavors deemed successful and to reform institutions and practices deemed troublesome converge around some core recognitions: education is complex, it has multiple valences, it is not easy to pin down as an "it;" to define, bound, package, and replicate. Education as an interdisciplinary field, as variegated institutions, as a diverse profession, as nationally driven, as globally connected, as place-based or virtual, as incidental or planned, as formal and postformal, as short term forays into subject matter, as lifelong learning or as unlearning is as elastic as the contexts couching its mention, the vocabularies marshaled to it, the purposes assigned to it, and the stakeholders invested in it. In disciplinary and institutional terms alone, education defies mapping. It is somehow reproductive of social castes, sorting processes and distributive logics, yet undermines tight predictive validity; somehow open to recombinatorial potentials for creativity, yet closed off to genuine innovation; somehow a passive recipient of the next reinvention of the wheel, yet furtively exceeds the parameters of checklist thought and indoor settings.

The substantive multiplicity around what education is, what it is for, who it is for, and whether it can continue in its current institutional formats is keyed to broader issues that pertain particularly to the historic invention of compulsory schooling as part of the claim to and formation of "the West."[1] The three "great bastions" undergirding the legitimation of Westernized educational provision and compulsory attendance law –

religions, the nation-state, and the concept of the individual – are now highly contested as sources of authority from a variety of perspectives. Contestation has emerged in academic and more popular accounts, from Indigenous and non-Indigenous critique, and from different sides of the party political spectrum.[2] Richardson (2009) and Khanna (2016) recognized, for instance, how European-derived essentializing narratives of nation-as-social-compact were made up in different ways around education, how claims to technology and science have been integral to new spatializations, and that both redraw maps that underscore the importance of the mélange of visions. In doing so, both recognized the glosses and risks inherent in different movements, the reductive tendencies, and the potential polarization. For Richardson, in high schooling, in what is presently referred to as the United States, the knowledge reduction became posited as the turn-of-the-20th century "scientific" versus "classical" curriculum debate, which saw different regions vying to represent "the new frontier" of a new Republic. In Khanna's *Connecteography,* the reduction is in regard to a "new frontier" of the 21st century, the "rosy" versus "gloomy" picture of a future predicated on technologically-driven interconnection, for which colorful political maps of national borders are now insufficient.

There is a subtle and inherent echo in the two arguments that is important to unpack. It includes and exceeds the dilemma of religions-as-belief-systems, nations-as-social-compacts and individuals-as-containers, and it includes and exceeds the recognition of reductionism and polarization that occur through representation. The echo and its reverberations pertain in this case to what happens when "systems" rewrite "being" and being enfolds within systems differently than before. This paper is a thought experiment that explores this echo specifically around what happens to and with Westernized conceptions of being-*as-self*, taking seriously the limits and

possibilities of onto-epistemological issues raised in current debates over education's framing, purposes, format, and direction.[3] Here, the rewriting of education will be explored via contemporary flashpoints that sit at the intersection of the "curriculum" and the "technological" and in the wake of the "great bastions' destabilization, reconfiguration, and/or demise.

There are several contemporaneous movements occurring in domains beyond education that have enormous potential to impact how education is defined, debated, and enacted. Movements such as artificial intelligence, Big Data, and neuroscience are not just coincidentally contemporaneous. They simultaneously trouble definitions and boundaries of the human; of capacity and ability; of mind/body conceptions; and of knowledge, truth and reality. For this special edition, I focus on the conditions of possibility and controversies related to Big Data's entry into education only, partly because it embodies the greatest scalar reach at the most immediate levels and partly because it reinvigorates entrenched polarized reactions related to prior movements.

Big Data sits perhaps ironically at the intersection between what Richardson has already illustrated and what Khanna is forecasting, between curriculum-as-reductive-representation and interconnection-as-dual-edged-flow. Between Richardson's understanding of how curriculum becomes curriculum and Khanna's new version of network mapping lies, then, a tangible locus and notable pivot, a movement where technology meets teleology and impacts trajectory – in this case what it could mean to "know" or perform and the "systems" in which such relations participate, simultaneously forging new narratives about selfhood (Baker, 2001).[4]

For instance, in her recent exposé titled *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, self-defined quant, Harvard University mathematics Ph.D., former hedge fund statistician, data scientist, and math-

lover Cathy O'Neill (2016) describes the problems that became clearer to her through observing the Global Financial Crisis from within:

> The crash made it all too clear that mathematics, once my refuge, was not only deeply entangled in the world's problems but also fueling many of them. The housing crisis, the collapse of major financial institutions, the rise of unemployment – all had been aided and abetted by mathematicians wielding magic formulas. What's more, thanks to the extraordinary powers that I loved so much, math was able to combine with technology to multiply the chaos and misfortune, adding efficiency and scale to systems that I now recognize as flawed (p. 2).

One of the flaws that O'Neill (2016) describes in education-related deployments of Big Data WMDs (weapons of math destruction) pertains to hiring and firing decisions of teachers based on algorithms to which teachers have no access:

> An algorithm processes a slew of statistics and comes up with a probability that a certain person *might* be a bad hire, a risky borrower, a terrorist, or a miserable teacher. That probability is distilled into a score which can turn someone's life upside down. And yet when the person fights back, "suggestive" countervailing evidence simply won't cut it. The case must be ironclad. The human victims of WMDs…are held to a far higher standard of evidence than the algorithms themselves (p. 10).

At stake (among many other things), and that which O'Neill's analysis is sensitive to, is the reclassification, intensification, and distribution of what is coded as the self (including shifting understandings of human abilities and capacities) and the allied conceptualization of knowledge and knowledge-production (including the historically mutating line presumed between truth and falsity, ironclad evidence and suggestive

opinion). Significantly, the alliances presumed between self and knowledge, and of both with reality, have long and provincial heritages that contemporary movements such as Big Data seem to reanimate and reconfigure.

Specific definitional debates and responses will be taken up below. In mainstream terms, Big Data typically puts massive databases in touch with each other across platforms in order to seek and/or generate patterns between data bits that no single human or no single computer could have trawled through before. As a movement and an analytics, it has a relatively trendy presence, a heavier reliance upon numbers as a coding and measurement system than on that which is posited as qualitative, concern for automation and patterning, a focus on learning, and a rapidly diffusing and relatively closed-system authority embedded in the requirements of coding. What it also has and implicitly shares with contemporaneous discourses like artificial intelligence and neuroscience is the tendency to link particular ascetic techniques to the task of preparation amid the positioning of life as a life of testing (e.g., have objectives been met?). This latter, less-inspected quality, overtly intersects with the field of education and brings Big Data into (thus-far awkward) conversations with subfields such as teacher education, early childhood, and educational policy. The aim here, however, is not to determine the "worth" of Big Data's presence but something more modest and tentative.

Rather than an exhaustive review of literature in Big Data research or adjudicating between polarities in debates that seem to recycle criticisms already directed at the quantification of reality, datafication in the social sciences, standardization in education, and neoliberalism in the West, the layers here emanate from a curriculum studies perspective, which typically has deployed interdisciplinary approaches to philosophy, history, sociology, and politics of knowledge and wisdom in a

limited range of transcultural contexts (Baker, 2010; Eppert & Wang, 2008; Watkins, 1999). My aim, then, is to collectively examine whether and/or how Big Data's self-knowledge-reality relations appear as elaborations of prior tendencies in Western "technologies of self," and to consider where the landscape seems to be altered so profoundly that the playing field seems irreducible to what has gone before (Foucault, 1985, 2005). In doing so, the concern is not to arrive at a new normative picture but to sketch a different kind of impossible "map," one that helps to articulate the conditions of possibility for the controversies that emerge when two fields intersect and when "mapping" might be seen as part of the problem.

Big Data and Binary Reactions

The entry of Big Data into the educational field has generated noticeable binary reactions. The debates, a sample of which are outlined below, beg obvious questions: what *is* data? When does something achieve the status of data? What is not data? And why has it come to matter? Moreover, if there is a "big" typically there must be a "little" or "small," so what is small data, what makes something Big and capitalized, and what does running the two terms together reference about changed environments for research and possibilities for thought? What synergies or new modalities are being indicated? And why now?

Data and Big Data

In English languages, the term data today appears as synonymous with multiple bedfellows such as *evidence*, *empirical*, and *fact*. Of these four terms, however, data is the last to enter English vocabularies. *Datum* entered in 1630, followed by the plural *data* in 1645, drawn from the Latin *dare*, for that which is given, present.[5] In the six

different senses of its deployment since then, three are particularly pertinent for contemporary debates in education, including data as 1) an item of information typically collected for reference, analysis, or computation; 2) something given or granted, known or assumed as fact, and made the basis of reasoning; an assumption or premise from which inferences are drawn (including in philosophy, referring to anything immediately apprehended by or presented to the mind or senses), and 3) a line, point, etc., forming a basis for measurement; a baseline, benchmark, or reference point (as in surveying) (OED online). Inhabiting these inscriptions are what the social sciences today enwrap as commonsensically related: information, senses, mind, reasoning, baseline etc. Yet, there is nothing commonsensical in data's coming-into-being or self-explanatory in its rise in orientations to truth-production.

Indicators of this are the questions posed via some indigenous cosmologies, which contest how narratives about one group of people are staged by another. In *Disciplining the Savage, Savaging the Discipline* Nakata (2007) asked, for example, if it is really possible to step outside history and leave one's baggage at the door, even when trying to be sensitive to data-gathering from across different cultural contexts. Nakata underscored how much is left unsaid and undertheorized in educational appeals to the "reasoning" based on "data." He noted that even in approaches such as grounded theory, valid methodological questions need to be raised about "the new referents and criteria" in interpretive frameworks for re-presenting data. In questioning the well-intended efforts of contemporary researchers to portray the complexity of Torres Strait Islanders to white researchers Nakata (2007) reflected on the lack of theoretical rigor in making appeal to the term data:

> Data gathered from Islanders and other people or from the archives as emerging out of a particular history? What is that history? Is it one where everything is

> continuous to Western systems of thought or one that is continuous with Islander systems of thought? Or is it a history of the contesting positions where one loses and others win? Or is it one of strategically moving between them, borrowing from and adapting to ensure the continuance of one alongside the other? In short, what provided the context for the data to have meaning?... How, for instance, will he [the researcher] see recollections by individual Islanders? Do we focus on data from the archive and the individual as belonging to a history that tells of an uninterrupted chain of events, or do we try to focus on the systems of thought in the messages that they convey (i.e., the context in which things are said and that which provides its meanings)? And do we seek, in either of these positions, continuities or links with other factors/events/readings to re-present to readers what was stated or written, or do we seek discontinuities, ruptures, thresholds and constraints to allow readers to see what conditions the possibilities in what is stated or written? (p. 166).

Nakata's final provocation here: what conditions the possibilities in what is written, stated, or counted, is rarely the question posed in historical accounts of Big Data's emergence, and this difference points to a disconnect that reverberates with multifarious effects in much that follows. Investigations at the heart of the new slogan, Big Data, do acknowledge that there is no single origin for its coining. The investigations have arisen in journalism, commercial realms, governmental realms, and academic disciplines. For example, in reflecting on his research related to a *New York Times* article about Big Data's emergence, Lohr explained some of the difficulties, which are more than etymological:

> The term Big Data is so generic that the hunt for its origin was not just an effort to find an early reference to those two words being used together.

> Instead, the goal was the early use of the term that suggests its present connotation — that is, not just a lot of data, but different types of data handled in new ways (Lohr, 2013, p. 84).

In commercial realms, Big Data has had purchase deeper and for longer than in most other domains. The IT company Gartner, for instance, captured this shift of not just lots of data but how data-as-information-assets is generated, gathered, treated, and processed by defining Big Data succinctly as "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" (Gartner IT Glossary, 2016).

Government bureaucracies' deployment of Big Data more commonly, however, follow Lohr's observation: it was not about the coining or visible presence of the words in any one federal or bureaucratic system, but about how "data" was being generated, collected, and handled in new ways, thus providing a springboard for Big Data's coming-into-being. For example, the availability of structured and unstructured pieces of information amid the integration of sources from multiple institutions has redefined lots of "pieces of information" now as "data" in government-based services. In the process, citizens are sometimes newly described as partners rather than recipients, and the public service in multiple geopolitical regions is construed as more participatory. This points to what Morabito (2015) called a "prosumer era" (proactive consumer in direct online democracies) in which the nature of the relationship between government offices and those who do not work in them is fundamentally changed by new sources of information and their flows. The key idea in these kinds of governmental turns to Big Data is transversal responsibility where private, non-profit, university, and government offices partner

with individual citizens or community-based organizations for greater effectiveness in civic problem-solving and maintenance with the promise of lower costs. New sources of information include social media, Crowdsourcing, the Internet of Things, consultation of public talent, and private-pubic partnerships, which point to "Big Society" politics undergirded by what Big Data facilitates. Here, citizens-as-the-public armed with new technologies are positioned as taking responsibility for themselves and others in activities ranging from building smart cities, to census taking, to traffic light maintenance, crime fighting, and providing emergency services. Thus, although the term Big Data seemed to "enter" rather than become coined in such realms, governmental approaches gave Big Data something upon which to hang its hat with dualistic effects. On the one hand, the specter of Big Brother arose in surveillance technologies in the collecting and storing of information to which the citizens concerned do not have access (recentralization). On the other, mobile technologies especially facilitated new imaginings of decentralization, of redistributing responsibility, and of new modalities for interaction and service-delivery. The governmental versions covered purposes and processes that ranged, then, from increasingly active, immediate online engagement and input to the removed portrayal of citizens and utility usage as a series of correlates.

In academic disciplines, Big Data has not been seen as necessarily entirely new but as growing exponentially, as coming from exterior interests, and as more indexical of some fields than others:

> Big Data, that is, data that are too big for standard database software to process, or the more future-proof, "capacity to search, aggregate, and cross-reference large data sets"… is everywhere…In some ways, this is not a new phenomenon. Those working in the commercial sector have been

> collecting and combing large data sets to improve segmentation of goods to customers and better understand their market for many years… Nor is it particularly big news to those working in certain fields in the natural sciences. Yet, in recent years, a far wider range of stakeholders have become more involved and more excited about the potential of Big Data (Eynon, 2013, p. 237).

In the discipline of economics, which hosts both the more corporate-oriented and more statistics-oriented subfields, the naming of Big Data emerged via innovative approaches to macro-econometric dynamic factor models (DFMs). Reflecting on his own coining of the term, and then realizing he may not necessarily have been the first or only to do so, Diebold (2012) explained the discipline-specific distinctions that generated the need for a new vocabulary:

> I stumbled on the term Big Data innocently enough, via discussion of two papers that took a new approach to macro-econometric dynamic factor models (DFMs), Reichlin (2003) and Watson (2003), presented back-to-back in an invited session of the 2000 World Congress of the Econometric Society. Older dynamic factor analyses included just a few variables, because parsimony was essential for tractability of numerical likelihood optimization. The new work by Reichlin and Watson, in contrast, showed how DFMs could be estimated using principal components, thereby dispensing with numerical optimization and opening the field to analysis of much larger datasets while nevertheless retaining a likelihood-based approach. My discussion had two overarching goals. First, I wanted to contrast the old and new macro- econometric DFM environments. Second, I wanted to emphasize that the driver of the new macro-econometric DFM developments matched the driver of many other recent scientific

> developments: explosive growth in available data. To that end, I wanted a concise term that conjured a stark image. I came up with "Big Data," which seemed apt and resonant and intriguingly Orwellian (especially when capitalized), and which helped to promote both goals. But there really is nothing new under the sun, and credit for the term Big Data must be shared (pp. 2–3).[6]

While the concerns that frame issues unique to journalism, commercial projects, government bureaucracies, and academic disciplines and their subject matter specificity do matter and change what matters, a general consensus across venues has arisen that the late 1990s and early 2000s was the key timeframe of Big Data's literal coining, with its uniqueness tied to three major qualities: the three V's, or volume, variety, and velocity, of that which appears as given, as data. Moreover, in many settings the historical and institutional boundaries between the above venues had already been blurred by the circulation, spread, and translation of particular logics and profit priorities, alongside the advent of specific technologies developed for different ends. In a celebratory account, for instance, Datafloq noted that this included the 1937 development of punch card reading machines by IBM for the Roosevelt administration's record-keeping, data-processing machines for codebreaking in WWII (e.g., Colossus), the upsurge in hiring of cryptologists during the Cold War, and plans for the first data storage center for tax information and finger prints in the 1960s. These were accompanied by innovations we think of as internet-based, including the WorldWideWeb in the 1990s, the advent of Hadoop in the mid-2000s and the realization that by 2010, five exabytes of information were being created every two days (Datafloq, 2016). The import here is rather obvious: the "machine-based" innovations, platforms, and the teleological frames that "humans" created them within contributed to the conditions of possibility for Big Data to appear as Big.

For scholars engaging the term, though, the fusing and partnerships do not necessarily map onto the ability to have conversations over the same things. In his discussant response to a panel of presentations, Graham's (2016) "Historicizing Big Data and Geo-information" (posted on the "Connectivity, Inclusion and Inequality Group" of the Oxford Internet Institute at the University of Oxford) raised issues that exceed the definitional and terminological:

> For all of us – scholars in this field – I wonder if we're all speaking about the same thing in this session when we talk about 'big data'. Are we talking about datasets that are a census rather than a sample of a population? Are we just using 'big data' as a proxy for 'digital data'? Are we using that term to refer to the whole contemporary apparatus of data trails, shadows, storage, analysis, and use? Are we using it to refer to digital unknowns – the digital black box? Is the term actually helping us as short-hand for something else? Or do we need more precise language if we want to make sure we're genuinely having a conversation? (Graham, 2016)

Collectively and at a minimum, though, it must be recognized that the shift into Big Data-style thinking gave "applied mathematics" especially a greater range of phenomena to which to be applied. For example, the machines and programs combining and trawling datasets did not just exceed the ability of any one scholar to do on their own, but exceeded what that scholar could do even if they had 30,000 consecutive lifetimes of average length to try. The available changes in software and hardware (greater storage and memory, faster processing speeds, enhanced device mobility, cross-platform and cross-institution coding compatibility) can be located, then, as part of a broader 20th century continuity and discontinuity: as elaborating prior corporate and natural science preferences for prediction, efficiency, and control and as a jolt or

rupture of an order of magnitude so profoundly large that it constitutes a genuine environment change, one which humans arguably have had little time to process.

In education, however, which has historically retained the status of a social or human science, the ground for application of or conversation over a Big Data movement has not been so self-evident or tidy. This is in part because particularly mainstream discourses in education, such as developmentalism (stagiest, phasal and linear) and holism (environmental, contextual, and nonlinear), while oppositionally conceived, both still place the rearing of the young within direct human-to-human interactions such as mother-child and teacher-children, framed often by specific institutional contexts of families and/or schools. Such contexts have almost always been supplemented by "tools," like reading glasses, chalk, markers, books, globes, paints, and boards, but not dominated by machines like projectors or computers even when available. So while there has been a prior sense of "data-as-information" and "technology-as-tools" in formal education, their nature, relevance, and role has winnowed and swayed and is difficult to generalize across policy and classroom settings in terms of specific deployments, kinds of feedback, or participation in a bigger picture.

This is something that post-qualitative researchers especially in the field have referred to as "Possibilities of 'Data'" (the latter term deliberately suspended). In such orientations, there is a refusal to assert a priori what is and is not data and a simultaneous recognition that without "data" and conventions around "it" the risk of lack of rigor and "anything goes" becomes heightened (Koro-Ljungberg & Maclure, 2013). The fundamental refusal to assert which "kinds" of "data" are better for which purposes in advance and once and for all alludes to a broad, contemporary sentiment derived from perspectives on educational processes as messy, complex, and irreducible – what might be thought of as "small data," "local data," or not-so-big data. In leaving the

question of "data" *as* a question, one that opens possibilities for multiple storylines and narratives around "children's development," for example, such post-qualitative stances produce and work the certainty/uncertainty boundary differently than Big Data:

> We worry about these uncritical notions, definitions, enactments, and treatments of data and thus want to provoke discontinuation of data as we have come to know of it through postpositivism, empiricism, text books, research training, and other grand narratives. This special issue is dedicated to (un)knowing and (un)doing data. Our challenge to the authors was to problematize conceptualizations of data as known, familiar, and inert objects and to imagine more complex, creative, and critical engagements with data in the conduct of research. Following this problematization, the word *data* in this special issue should be read as "data," data (under erasure), data-undone, data-rethought, data-particles, or maybe data-becoming (Koro-Ljungberg & Maclure, 2013, p. 219).

Such an orientation, in refusing to concede on the one hand that "data," whether numbers or not, are automatically to be invoked as transparent capturers of truth or as telling the best or whole story, is acutely aware on the other that "data" is, can be, and must be seen as necessary, relevant, and purposeful (Koro-Ljungberg & Maclure, 2013).

While the above quote underscores a core feature within many social science logics, including educational ones - the tendency to continue making use of the tools and categories ("data") which it criticizes - this is not the same way that Big Data stylistics necessarily construct storylines emerging out of corporate, governmental, or natural science settings (Derrida, 1967/1978). Whereas for Koro-Ljunberg and Maclure (2013) the uncertain and unknowable is a moving target that is relative, that cannot be "solved" permanently, and that does not sit lying in wait for discovery, the epistemological

assumption underlying some versions of Big Data logic is that the three V's will indeed enable fixed and pre-existing truths, universal laws, and/or causality to be revealed. The major early points of Big Data's entry into education, then, were in those parts of the field that sat closer to such sensibilities such as educational psychology, educational technology, and governance in higher education. While a pie can be carved in many ways, especially when emergent, Big Data has largely manifested in two key areas that cut across such subfields as *educational data mining* (EDM) and *learning analytics* (LA). However there is disagreement as to whether these are so discrete in the first place, whether EDM has a specificity beyond other forms of data mining that its educational contexts give it, and whether LA are more about the child, students, and teachers or more about an organization, an administration, or a system (Baker & Inventado, 2014; Siemens & Baker, 2012).

Educational Data Mining

While it is a cliché in education to position any new method or technology as potentially dual-edged, as that which can be used for good or for ill, in regard to EDM, the spread, power, and exclusivity of access to data mining (and subsequent image-management) presents new issues. EDM is defined differently depending on the source and historicized somewhat differently as well. The aim here is not to offer a comprehensive repetition of historicizations, definitions, and debates and how the different camps and orientations formed across the late 20th and early 21st centuries, but to give a sense of the binary reactions to a few pertinent issues raised through, around, and in response to EDM. What makes them pertinent is that, at one level, responses for or against EDM might be raised in regard to many "quantification of

reality" movements in education, not just Big Data; and, at another related level, the issues point to more EDM-specific flashpoints and controversies.

If, at a minimum, it is accepted (and it may well not be) that EDM is a variation of data mining that has its own specificities peculiar to educational contexts, then the major difference it invites beyond regular statistical analysis, traditional quantification techniques and methods, and educational measurement and standardization discourses would be the size, speed, and range of what can be drawn into focus for analysis. Many ethical implications have been debated in other disciplines, but until recently in education the question of how data has been and is being used has been less investigated (boyd & Crawford, 2012; Romero & Ventura, 2010). Four key points of reflection and flashpoints have arisen as education engages with such questions, including privacy, applicability, interpretation, and structure/equity.

The lack of direct conversation about Big Data and children in early grades is somewhat surprising, especially because where children are involved in institutions that they are forced by law to attend, their vulnerability seems higher and the need for democratic dialogue more urgent. Recent literature demonstrates succinctly, though, how in settings beyond compulsory education, and in the tertiary sector unique sets of ethical issues are arising to prompt reflection on the considerations that are then unleashed (Kei Daniel, 2016). The low hanging fruit here seems to be identification of particular students as "at-risk" and what to do about it. In using Big Data for predictions of university student drop out from online courses, for example, new implications emerge amid all the positive effects such a study could have. Issues concerning privacy, informed consent and protection of harm, for instance, incite further ones:

> We need to think carefully about the social implications of this kind of use of the data. What happens to students who are most likely to drop out? Do we

> tell then, support them, (which has economic implications), or let them sign up and take their money…? What happens to serendipity in a system where all educational choices are based on recommender systems? What kinds of learning can a student truly keep "private"? Does the potentially highly public and trackable nature of learning have impacts for the learning process? There are any number of questions of this kind, which need to be asked and crucially considered every time data is analyzed and used… Big Data have a kind of kudos that needs to be treated with care, as the values that are designed in to the analysis process are not always properly considered or made explicit (Eynon, 2013, p. 238).

The normal/abnormal line assumed here (drop out as "abnormal" or less than desirable), while coming from a well-intended place of wanting to help, assumes a particular view of linearity and ability within systems. Once a norm is established, it urges certain students forward as sites of action or as problems, and the logic moves into how to remove or rehabilitate the deviation from the norm so as to make it disappear or become less disruptive.

The difficulty of presuming tidy linear causality in regard to human behavior and assuming correspondences between sites and systems is raised in other versions of privacy issues surrounding Big Data's role in MOOCs. Morabito (2015, p. 47) described how MOOCs can open new income streams for universities through employment recruiting services, syndication, and sponsoring, as well as by advertising income, selling student information to potential employers and advertisers. While Morabito noted the advantages, opportunities and challenges of the role Big Data plays in such settings, others have cautioned that many factors cannot be cross-institutionally copied, especially culturally specific orientations to a public/private line, institutional cultures,

and teaching strategies; making recommendations, policies, and suggestions difficult to generalize across universities, online programs, specific classes or even examinations (Prinsloo et al, 2015).

The wider questions of what kinds of data should be combined and analyzed and the purposes to which they should be put relates to a second set of challenges around applicability. Baker & Siemens (2011) noted initial sites of application, different prediction models, structure discovery algorithms, kinds of relationship mining and the spread of analytical practices from primarily business intelligence, psychology and educational measurement:

> Much of the early work in EDM was conducted within intelligent tutoring systems (as described in Koedinger and Corbett, 2006) and much of the work in LA began in web-based e-Learning and social learning environments. In recent years, this has extended to a wider variety of educational situations, including data from student collaboration around learning resources (Martinez et al., 2012), science simulations (Sao Pedro et al., 2013), teacher newsgroups (Xu & Recker, 2011), and school district grade data systems (Bowers, 2010). (p. 7)

The worth of such approaches within the wider field of education remains an open question. Prinsloo et al (2015) noted, for instance, that Big Data cannot be entrenched as automatically good or bad, nor does this make its presence neutral. They asked instead: "In order for big(ger) data to be better data, and to result in more effective and appropriate teaching, learning and support, what are the issues that we need to consider?" (p. 9). This does not resolve cross-cultural conflicts in how effective or appropriate teaching would be defined. Imagine that in some classrooms effective teaching may mean making sure that every student knows via the curriculum content

exactly how many American presidents were slave owners. In others, it might entail preventing students from having that knowledge. What constitutes the effective and the appropriate are not easily settled, and the "techniques" for getting students to remember on the one hand and forget or never know on the other cannot be presumed a priori to be neutral when such deep emotion is involved.

For Eynon (2013), the reality of clashing values means ensuring an understanding of the kinds of research that can and cannot be carried out using Big Data. Eynon suggested that rather than resolving or adjudicating such cultural differences Big Data seems to enter the picture on one side of them, tilting the field in certain directions: "The availability of Big Data limits the kinds of questions we can ask, as we can only study data we can collect or already have. We do not have data for everything thus the availability of Big Data shapes what people research and the question that they ask" (p. 238). Here, the issue also becomes one of expressing meaning and offering interpretation, which harkens back to the humanist/behaviorist split of 20th century psychology. If Big Data represents the threat of domination of education again by applied mathematics, this time without the Sputnik crisis of the 1950s to back it up, and given its funding and outer forces that support its spread, what happens to the other strategies of knowledge-production that social sciences have developed and found valuable, not because they "imitate" the corporate sphere and natural sciences but precisely because they don't?

Under this line of reasoning Big Data is described as techno-solutionism and techno-romanticism (Morozov, 2013; Selwyn, 2014), essentially producing and examining patterns, and as typically telling us more about what people do than about what they say they do (as in surveys or other techniques): "While being able to measure what people do may be valuable, it is not sufficient for all kinds of social science

research. We also need to understand the meanings of that behavior which cannot be inferred simply from tracking specific patterns" (Eynon, 2013, pp. 238-239). boyd and Crawford (2012) give an example of this: while relationships between people may be measured using certain metrics like address book contacts, number of emails sent, and so forth, this does not necessarily reveal a standard meaning or equate to the value people place on those connections.

It is important to note, however, that amid such caution a fundamental critical ambiguity, attended to in subsequent sections, is kept in play: nervousness over EDM and interpretation takes at least two forms. First, suspicion over the slippage from pattern-making (or identification) to correlate and from correlate to cause; and second, the presumed divide between measurement and meaning, quantitative and qualitative, fact and interpretation. In seeking to tether the domination of the former, the numerical, over the latter, the meaning-making, the caution expressed in responses to EDM along these lines makes the latter nonetheless dependent on the former for definition and existence. In other words, rather than contesting the idea that numbers are objective or transparent in the first place, such responses leave the place and role of numbers within an EDM epistemology untroubled, trying to contextualize what numbers can be made to *mean* rather than seeing them as emerging from an already full and complete meaning-system and worldview. The perspectivalism rather than objectivity of number has been contested for centuries - insights that fields such as ethnomathematics have already underscored in pointing to cultures that do not organize reality around the use of number and others that do but are not base 10 (Ascher, 1991; François & Kerkhove, 2010).[7]

Last in regard to educational data mining are the set of challenges that have been posed around issues of structure/equity, and how Big Data may reinforce and perhaps

even exacerbate existing social and educational inequalities in a variety of ways. Whose data traces will be analyzed using Big Data and whose will not appear in the picture is one concern. For Eynon (2013) this implies those with access to the technologies and time to use them, meaning those who are better off, will be represented in the research on social media,. In addition, which researchers have access to which datasets and which are owned by commercial companies, governments, non-profits, etc., remains a concern. This implies that if understandings of reality are to circle more and more around data mining the advantage of a small group of affluent stakeholders and owners of infrastructure is even greater, and control over image-management more exclusive.

For O'Neill (2016), however, Big Data's inequality effects in education have a different kind of reinforcement problem. In particular, reinforcing the intersection of racializing and classist prejudices – what has been labelled algo-racism. For O'Neill these are widespread and already-realized consequences that come in part from the prejudices built into programming and its representational logic. O'Neill argued that the logic from the corporate sphere such as banks has moved into education and goes something like this: *some* people at a certain zip code will default on their loans or not pay their credit card bill on time, therefore loans or credit cards offered to *all* those at that zip code will incur higher interest. Bad zip, bad customer. In education, instead of zip codes its test scores. Algorithms that sit at the intersection of addressing things like uneven test scores between communities of children and the role of teachers and school principals assume direct causation between one representative number and another – bad test scores, bad teacher.

For Amrein-Beardsley who runs the blog *Vamboozled* about problems with value-added models (VAMs) and systems in education, the issue is the logic used to make connections between bits and pieces of information and the conclusions drawn. In

2016 on her blog, quoting the recent US Department of Education (Chiang, McCullough, Lipscombe & Gill, 2016) study, which concluded that VAMs were not effective for evaluating school principals, Amrein-Beardsley underscored how the study noted the problem with transient factors and the lack of predictive validity. Predictive validity was defined in the study as "the extent to which ratings from these measures accurately reflect principals' contributions to student achievement in future years" In particular, Amrein-Beardsley blogged how the study could not identify any factors that drew a principal's performance and a student's test scores into a directly causal and certain relation in terms of such definition[8]: "In short, 'A measure could have high predictive validity *only if* [her emphasis] it was highly stable between consecutive years [i.e., *reliability*]…and its stable part was strongly related to principals' contributions to student achievement' over time (i.e., *predictive validity*).".

Paige (2016a) also blogged about his recent study of the intersection of VAMs and teacher evaluation law, noting not only the lack of predictive validity, but how the hasty adoption of VAMs in evaluation and employment law actually complicates efforts to improve teacher quality, especially at the local level. He argued that their uptake relied upon the following: "that VAMs will effectuate teacher termination with greater ease because nobody besides the advanced statisticians and econometricians can argue with their numbers derived. In other words, if a teacher's VAM rating is bad, then the teacher must be bad. It's to be as simple as that. How can a court deny that reality?" Paige noted that legal circles have subsequently been following cases brought against certain school districts and analyzing their rulings. He argued that VAMs' costs vastly outweigh their benefits and noted that in the lawsuits now coming from discredited teachers the judiciary has favored what administrators' impressions are of teachers and teaching, not algorithms. In such cases, unreliable, arbitrary, or uncertain bases for

decision-making over someone's employment, which he argued VAMs have been shown to be, may actually strengthen the case of the teacher who sues because they would be considered to have lost employment on an infirm basis (Paige, 2016b). Paige (2016a) subsequently advocated for a complete removal of VAMs in high-stakes decisions, blogging recently: "Bluntly stated: VAMs are a statistical 'hot mess.' The American Statistical Association, among many others, warned in no uncertain terms that VAMs cannot – and should not – be trusted to make significant employment decisions. Of course, that has not stopped many policymakers from a full-throated adoption of their use in employment and evaluation decisions."

Such flashpoints have led to some already-expressed cautions, however, in regard to calls for different kinds of accountabilities such as algorithmic accountability (Buckingham Shum, 2016) and more precise identification of just where data makes a difference and which kinds (Prinsloo et al, 2015), pointing to a kind of double consciousness that has accompanied the movement. In regard to educational data mining, Eynon, boyd and Crawford, Amrein-Beardsley, O'Neill and Paige (2013; 2012; 2016; 2016; 2016) are quite overlapping in the dual-edges given to the "kind of" innovation that Big Data is seen as. Eynon (2013) notes, for instance, that exciting areas for research and practice in learning, media and technology fields especially can open up "Yet, as a community we need to shape the agenda rather than simply respond to the one offered by others" (p. 283). Arguing that the focus needs to be on broader ideals, which are distinguished from technical fixes seems a commonly held position: "it is easy for scale and processing power to look impressive. Indeed the conclusions from Big Data studies are likely to offer some neat answers at first glance, particularly in comparison to the messiness of more established social science approaches… The debate needs to be about more than overcoming "barriers" to Big Data (personnel,

equipment, cultural mind sets, etc), so much as ensuring that we use these techniques to empower researchers, practitioners, and other stakeholders who are working in the field" (Eynon, 2013, p. 240). As alluded to above and discussed further below however, the distinctions between ideals and technologies, between the real field workers and numbers, and between qualitative and quantitative arguably remain part of Western "technologies of self" that sustain the very problems that both Big Data and its critics refound.

Learning Analytics

While educational data mining and especially VAMs have attracted stinging criticisms in more recent years, many of the concerns that have been raised within educational settings and outside of them have been overtly acknowledged in Big Data's other point of entry into the field – learning analytics. In a recent overview of the state of the field, Buckingham Shum and Deakin Crick (2016) used the softer language of support. This stands in noticeable contrast to being closed off to contestation over whether someone should be fired on the basis of an algorithm to which they do not have access. The emergent field of LA is described in ways that seem relatively more focused on its human participants along several related lines: "Learning analytics approaches offer in general different kinds of computational support for tracking learner behavior, managing educational data, visualizing patterns, and providing rapid feedback to both educators and learners" (Buckingham Shum & Deakin Crick, 2016, p. 6).

Moreover, their overview of latest research addressed overtly some of the earlier expressed concerns regarding Big Data mining and prediction, noting the presence of power and values:

> These programs expand the boundaries of 'the system' and the stakeholders we should consider in learning analytics, drawing attention to the ways in which power is (re)distributed by such infrastructures, and the many levels at which values are baked into them, with the risk that they become invisible, and unaccountable (pp. 15–16)

In trying to make the rationale for learning analytics explicit, the most prominent point of appeal is a description of reality as in constant change or flux and the need for skills and dispositions to cope with novel, complex situations, turbulence, and a jobs market place "where routine cognitive work will increasingly be automated" (Buckingham Shum & Deakin Crick, 2016, p. 7). In particular, LA approaches have implications beyond classrooms, insofar as they have the potential to assess particular skills, competencies, and dispositions "which are important precisely because they need to be displayed in interpersonal, societal, and culturally valid contexts" across an "arc of life" that exceeds formal learning settings (Buckingham Shum & Deakin Crick, 2016, p. 8).

Because a key focus of learning analytics is computational support for improvement and/or perfection of particular competencies, especially those described as 21st century ones (C21 competencies), LA approaches have focused heavily on getting feedback to the learner under the view that in complex, self-organizing systems feedback is the main source of improvement in performance and for self-regulation. Under this conception of self, the learner is primarily meant to take greater responsibility for their own learning in alignment with a system's operation.

In the context of the "quantification of reality" and within competitive settings, the qualities that learners are thought to need are positioned as signs of adaptation to a rapidly changing world where a short list of similar competencies, dispositions and

skills remain rather overlapping and vague, yet familiar. C21 competencies are variously defined as: critical thinking, creativity, collaboration, metacognition, and motivation; cognitive skills, interpersonal skills, and intrapersonal skills; knowledge, skills and attitudes, value, and ethics; ways of thinking, ways of working, and tools for working and living in the world; or, mindful agency, sensemaking, creativity, curiosity, belonging, collaboration, hope and optimism, and orientation to learning, depending on the sponsoring body, corporation, government, or academics doing the defining (Buckingham Shum & Deakin Crick, 2016).

The computational support is dedicated to delivering actionable insights to educators, students, and other stakeholders who constitute the learning system in question by forging new links between the body of learning sciences research and how data, algorithms, code, and user interfaces come together through coherent design in order to automate such analyses. Foci can be such areas as learner's dispositions and engagement, quantifying conscientiousness through educational games, language technologies that illuminate the quality of interpersonal interaction in textual discourse, etc.

The quantification of reality in education is not dedicated here to producing numbers for the sake of it, however. It is acknowledged as challenging and as a core problem: "Quantifying the deeply personal qualities in order to feed back and strengthen them, without in the process reducing them to meaningless statistics, is at the heart of the learning analytics challenge" (Buckingham Shum & Deakin Crick, 2016, p. 9). The difficulty of translation ensues and in some cases is acknowledged. Some research which positions itself as Educational Improvement Science, for instance, describe what are called non-cognitive factors like mindsets and dispositions such as resilienceas having "for some time been quantifiable from self-report survey measures"

which, in turn, generate "behavioral proxies" that are now beginning to emerge via video gaming e.g., behavioral proxies for qualities such as persistence or perfectionism in children (Buckingham Shum & Deakin Crick, 2016, p. 12). Here, an improvement study is not rendered as cynical and cv-building but as helping and as participatory. It "answers the question 'where should we target our analytics?' because it provides a participatory methodology to work with educators to identify their most pressing challenges, and key drivers; 2) analytics provides new ways to track those drivers and provide the rapid feedback loops critical to improvement cycles, answering 'did we make a difference?'" (Buckingham Shum & Deakin Crick, 2016, p. 11).

Alongside targeted feedback within quantification projects are the importance of visualization projects. LA is dedicated to developing approaches "that demonstrate the power of quantifying and making visible a quality," like student effort. Student effort is a quality that "would otherwise remain intangible and therefore hard to talk about or improve" (Buckingham Shum & Deakin Crick, 2016, p. 13). The rationale is founded on the assumption that by characterizing or coding something as intangible within previous understandings, it has remained impossible to discuss. Here, concern revolves around visualization-as-consensus. Getting stakeholders to agree to definitions, to shared aims, and to metrics is crucial so that at each level of a system the accountability can be consistent, the performance quantified, and the feedback appropriately directed. This is not intended as a sinister system that just spits out judgments. It is targeted at the importance of usefulness and reflection to the concept of improvement: "The role of technology in these tools is to aggregate quantitative data and display it in various summary forms, including visualizations, in order to provoke useful educator and student reflection" (Buckingham Shum & Deakin Crick, 2016, p. 14).

Amid the quantification, visualization, and instrumentalization integral to identifying learning analytics as such, future challenges are noted – the interdisciplinary nature in which expertise in a variety of fields would be required, and how to encourage the engagement of students, teachers, and leaders alongside academics and technologists. This is far from the impersonality associated with data mining and the hiring and firing of teachers whose only recourse seems to be a lawsuit. Moreover, issues for further attention that are identified involve frank discussion of tendencies in some sectors of the emergent field. Akin to the problem of predictive validity in a school principal's performance is the issue of the absence of predictive modelling of "student success":

> Looking to the future, when C21 learning analytics have matured, and in concert, assessment design more explicitly values those competencies, teams will no doubt aim to develop "C21 student success" predictive models. However, caution is advised. Consider, for example, if it makes any sense to present an amber traffic signal to a student because s/he is not behaving as an archetype derived from the activities of previously creative or curious peers, tackling open-ended, authentic challenges. As a complex system, there are so many variables. The prospect of software acting autonomously and "adaptively," on the basis of a classifier using intrapersonal and interpersonal constructs, is fraught with ethical considerations – worthy of deep reflection and values-sensitive design (Buckingham Shum & Deakin Crick, 2016, pp. 14–15).

Caution is overtly advised here, then, as the differences between open, adaptive systems and closed systems that are just seeking prediction for course completion are elaborated. In positioning learning environments as complex systems LA approaches acknowledge

how complexity makes discussion of purposes paramount, which involves establishing processes that matter, underscoring how a system operates at several levels simultaneously, making use of feedback loops, and encouraging the learner to see themself as a self-organizing system. Under this view, pre-determined scripts, lockstepped curriculum plans, and fixed syllabi can no longer suffice as a delivery mode. Describing a shift already noted in systems' theoretical accounts of curriculum studies, what Osberg, Biesta & Cillers (2008) characterized as the move from curriculum-as-representation to curriculum-as-emergence, LA approaches ideally reposition the teacher as informed facilitator and designer, not as superfluous. Teachers or tutors are

> required to received [sic], collate, analyze, and respond to complex data about real learners in close to real time and make pedagogical decisions in situ, in an on-going cycle of improvement. They function more like "learning designers" where they attend to their purpose, the context in which they operate, and the needs of students to synergize these into next best actions. They manage what 'emerges' rather than simply "delivering expert knowledge" (Buckingham Shum & Deakin Crick, 2016, p. 17).

Last, LA approaches also focus on governance and responsibility issues, which implicitly recognize that while there are multiple stakeholders, only some are going to be held responsible for outcomes called improvement or change. This has historically left administrators and leaders in difficult positions, as well as teachers. What, for instance, happens to an administrator's job scenario if one person or group of people "mess things up" (however locally defined) while 95% do not? In discussing complex systems perspectives and the organizational dynamics of introducing LA approaches, the *diffusion* of decision-making through the principle of self-organization becomes crucial:

> This capacity of learning analytics to enable self-organized decision-making at every level has implications for leaders, policy makers, and researchers, too. How can we align all stakeholders around a shared improvement aim whilst also enabling responsibility, authority, and accountability to be aligned appropriately at each level? What sort of leadership analytics and dashboards can support this challenge? How can governments re-think accountability frameworks to support this? Improvement science is an approach that does justice to the micro-level of learning at the same time as an overall macro level improvement aim and learning analytics… are a powerful resource to support this approach because they allow feedback at different levels on processes of learning to the people responsible for change (Buckingham Shum & Deakin Crick, 2016, p. 18).

In other forums, however, concerns are expressed about this potential. In a 2011 blog thread on Learning and Knowledge Analytics, for instance, which involved academics who publish in the field, the hopes and fears of LA's presence in education were candidly expressed, more so than in many peer-reviewed publications. Learning analytics was run together with quantification/statistics/measurement and elicited strong binary reactions. Responding to the prior content, George Siemens - who had initiated the thread, founded the Society for Learning Analytics Research in 2011, and hosted the first Learning Analytics and Knowledge conference (LAK) - wrote at a length worth quoting because it succinctly summarized many of the critical approaches prior to 2011 and published since then. Siemens, who does not see educational data mining and learning analytics as necessarily so distinct, explained:

I'm not content with an emotional reaction or vague dismissal of analytics. Simply reacting negatively to numbers and quantification without some basis on which to explore the critiques is on par with superstition.

I see several concerns arising in relation to analytics:

1. It reduces complexity down to numbers, thereby changing what we're trying to understand

2. It sets the stage for the measurement becoming the target (standardized testing is a great example)

3. The uniqueness of being human (qualia, art, emotions) will be ignored as the focus turns to numbers. As Gombrich states in "The Story of Art": The trouble about beauty is that tastes and standards of what is beautiful vary so much". Even here, we can't get away from this notion of weighting/valuing/defining/setting standards.

4. We'll misjudge the balance between what computers do best...and what people do best (I've been harping for several years about this distinction as well as for understanding sensemaking through social and technological means).

5. Analytics can be gamed. And they will be.

6. Analytics favor concreteness over accepting ambiguity. Some questions don't have answers yet.

7. The number/quantitative bias is not capable of anticipating all events (black swans) or even accurately mapping to reality (Long Term Capital Management is a good example of "when quants fail")

8. Analytics serve administrators in organizations well and will influence the type of work that is done by faculty/employees (see this rather disturbing article of the KPI influence in universities in UK)

9. Analytics risk commoditizing learners and faculty - see the discussion on Texas A & M's use of analytics to quantify faculty economic contributions to the institution

10. Ethics and privacy are significant issues. How can we address the value of analytics for individuals and organizations...and the inevitability that some uses of analytics will be borderline unethical? (Siemens, 2011).

Unlike EDM, which generates criticisms regarding machine-driven, impersonal, anti-contextual, and removed assemblages of data, the range of concerns expressed around LA acknowledges the more concerted effort to have 'the human" involved in the loop and as agentively engaged in establishing the feedback processes which are meant to both position and help them. Nonetheless, there is a core concern in both EDM and LA with eradicating uncertainty and unintended consequences, with resisting and/or confronting an inscription of dynamic systems as nonlinear in practice, and a move into trying to make certain, efficient, clear and pure, causal sequences from instruction (behavior), to feedback, to performance conceived as improvement. Siemens' list seems to sit at a specific and familiar intersection, then: the dissension over "ethics" or "morality" of what humans "ought' to attend to or "normalize," what humans are thought capable of, and what new technologies make available to the imagination.

Beyond the Binary?

The above outlines a basic schematic range in the debates over Big Data's entry into education. If one were to "stylize" the responses, Big Data appears as avante-garde, or, as leading education into a new era, in particular, a more certain, more positive, and

less fuzzy direction by deciphering universal and replicable laws of learning, patterns of feedback delivery and behavior, and probabilities of outcomes. Here, previously unavailable diagnoses and experimental modalities can transform education's understanding of its nebulous self, especially via data mining. Moreover, by delineating with greater specificity the macro-micro dynamics of pedagogy via learning analytics' more inclusive, student-friendly, disability-aware, and personally tailored programs, with immediate feedback, what seems on offer are more responsive and adaptable performance indicators for meeting clear goals, with the latest competencies necessary for digital natives and new millennials. This presents new ways of being a teacher and a student that enable relevant, real-world kinds of flourishing which would otherwise be thwarted. The silos and 20th century tendencies to cling to static and stultified versions of curriculum, learning, and evaluation have contributed to the reproduction of the very problems so often complained about. Big Data enables the link between the local and the global to be made manifest and concrete, to better understand where to place funding and energy to solve entrenched problems. New sources of information, new methods of relating them, and new approaches for applying them through systems-integrated feedback loops are needed to move education forward in recognition of its multiple stakeholders and in partnership with institutions that lie beyond the classroom in new 21st century circumstances.

On the other hand, Big Data is positioned as an extension of quantification and standardization strategies; contributing to a limited range of questions; suited to computer-based coding logics; and driven by fear of uncertainty, discomfort with ambiguity, and desire for control via prediction. Its unspoken abjection of multiculturalism via standardization of only certain cultural norms demonstrate the fear of genuine diversity at its heart, amid efforts to minimize challenges to existing

power structures. Giving different worldviews or cultural processes "numbers" doesn't account for "context." The imbalanced distribution of privilege and resources are built into the language of data - a disguise for very emotional desires for controlling the rules of the game through coding, algorithms, and ownership of infrastructure. The ethical issues it raises as well as its potential for colonization are to be tethered by addressing first a broader democratic discussion about the different purposes of education within rapidly changing social compacts. Here, the humanistic capacities of stakeholders and the critical orientations of researchers exceed automatic reverence for combining or combing large data sets or elevating the disciplines of psychology, mathematics, and computer science as though definitive of being and learning. Especially if the results of such studies are to support workers, students, and participants in the field who are from diverse backgrounds, living in conditions that are not even or equal, and who hold multiple worldviews that are not necessarily reconcilable within or beyond "the West," Big Data in education needs to better understand its prejudices and its place.

The binary reactions do not leave researchers in a strong position from which to agentively reconvene conversations about the politics of knowledge-production outside of recycling the same criticisms and responses that have already been raised in regard to quantification and standardization in the field. Consider the dilemma that many researchers in the social sciences now find themselves in – data scientists or not. On the one hand, a researcher can invent a category, collect "information" that they say belongs to the category, tell a story around (numerical or otherwise) nodal points, and claim it is empirical. This raises the typical rejoinder: if selective use of whatever is designated data can be deployed to legitimate any view we like (or dismiss ones we don't), then what is the point of it and what are researchers really involved in? On the other hand, the specters of relativism (anything goes and all views are equally valid) and nihilism

(there is ultimately nothing, so why bother?) quickly appear on the horizon, raising issues of potential obsolescence - institutions beyond their use-by dates. Why have academic disciplines, universities, research, and knowledge-production at all, then? Are there really experts and exclusive practices that must be learned? Why not live in a world where you can just launch your opinion, pretend its justifiable with either numbers or interview transcripts, and then get to the "real" work of using whatever narrative has been constructed for persuasion and image-management with "data" thrown in from the side, a bit like bubble wrap around something breakable – a buffer to stop the slippage. If authority and legitimation are ultimately founded on consensus, and if consensus does not need an expert, fact-checking, or rigorous review, then the castle of research has finally been stormed - and fallen. But then what? Endless party politics and echo chambers? Infinite he-said, she-said?

This stark portrayal of conversations that readily arise in both formal and informal academic settings of late and of which Big Data is but one flashpoint indicates a series of controversies that are not easily resolved by staying within the logic that created them (if, that is, resolution is indeed a goal). Derrida's (1990) work has for a long time already centrally engaged this problematic of "flattening" relative to "exclusion," not the least of which appears in his discussion of the force of law and the mystical foundations of authority and his deconstruction of the nature/nurture binary in the human sciences and anthropology (1970). I want to engage the issues here from a different direction, by approaching the conditions of possibility for such binary positions and quandaries which form at this moment around "data" and the emergence of Big Data in education. To do that, several longer, albeit punctuated, journeys are required to take a step back from the immediacy and seduction of binary reactions to new phenomena.

Much has already been made available regarding the history of numbers, mathematics, statistics, and the quantification of reality (Dear, 1995; Hacking, 1990; Porter, 1996; Stigler, 1990; Stiegler, 1994; Poovey, 1998). The aim is not to replay those debates here. Ultimately though, such a trajectory would not matter if numbers-as-data could not make a mark on someone about something. Imagine presenting a series of numbers to someone who has never seen "information" coded as figures, asking them what they are going to do about it, and then getting the counter-question "What's your point?" The debates around Big Data might be understood more broadly, then, to at least entail something in addition to a quantitative/qualitative rift. First, such debates are generally accepted as being carried out between humans who are thought to have selves that hold specific, different, or consensual opinions that can be launched with or without overt reference to "data" or to "number." Belief in self is a repetitive pattern in Western thought, yet it seems to emerge out of a unit repetitively characterized as social, as something more than a One. To trace how data becomes available as a construct, then, the possibilities for knowledge-production that preceded its naming and the possibilities for pinning Becoming/Being to a selfhood that would be impacted by "data" will be attended to.

Technologies of Self and Microstrategies of Conversion

Foucault (1985; 2005; 2008) argued that in the West truth effects and truth-production are not reducible to the exclusionary effects of archaeologies of the academic disciplines or genealogies of institutional systems such as asylums, clinics, and schools, but that investigation of how a being came to see themselves as human in specific ways and in particular, as able to do particular kinds of work on something

believed to be a self and upon others was also required. This is most forcefully articulated in *The Hermeneutics of the Subject* lectures and *The History of Sexuality* (volume 2), *The Use of Pleasure*. In the opening of the latter, Foucault (1985) explained the long detour his research had taken, arguing that he was not writing a history of moral behavior, nor a history of codes and regulations, but a history of "truth" (he suspended the term). In the newer project, he argued that the research must take a different direction: "It seemed appropriate to look for the forms and modalities of the relation to self by which the individual constitutes and recognizes himself *qua* subject" (Foucault, 1985, p. 6). In doing so, Foucault deployed as substitutable the terms technologies of self and techniques of the self, as well as art/s of existence and aesthetics of existence (Hattam & Baker, 2015), elaborating the unique dimensions of his detour in which "the modern individual" turns into an object of study, becoming a particular kind of subject:

> What I mean by the phrase [arts of existence] are those intentional and voluntary actions by which men not only set themselves rules of conduct, but also seek to transform themselves, to change themselves in their singular being, and to make their life into an oeuvre that carries certain aesthetic values and meets certain stylistic criteria. These 'arts of existence', these 'techniques of the self', no doubt lost some of their importance and autonomy when they were assimilated into the exercise of priestly power in early Christianity, and later, into educative, medical, and psychological types of practices. Still, I thought that the long history of these aesthetics of existence and these technologies of the self remained to be done, or resumed (Foucault, 1985, pp. 10–11).

Crucial in this enfolding of questions of "being" to a self and aesthetics of existence was the what and the where of "the real." The narratives that conferred

legitimation on the what and the where of "the real" or under what conditions it could be recognised or experienced over the non-real required interrogation. Rather than automatically invoking disciplines or institutions as the sites or arbiters of the real, Foucault's (2005) *Hermeneutics* lectures in particular were preoccupied with identifying "the movement by which, in ancient thought, from the Hellenistic and imperial period, the real was thought as the place of the experience of the self" (p. 465). This shifting triad of self-knowledge-reality eventually becomes "institutionalized" within contemporary "disciplinary societies" in multifarious but not infinite ways, ways in which education and more recent movements like Big Data continue to participate.

In Foucault's (1985) positing of different technologies of self, however, technology does not refer to a tool separate from humans, to electricity outside the body, or something prehensile, but to practices and operations within a teleological frame, practices "which permit individuals to effect by their own means or with the help of others a certain number of operations on their bodies and souls, thoughts, conduct, and way of being, so as to transform themselves in order to attain a certain state of happiness, purity, wisdom, perfection or immortality" (p. 18). On this view, the terms *curriculum* and *computers* would both be understood as technologies.

Different technologies have relied upon different relations between "truth" and "the Subject", different arts of existence. The post-Cartesian epistemes were rather different from what had gone before. First, post-Descartes recognition of self had to go through a construct called *mind*. Second, the notion of "care of oneself" (*epimeleia heautou*) had been subsumed, or even forgotten, in favor of the trope "know yourself" (*gnōthi seauton*). This idea of *epimeleia heautou*, for Foucault (2005), is a defining mode of thinking for ancient Greek and Roman philosophers and early Christians (from the 5th

century BCE to 5th century CE) and can be considered as "an attitude towards the self, others and the world" (p. 10).

The big-picture upshot is that in a Cartesian epistemology, one does not need to put oneself necessarily in the right condition to receive truth. The search for what Descartes called clear and distinct ideas can take place instead via following appropriate methods; knowledge is believed to be produced in a site ejected from the body (knowledge as *connaissance* rather than as *savoir*), and through this, consciousness is standardized as access to Being via mind. For Foucault (2005), Descartes thus constitutes that pivot point *when science splits from philosophy and spirituality*.

Such hermeneutics of the subject typically involve a perception of body as flesh encasing a more highly valued ontological principle – soul or its analogues. The specificity and inscription of soul, self, and the subject remain to be seen in each case, changing according to the philosophical or theological school and the preferred technologies of self. Significantly, amid the range, Foucault (2005) posited an important and constant presence of *an image of return* in the soul-self nexus that marks conversion and its relationship to truth-production. The microstrategies of conversion and this relationship to knowing are apparent in Hellenistic and Roman thought, while the relation of return is refigured in early and later Christian thought. Such strategies include how self withdraws from its surrounds (Platonic *epistrophé*), how self flees from self (medieval Christian *metanoia*), and how self splits from self to be under its own eyes (modern *subject-as-object*).

Although the image of return (to a homeland, to God's law, and to a true nature, respectively) continues to operate, the soul-self nexus is constituted of a different "substance" in each tradition. Amid these differences, Foucault (2005) noted one further continuity across ancient and medieval thought - the presence of the Other-as-master:

"the self is actually something that always has to go through the relationship to someone else who is the master" (p. 58).

Three major lines of evolution are described, and recurring questions arise within the changing specificity of self-knowledge-reality relations, including: 1) *dietetics*: the relations between care of the self and medicine, treatment of the body, and regimen which become increasingly intertwined in the history of the West (Foucault repeatedly uses the terms West and Western); 2) *economics*: is the relation between care of the self and social activity - including private duties of the family, landowner, and slave-master which constituted the domain of ancient Greek "economics" – compatible or incompatible? and; 3) *love*: must care of the self go through the love relationship or should care of the self and the erotic be separated? (Foucault, 2005, pp. 59–60).

The range of technologies and microstrategies of conversion, with their commonalities was not surprising for Foucault, given the long history of debate before and after the spatialization of spirituality, philosophy, and science over care of the self and the presumption that "the self" was the site of experience of the real. He did not see such practices, technologies, or aesthetics as exclusively Western, however. The range includes practices he posited as typical well before Plato and beyond Greece: "there was, if you like, an entire technology of the self related to knowledge (*savoir*), whether this involved particular bodies of knowledge (*connaissances*) or overall access to truth itself. The idea that one must put a technology of the self to work in order to have access to the truth is shown in Ancient Greece, and what's more in many, if not all, civilizations, by a number of practices" (Foucault, 2005, pp. 46–47). This includes *rites of purification*; techniques for *concentrating the soul,* such as avoiding dispersing the soul, breath, or *pneuma*; *withdrawal* or disengagement from the world; and the *practice of endurance* where painful and hard ordeals are borne or temptations resisted.

The Hellenic, Roman, Christian, Gnostic, and post-Christian thought with which the *Hermeneutics* lectures were more directly concerned foregrounded and backgrounded these measures differently. It is important to outline how so, for such a glimpse opens up the ways that education and Big Data can be seen participating today in dis/continuities in the onto-theo-philosophical assumptions and heritages that drive 21st century claims to truth, reality, and description of competencies.

Platonic *Epistrophé* and the Return to Homeland

Big Data is predicated in part on the idea of producing or finding patterns via connecting large databases with each other. Finding *meaningful* patterns amid the noise of all the information points that are available, with only necessary variables filtered in, is thought to offer clarity, precision, and a broader view that a local perceiver is just not capable of when embedded in the coalface of teaching. Its promise is an Apollonian eye above and beyond the classroom interface. To a certain extent, this more "removed" and seemingly global and clearer aesthetic reflects elements of the Platonic *epistrophé*, in modified form.

In Platonic *epistrophé* the central theme is the liberation of soul from body, which requires a withdrawal from certain things in the external world and return to the self in a better "corrected" state. But what is that self to which a reliable return can be effected? While different versions are articulated in different Platonic texts, in *Alcibiades* especially, the soul-as-subject rather than soul-as-substance is revealed in its complexity as the "thing" of the self, as that which is the subject prior to the action, behavior, or goal in the external world that "the body" enacts. The link assumed between knowledge and self-formation, posed in the form of veridiction, thus downplays the status of body and is here constitutive, opening a place for the role of

philosophy and a heightened status for soul. Once predicated on opposition between this world and another and a dissatisfaction with how the immediate external world can corrupt, transformation is posited as necessary. In assigning a privileged role to philosophy in the form of knowledge-as-*savoir*, coupled with a rejection of aspects of the outside world, a confluence is enabled between *ascetic techniques*, the *task of preparation,* and the inscription of "life" as a *life of testing* (Foucault, 2005).

While different responses and formulations in later Hellenic and Roman thought are elaborated in regard to care of the self tropes, for example, between the Stoics and the Epicureans, Foucault (2005) argued that, in general, up to the end of the second century CE, the major principles organizing practices of the self included at least two aspects that in Western educational theory today are quite counter-intuitive. First, *integration* in which the intertwining of the practice of the self with the art of living was assumed. Here, care of the self was not a confined pedagogical practice reducible to teaching youth within a single institution like a school. The practice of self was no longer a turning point marking the difference between childhood and adult life, for it involved a critical aspect rather than a training function, correcting rather than teaching, and often in a dialectical one-on-one relation. This meant that "self"-development was not seen as preparation for a "life" that was to come after compulsory schooling. While the objective of the practice was preparation for old age as the site of full emergence of the subject, the practice of the self was at one with, or merged, self and life (Foucault, 2005, pp. 126–27), an integration guaranteed not by endless sensory stimulation but the reverse, the turning away from appearances and the recognition of their illusory or deceiving nature.

The second characteristic expressed in Hellenic and Roman thought was its unqualified and inclusive nature, for it appears that everyone without any prior

condition of status or any technical, professional, or social aim could practice care of the self. However, while all could practice, only some were considered capable of the full and complete status of the subject of the self, those closed off in religious settings or cultural segregation where the techniques were refined. This reinforced the dominant image, the image of return, characterized as that in which "We must turn away from everything that is not part of ourselves but which grabs our attention, our diligence and arouses our zeal. We must turn away from this in order to turn round to the self. Our attention, eyes, mind and finally our whole being must be turned towards the self throughout our life" (Foucault, 2005, p. 206).

Several aspects of *epistrophé* reverberate in the presumptions of Big Data's rewriting of self. First, the version of withdrawal and return, purified and perfected beyond errors, operates similarly via selectivity and cleansing – only some things are turned away from while others are drawn into mattering as relevant or meaningful data. The purification is sometimes unidirectional, especially in EDM where "the human" is not put into the loop to speak back to the classifications and categories or create new ones. Second, the self's locus is positioned as beyond the immediacy of a classroom, or any institutional, interface. In EDM, for instance, to "know thyself" requires moving away from the messiness of daily pedagogical action and taking up the Apollonian eye of the observer who codes behaviour from a distance. Third, the definition of knowledge is driven by presumption of a pre-existing homeland – a place lying in wait for (re)discovery – which does not explicitly announce its commitment to absolutism. The selectivity operationalized - *which* appearances are turned toward and *which* are turned away from in the reconstruction of homeland - is, then, rarely made explicit in some incarnations of Big Data. At a broad schematic level, the "action" and practices born of belief in an *epistrophé* refigures homeland as associations between

filtered pieces of information ("data"), creating new territorializations of self as a) stripped down (epistemologically purified of errors in perception) and b) systemically integrated into a prefigured absolutist atmosphere (predetermined).

There is latitude in *epistrophé*, however, that falls away in later eras, and this reflects the difference that learning analytics has taken up. Rather than morality defined in terms of "obedience to a system of rules," personal ethics in Antiquity for Foucault (2005) were understood in terms of an "elaboration of one's own life as a personal work of art" or "the search for an aesthetics of existence" (p. 49). Importantly for Foucault, Greco-Roman philosophy relative to the later priestly control of Christianity indicated that "people are not told what they ought to be, what they ought to do, what they ought to believe and think . . . people are left to make up their own minds, to choose, in the light of all of this, their own existence" (Foucault, 2005, p. 50). Here, Foucault moved beyond a simple liberal notion of choice or curiosity and potentially referred to the seeding of double-edged dynamics he described elsewhere (discussed further below) as a "technology of the environment"- as those practices that contribute to an atmosphere, to be subtly absorbed as ambient – practices that allow certain functions and give the appearance of latitude within a limit that is not meant to be grasped, pointed out, or manipulated, but creatively operated within.

Medieval Christian *Metanoia* and Self-Transformation

Attempting to reduce human behavior, performance, and potential to algorithms is no easy job (O'Neill, 2016, p. 5). Via medieval Christian *metanoia* as a microstrategy of conversion, the imperative to remake the self as obedient and dependent on the rules that have garnered cultural and legitimating authority or

"force" becomes more noticeable. Such rules come through Church-based and Biblically-referenced sources, not through anything akin to today's databases or an ancient Hellenic pantheistic pantheon in which "humans," animal, and gods could transact, mate, transmogrify, and reincarnate. Medieval Christian *metanoia* operates within an "examine yourself" framing in which *metanoia* is inscribed as a positive term entailing both penitence and radical change of thought and mind.

The term also draws upon the theme of a world opposed to another world, but transformation occurs in a different manner. The Platonic *epistrophé* invoked the theme of "conversion" based on the gap between this world and another which involved studious movement from the world below to that above, finding perfection when the soul is returned to its source and places itself back once again within the eternal movement of being (Foucault, 2005, p. 216). The "inbetween" worlds of Hellenic and Roman thought, neither fully the Platonic *epistrophé* of "know yourself", nor the Christian *metanoia* of "examine yourself" entailed a more ambiguous "turning your gaze on yourself", the content of which means turning the gaze away from others and from things in order to establish your own aim. This is not quite the same, Foucault suggested, as what happens within Christian *metanoia* where there is the *formalization* of a call *to constitute oneself as an object* and where, for instance, *vigilance* to the precepts and thoughts arising within and their sources, signs, and analysis, occurs.

Another of the differences that Christian *metanoia* entailed was that of *sudden* change: "Whether or not there is preparation, development, effort, ascesis - conversion anyway requires a single, sudden, both historical and metahistorical event which drastically changes and transforms the subject's mode of being at a single stroke" (Foucault, 2005, p. 211). In addition, this upheaval involves a transition from one type of being to another, from death to life, mortality to immortality, devil to God, etc. And

third, this must occur in the subject, a dying of oneself, and being reborn in a different self. A unique and dramatic event is thus required in which there is a passing from death to life, from darkness to light, in which a rupture within the subject occurs, enabling renunciation of the former self.

At a surface level it is easy to see here the resemblance to how Big Data's entry into education potentially changes the "self" of schools, districts, teachers, students, and universities via a kind of *metanoia* that entails a breach with a former self's version of performativity. The new version of being ushered in may be received as though educational professionals are meant to be born again by revelations in the form of pattern-recognition via Big Data analytics. Instead of opposing one world with another world like heaven/hell, however, there is a flatness and transversality to Big Data movements in education that cuts off the possibility for transcendence to anywhere else. The passage from darkness to light indicative of medieval *metanoia* is forged in Big Data through interconnection not monasticism, by buying into multiple points of contact not stepping out, indicated by reverence to how a computer program and an algorithm enable you to see your "self" in and as correlates. Like Christianity, though, the seeing can only take place in the terms that such systems can cope with in prefigured ways.

Big Data's commitment to linear logic, to largely numerical databases, as part of a system of reasoning that seeks predictive validity between select variables, is far removed at one level from placing all causality in the hands of an invisible Creator, but not so far on the other from making appeals to analyzing the *worth* of doing something. When O'Neill (2016) argued, then, that algorithms are "an opinion embedded in a code," the insight cannot help but arise as direct contestation of *metanoia* mentalities which indicate that the Church's hold on village life in Europe has indeed been broken but

perhaps in content more so than form: "Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful were beyond dispute or appeal. And they tended to punish the poor and the oppressed in our society, while making the rich richer" (p. 3).

This raises the possibility of excavating heritages in Big Data's genealogy that see it wedded to and stuck within familiar definitions of what analytical rigor "ought" to look like and the problem of causality. If one substitutes the language of Christianity for the language of Big Data it is clear how familiar at the level of practices of self the new master is: instead of linearity, Christian vigilance toward whether continuity between self and aim was sustained; instead of numbers and data points, sign after sign, list after list, of what arises internally as a thought and as a desire; and instead of patterns implying correlates and in some cases cause, an analysis of how and whether the thought arising is good or evil, should be carried out as a behavior, or if it was carried out, whether it was inspired by good or evil thoughts, and thus lists of virtues and vices.

The point here is not the inherent goodness or otherwise of Big Data or Christian variations on conscience, however, but whether the innovation that is offered at the level of analytical practices and the search for patterns and causal explanation, reducible to information that must take already prefigured forms and adhere to certain norms, is really that innovative? Under this line of questioning, *metanoia* prompts a search for more and more data of different kinds and range, allied with a rigorous inspection dedicated to governing the self more efficiently to achieve its aims. The questions that such reverberations of *metanoia* leave us with, then, include whether Big Data is to be understood as a game or landscape-changer, new frontier, and new era; or more as an elaboration of prior tendencies, within which echo practices of self in regard

to the rules of a teleological and salvific game established by a specific clergy and reinforced with inducements - inducements that assume that putting a *line and a link* between self and aim is "good" and that tightening it is even better?

Modern Subject-as-Object

My sense is that there is something more and something different going on in contemporary movements like Big Data, however, than mere echoes with prior modalities of relating self, knowledge, and claims to reality-as-morality. Recognizing patterns, whether putatively here in historiographical and philosophical terms or in numbers related to more numbers, is both easy and endemic. It is, however, in the chasms forged and then variously targeted for reunification that the possibilities and the limits in Big Data's emergence most sit, and that most distance it from compulsory education's more developmental and holistic tendencies and differential understandings of the social. At the heart of this chasm is difference-making "itself" – claims to difference/identity as the centerpiece of a modern episteme, in which distinction from, rather than similarity to or resemblance with, becomes the basis for claiming that one has produced knowledge.

In medieval Christian *metanoia*, it is clear that one is working on the "worst" parts of oneself in order to purify for salvation. There are rules, though, for what constitutes the worst and best. These are expressed and administered by the clergy-as-proxy and interlocutor and the sources of authority such as the Bible. These sources reinforce the route to knowing and enable the (old) self to be split so that a new self can emerge. In the making of "the modern individual", of subject-as-object, the master is internalized in new ways that formalize method as an exteriorized process. One is no longer going overtly through a personified authority like a minister, priest, teacher, or

God but increasingly going through the regime of truth called science, which constitutes the organizing principle for self-analysis and transformation. Here, in the post-Cartesian moment in which Foucault places the beginning of the modern history of truth, "knowledge and knowledge alone" (as *connaissance*) gives access to truth. Relatively speaking, the philosopher or the scientist can recognize truth and have access to it solely through the activity of knowing, without anything else being demanded, without having to change or alter their being as a subject.

Whereas induction into procedures, such as following the rules of training, adhering to certain methods, observing the structure or nature attributed to the object and so forth, may be required, these activities do not generally concern the subject's *being*. "The subject" here does not have to go through a conversion process or put themself necessarily "in the right condition" in order to receive truth-as-revelation. They do not have to purify themself through meditations and prayers in order to upgrade their circuitry and withstand "the light" that God's brilliance will shine upon them when they are apparently ready for it. Rather, "the subject's" access to truth is defined within knowledge-as-*connaissance,* and this marks a different age of the history of relations between subjectivity and truth:

> The point of enlightenment and fulfillment, the moment of the subject's transfiguration by the 'rebound effect' on himself of the truth he knows, and which passes through, permeates, and transfigures his being can no longer exist. We can no longer think that access to the truth will complete in the subject, like a crowning or a reward, the work or the sacrifice, the price paid to arrive at it (Foucault, 1980, pp. 18–19).

On the surface, this might be the most obvious site of disconnect between education's versions of developmentalism and holism, and Big Data's entry. The direct

human-to-human interaction of prior orientations seems more aligned with the *metanoia* conversion process, whereas EDM and LA seem relatively more aligned with allowing more things to run on automatic method, exteriorized from body and made visable, regulatory and formulaic, no matter how much the human is placed in the loop or enters into the connection between data points or feedback units being assembled and distributed.

Again, on the surface, this is a crucial distinction, for it seems to underscore the disconnect in kinds of rationality between other educational approaches and Big Data. Chertok and Stengers (1992) argued, for instance, that different versions of rationality already separated the theoretico-experimental sciences from the ethico-redemptive ones in the 19th century and that this continues:

> The theoretico-experimental sciences are distinguished by the practice of making their version of "reason" depend on the power to "give reasons" for or to explain phenomena. This version of reason thus presumes the power of predicting outcomes, of controlling in order to replicate, or purifying to insure the implication of a theory – the power, in sum, to make a phenomenon 'admit' its truth (p. ix).

Key in the theoretico-experimental sciences is a version of rationality in which purification - not of self, internal desires, or gripping vices but to a clear cause or set of causes - operates as the main analytical and narrative principle. In addition, in such a purification process the search for causal variable/s to explain an effect that is *replicable across contexts* is that which is thought to produce knowledge. The mantra of VOR (Validity, Objectivity and Reliability) remains paramount. In the case of Big Data, reason – whether enacted by human or machine programmed by a human - is thought to be demonstrated by definitive coding of phenomena, quantifying performance,

visible behavior or declared attitudes, or by error correction via feedback loops where the patterning becomes both the truth and causal and where graphic or visual display is the key format of representation (Myer-Schoenberger & Cukier, 2014; Long & Siemens, 2011).

In "the soft sciences" which Chertok and Stengers (1992) refered to as the ethico-redemptive sciences, however, rationality is comported and recognized as rationality differently. Different tactics of reason, evidence, and causality emerge on the basis of *intersubjectivity*. This difference, Chertok and Stengers argued, is captured most ardently in the difficult case that animal magnetism and hypnosis have posed to "reason" in the West and especially its key feature - *suggestibility*.

In their analysis of the formation of mind sciences, Chertok and Stengers (1992) argued that it is in psychoanalytic theory and practice especially that the difficulties of reason arise, for such a human science "does not simply reproduce the model of other rational practices. The 'heart' to which psychoanalysis addresses itself is not conceived in such a way as to guarantee a science resembling other sciences (by contrast to the role 'behavior' plays in experimental psychology, for example)" (p. ix). Rather, two differences mark the uniqueness of the human sciences. First, one objective was to create a practice that would render intelligible the obstacle that "heart" poses to the efforts of "reason" (Chertok and Stengers deliberately suspend the terms). In today's terms, it would be like asking why giving feedback to someone does not necessarily result in improved performance. A second objective was to create a practice that would *not* be limited to making "heart" an object of science, like any other, only more complex. Imitation of the theoretico-experimental sciences was not the aim. Again, in today's terms, it would be like understanding that trying to turn teaching into physics was never the desire or aim; and instead, the project was understanding the unique

specificities of teaching on its own terms. This may in turn expose the limits of physics, but this is not the goal, concern, or reference.

The point here is not about hypnosis *per se* or recommending it. Rather, Chertok and Stengers (1992) focused on an event such as hypnosis because it has had a troubled relation with scientific reason. A hypnotizer impacting a subject's behavior so thoroughly is indexical of limits in theorizing causality and consequences, including the limits of organizing, enforcing, and evaluating behavior through control of purified variables. Who shall we say is responsible for the behavior observed? Whose self was it that was performing the command? To that end, the practical invention of the hypnotic relation "has had the effect of providing a privileged terrain where "heart" and scientific reason confront each other, a terrain where proclamations of rational conquest alternate with admissions of defeat" (Chertok & Stengers, 1992, p. x).

The different versions of rationality in different disciplines reflects much of the "back and forth" so highly valued in educational approaches and often problematized if not despised beyond them. Again, the issue is not "Why not study hypnosis better and see what we get?" or "How can we control what happens better?" Rather, the point is about how disciplines form differently; *it is about the cost of becoming an object of study in certain terms*:

> To the extent that hypnosis has indeed become (principally in the United States) a phenomenon subject to experimental research, the controversies surrounding that research teach us less about hypnosis itself than about the *price* the experimental ideal requires a phenomenon to pay if it is to become an object of study. And it is not a matter of hypnosis being incapable of paying the price, but rather of that payment making the hypnotic relation not a scientific object but instead a deceptive mirror where the very

> ambition to submit it to science is reflected. In that respect, we believe that the French attitude toward hypnosis – the "irrational" character the French attribute simultaneously to hypnosis and to interest in it – uncovers a question left in the shadows of the American attitude that hypnosis in "normal." Can rationality be defined by a standard pertaining to the phenomenon to be studied and not to those who study it? Can the operation of purification, of creating the experimental setting, be conceived as "right" of reason? (Chertok & Stengers, 1992, p. xiii).

Rationality as what or who, and right of reason as design or designer are the problems brought to noticeability through the hypnotic relation and the effort to submit it to the experimental ideal. Here, the historical account "suggests" that a nexus, sometimes construed as a problem, in which cause-effect, visibility, and the "proper character" of the researcher who reflexively moves between "heart" and "reason" composes the domain of human sciences differently from other disciplines.

Suggestibility becomes a *problem*, then, only from some perspectives: "the question of suggestion always arises when 'heart' and 'reason' are no longer conceived as being in opposition, when 'heart' is no longer considered an obstacle to the legitimate power of (theoretico-experimental) reason" (Chertok & Stengers, 1992, pp. xvi–xvii). Because "the infant's relations with its caretakers are already characterized by what we should recognize as a form of suggestion" (p. xvii) the social sciences which focused on direct human-to-human relations could not so readily make a phenomenon admit its truth via a purification process: "suggestion puts 'truth' in question, that is, it problematizes the possibility of constructing a theory on the basis of experiment or experience. Suggestion is impure; it is the uncontrollable par excellence" (pp. xvi-xvii).

A more ethico-redemptive orientation sees unintended consequences as built-in, then, not necessarily as something that can or ought to be eliminated within a social efficiency model that wants the fastest and cheapest line between self and aim. It embodies attention to continuous movement between heart and reason and attention to the failure of reason. This movement and attention is accepted as the definition of rationality, with causality understood as impossible to parlay into purifications that are replicable across contexts: "The greatness of psychoanalysis resides, we believe, in the fact that its failure forces us to pose the problem of "reason" itself, and more precisely, the problem of the model of rationality guiding modern sciences" (Chertok & Stengers, 1992, p. viii).

Like Chertok and Stengers' historical account, in Foucault's (2005) technologies of self, the emergence of ethico-redemptive sciences did not mean the absence of rationality or of rigor. Neither did the formation of such disciplines mean the end of spirituality nor necessarily total divorce from it in all forms of research. Post-Descartes, when science putatively splits from spirituality and philosophy, the role of practices and concepts historically involved in spirituality still circulates as a queston in certain orientations, including and especially in psychoanalysis and Marxist approaches. For Foucault, these approaches, while familiar within the human sciences, are not *exactly* sciences. They should not seek assimilation into science, but *are* forms of knowledge, of the order of *savoir* rather than a methods-driven *connaissance*, and as such, house certain reverberations:

> We should not forget that in those forms of knowledge (*savoir*) that are not exactly sciences, and which we should not seek to assimilate to the structure of science, there is again the strong and clear presence of at least certain elements, certain requirements of spirituality….It goes without saying that it

> would be completely wrong to identity these [forms of knowledge like Marxism and psychoanalysis] with religion….However, if you take each of them, you know that in both Marxism and psychoanalysis, for completely different reasons but with relatively homologous effects, the problem of what is at stake in the subject's being (of what the subject's being must be for the subject to have access to the truth) and, in return, the question of what aspect of the subject may be transformed by virtue of his access to the truth…which are once again absolutely typical of spirituality, are found again at the very heart of, or anyway, at the source and outcome of both of these knowledges (Foucault, 2005, p. 29).

Such insights, which explain the differences in "the feel" of education's versions of developmentalism and holism relative to Big Data, help account for some of the stinging critiques, which see/frame Big Data as "removed' or heartless. It may also help to explain how the concepts of cost, the price paid, and efficiency are being redefined. For instance, the "price paid" occurs in Big Data's uptake, not in advance in the form of the "the subject" getting themself into the right condition; but afterward, as an imposition of a specific subjectivity through a lateralism that is meant to be "chosen" and "adapted" to.

This is sometimes posited as a characteristic of neoliberalism and met with the critique that the coming into "being" of "knowing" as sets of information-points produces "reality" as "quantity" or "unitized," which then requires "the subject" to modify themself accordingly, a posteriori. Such a "flattening" of knowing does not offer an out or a beyondness, to an other-worldly or heavenly ethical or principled system, where a transcendent Platonic Form might reside revealing absolute Truth. It does not point either to a *metanoia* or even an Enlightenment where the goal of the religion

might be to eventually teach "the subject" to exceed all religion. Rather, the critique here is typically that Big Data's version of logic and of theoretico-experiemental rationality pins one and all into a network (digital, electrical, financial and juridical) where there is no in or out, above or below. Along this line, neither transcendence nor immanence are possible. Systemic integration operates instead as enchainment and repositioning into a new trope of associationism across complex interconnected platforms, in which there is something other than gods (*epistrophé*), God (*metanoia*), or the nature of Man (modern individual) operating as the master and decisive agent, judging how well you use your "agency" and how much you have demonstrated "mastery" - an enchainment that "the subject" was asked, encouraged, rewarded, and made, to actively encircle around their "own" legs and champion as competence.

At the same time, there is something utterly repetitive and salvific within this very line of critique, not just in "rescue us from Big Data!" mentalities but also in "Big Data to the rescue!" mentalities. The concern for *improvement*, however organized, is a moral judgment premised on dissatisfaction with a present state, however defined. While the coming-into-being of Big Data and cautionary responses to it may expose different conceptions of rationality in the disciplines and their "feel," the very concern to demonstrate rationality in some format and for improvement to result taps into modified spiritualist heritages around practices of selfhood. What is remarkable, however, is not the complicated lineage where human sciences retain some aspect of practices once historically associated with spirituality-as-religion and their own version of rigor, without such human sciences *being* religions or *being* hard sciences. Instead, what is remarkable are the contemporary tendencies that fuse the silos and that may give pause for thought in other ways about whether "differences" are as oppositional as once claimed. This leads us, finally then, to consideration of such contemporary

tendencies which potentially change the game and the landscape of Westernized conceptions of selfhood.

Technology of the Environment

The above has considered some common controversies around Big Data and offered an outline of some conditions of possibility for its entry into education via technologies of self and microstrategies of conversion. Big Data arguably elaborates some prior tendencies in practices of self-knowledge-reality construction in a variety of ways: 1) from Platonic *epistrophé*, cutting through the noise for an absoluteness about the nature of things revealed as incontestable Truth; 2) from medieval Christian *metanoia*, more data and better kinds, with closer inspection and tightening of self to aim; and 3) from the modern reflexive subject-as-object, consideration of the most efficient and best places to place analytics and data.

What seems to exceed such elaborations, echoes, and reverberations, however, is where the "sites of action" now lie. In the footnotes of his lectures based on pages of manuscript scribblings on "the birth of biopower," Foucault (2008) offered a glimpse of a different way to understand how the advent of Big Data, which came after him, and of education's often binary responses to outside movements, which came before him, constitute changed circumstances and a new research-practice environment, irreducible to what has gone before. Via a lecture on "American neo-liberalism" at the Collegè de France we might understand education's multiple valences with which we opened this paper and Big Data's morphing uptake as playing on the same field, as but two sides of the same coin, and of not so much technologies of self but a new "technology of the environment" (Foucault, 2008, p. 261).[9]

Foucault (2008) led up to this recognition by noting that liberal governmentality of the 19th and 20th centuries was both legalistic and normalizing, with disciplinary regulation being the switch-point between the two aspects. One of the problems here was the insufficiency of making more and more laws for regulation – it wasn't cost-saving to do so, and the implication was that there was no guarantee that people followed them anyway: "the great idea that the law was the principle of governmental frugality turns out to be inadequate – because '*the* law' does not exist… You [can have?] as many laws as you like, the overflow with regard to the law is part of the legal system" (Foucault, 2008, p. 260).

He goes on to note that because the law can only function ballasted by something else that is its counter-weight, its interstices, two things are necessary: 1) to change the conception of law or at least elucidate its function: "The law is that which favors the game, i.e.,… enterprises, initiatives, changes and by enabling everybody to be a rational subject i.e., to maximize the functions of utility" and 2) to consider calculating its enforcement instead of supplementing it with regulation, planning and discipline: "that is to say we must not supplement it with something else, but with that which must simply give it force but while saying clearly that this enforcement is basically the main element, because the law does not exist without it, because it is elastic, because it can be calculated" (Foucault, 2008, p. 260).

If the new operating principles were to be built around making everybody a rational subject or player *within* the law rather than "a subject" imagined as a site of application *of* the law, maximizing the function of utility and force via elasticity, the question remains as to how to rationalize this enforcement. Key here is action on "the environment" rather than directly on "the subject," whom the past had shown, didn't always obey or respond to feedback in desired or predictable ways. For Foucault

(2008), such a new rationalization is achieved "through the calculation of costs, the utility of the law, and the cost of its enforcement and by the fact that if you do not want to get out of the law and you don't want to divert its true functions as rule of the game, the technology to be employed is not discipline-normalization, but action on the environment. Modifying the terms of the game, not the players' mentality" (p. 260).

While this analysis is a stepping stone to what Foucault (2008) proposes later in *The Use of Pleasure* and *The Hermeneutics of the Subject* – the entwinement of changes in game and mentality - it emphasizes what he called a "radicalization" that moves from Old world to New World, from German liberalism to American neo-liberalism:

> We have here a radicalization of what the German ordoliberals had already defined with regard to governmental action: leave the economic game as free as possible and create a *Gesellschaftspolitik*. The American liberals say: if you want to maintain this *Gesellshaftspolitik* in the order of the law, you must consider everyone as a player and only intervene on an environment in which he is able to play. An environmental technology whose main aspects are: the definition of a framework around the individual which is loose enough for him to be able to play; the possibility for the individual of regulation of the effects of the definition of his own framework; the regulation of environmental effects; non damage; non absorption; the autonomy of the environmental spaces (Foucault, 2008, p. 261).

This points to something beyond "disciplinary societies" with their normalizing examinations, ascetic techniques, teleologies, and definition of life as a life of testing. While perhaps not without those elements, the landscape-changer here is the appearance of action on the environment rather than on the "spirituality" or "soul" of the player: "Not a standardizing, identificatory, hierarchical individualization, but an

environmentalism open to unknowns and transversal phenomena. Lateralism" (Foucault, 2008, p. 261).

The search for causality moves from "subject" to "system," or rather, "environment" in which nonlinearity is still taken as an "illness" or problem to control but expected to a certain extent within the game. This requires an acceptance at some level of an uncertainty *within* the predictive modality of game theory, born of a transversal connectography that puts system-environment in touch with system-environment. In field connected to field, education to Big Data and Big Data to education, the unintended consequences integral to understandings of the nonlinear intersubjectivity and suggestibility that once marked the "mystery" of the social sciences and human-to-human interactions become embodied and rationalized within a probability reasoning in search of new patterns. These are new patterns that, in turn, try to establish new criteria for recognizing and controlling the unintended, forcing decision-making into quantum-like considerations, while rewriting cognition as neuralized discernment to which circulating, transversal, and incoming data environments attend first.

Conclusion: Old Maps, New Maps, No Maps?

The impossible "map" sketched here - of reactions to EDM and LA as part of Big Data's emergence, mutations in technologies of self (*epistrophé*, *metanoia*, *subject-as-object*) to which Big Data and responses to it are indebted, the advent of particular vocabularies (datum, data, and Big Data), and more recent disciplinary and institutional formations dedicated to different kinds of rationality amid changing "social compacts" marked by technologies of the environment - offers one angle on flashpoints between

the seemingly subjective realm of education and the seemingly objective realm of Big Data.

This paper has not intended to solve the problems posed by such flashpoints, differences, and changes in processing power. Instead, it has offered a vantage point on 21st century repositories of self-knowledge-reality, in which the three evolving lines that Foucault identified as recurring and integral (dietetics, economics and love) implicitly reshape and are reshaped by new-ish versions of "the master" through which care of the self (involving transcendence) and technology of the environment (no transcendence/immanence possible) are imagined to take shape. In the current lateralism that education-meets-Big-Data is but one example of, the capacities or abilities attributed to human selfhood are continuously stretched by how connection changes the nature of an object (Cooper Ramo, 2016). Redefining and/or restricting the imagination of improvement, exceeding the augmentative, and moving into new kinds of enfolding between "system" and "being-as-self" helps relocate selfhood beyond the flesh and beyond simple appeals to either "number" or "cultural contexts."

Intersubjectivity is not just mother-infant here and cannot escape the play of the mechanic, which radically transforms the price paid for access to "truth." In educational versions of holism and developmentalism, opposed as they sometimes are, the subject must pay a price for access to truth, confessing his or her "false consciousness" so that the putatively superior academic rescuer, researcher, psychologist, or interlocutor can substitute a new one. In Big Data, the true nature (data) to which the self returns does not go through an a priori confessional practice that involves putting oneself in the right condition (data literacy) to read the graphs or printouts on the screen. Rather, it entails the language of modification, adaptation, amplification, augmentation, and support that

locates Becoming in the technologies of the environment and Being, uncontested and godlike, in the code.

From a curriculum studies perspective, the questions that such an impossible "map" can raise are many and significant: If Big Data is but an opinion embedded in a code why not analyze opinions and prejudices first and/or instead? If coding is an enactment that participates in desire for control, who or what owns infrastructure, the storage, and who gets access to coding? Are new networks "truly" flattened out, so that anything coming from anywhere can impact everywhere, or is it more lopsided than that? What happens to strongly place-based and/or indigenous and/or so-called "non-Western" conceptions of "being" amid the standardization that programming generates, whether that is programming of a belief in the naturalness of self (curriculum) or of computing? Are "humans" to accept more and more levels of representation through numerical optimization? Can non-representational theory engage a more multimodal archive rather than just a screen or print-based one? In short, might the disenchantment and re-enchantment of Western versions of science, spirituality, and rationality, in whatever forms, generate a new fusing of disciplines that offers new maps? Better maps? Or no maps? In short, in its beginnings, does Big Data "truly" get "us" beyond massive social and educational problems (called improvement) or simply trade on old ones and in the process unleash new ones (called reflexivity)?

For Richardson (2009), the reduction of the high school curriculum to apparent debates over scientific versus classical presented an opportunity to demonstrate how "the nation" was constructed as such through disparate and irreducible differences, emanating as regionality. For Khanna (2016), the rosy versus gloomy scenarios of globalization rhetoric remain unhelpful amid the need for new maps that are required to understand the federations, alliances, networks, and metropoles that exceed

geopolitical thinking. The controversies that might emerge from quite different yet related onto-epistemologies coming into contact, education and Big Data, present new opportunities for asking different kinds of questions about what Latour (2013) described not so much as matters of fact, but rather matters of concern, inviting consideration of whether and where the appeals to "the real," "the self," and "knowledge" – if any - will be anchored, and whether that matters as the 21st century continues to unfold.

References

Ascher, M. (1991). *Ethnomathematics: A Multicultural View of Mathematical Ideas*. Belmont, Canada: Wadsworth Inc.

Amrein-Beardsley, A. (2016). U.S. Department of Education: Value-added not good for evaluating schools and principals. [blog post]. Retrieved 12/05/16 from: http://vamboozled.com/?s=Chiang&submit=Search&__bcf_gupi=1DDFF236347B00016DDE19C0BBC064001DDFF236347C0001B232142E15ECE000.

Baker, B. (2001). *In perpetual motion: Theories of power, educational history, and the child* (Vol. 14). New York: Peter Lang.

Baker, B. (2010). Provincializing curriculum? On the preparation of subjectivity for globality. *Curriculum Inquiry, 40*(2), 221-240.

Baker, B. (2016). From technologies of self to discourses of vision: The conditions of possibility for educational neuroscience. Keynote presentation, *Possibliites of "Data" Conference & Summer Institute*, University of Tampere, Finland, Aug 2016.

Baker, S. R., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning Analytics: From Research to Practice* (pp. 61-75). New York: Springer.

Baker, R. & Siemens, G. (2011). Educational data mining and learning analytics. Online paper retrieved from: http://www.columbia.edu/~rsb2162/BakerSiemensHandbook2013.pdf.

boyd, D., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society, 15*(5), 662-679.

Buckingham Shum, S. (2016). Algorithmic Accountability & Learning Analytics (UCL). Retrieved from http://www.slideshare.net/sbs/algorithmic-accountability-learning-analytics-ucl

Buckingham Shum, S., & Deakin Crick, R. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics, 3*(2), 6-21.

Chertok, L., & Stengers, I. (1992). *A critique of psychoanalytic reason: Hypnosis as a scientific problem from Lavoisier to Lacan*. Stanford, CA: Stanford University Press.

Chiang, H., McCullough, M., Lipscombe, S., & Gill, B. (2016). *Can student test scores provide useful measures of school principals' performance?* . Washington, DC: Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Cooper Ramo, J. (2016). *The Seventh Sense: Power, Fortune, and Survival in the Age of Networks.* New York: Hatchette Book Group.

Datafloq (2016). A short history of Big Data. Retrieved 05/12/16 from https://datafloq.com/read/big-data-history/239).

Dear, P. (1995). *Discipline and Experience: The Mathematical Way in the Scientific Revolution*. Chicago, IL: University of Chicago Press.

Derrida, J. (1978). Structure, sign, and play in the discourse of the human sciences (A. Bass, Trans.) *Writing and difference* (pp. 278-294). London: Routledge and Kegan Paul.

Derrida, J. (1984). Deconstruction and the other. In R. Kearney (Ed.), *Dialogues with contemporary continental thinkers - the phenomenological heritage: Paul Ricouer, Emmanuel Levinas, Herbert Marcuse, Stanislas Breton, Jacques Derrida* (pp. 105-126). Manchester: Manchester University Press.

Diebold, F. X. (2012). *A personal perspective on the origin(s) and development of "Big Data": The phenomenon, the term, and the discipline*. Unpublished manuscript, Retrieved from http://www.ssc.upenn.edu/~fdiebold/papers/paper112/Diebold_Big_Data.pdf.

Eppert, C., & Wang, H. (Eds.). (2010). *Cross-Cultural Studies in Curriculum: Eastern Thought, Educational Insights*. New York: Routledge.

Eynon, R. (2013). The rise of Big Data: What does it mean for education, technology, and media research? *Learning, Media and Technology, 38*(3), 237-240.

Foucault, M. (1985). *The Use of Pleasure. The History of Sexuality (Vol. II)* London: Penguin.

Foucault, M. (2008). *The Birth of Biopolitics. Lectures at the Collège de France 1978-1979* (G. Burchell, Trans.). New York: PalgraveMacmillan.

Foucault, M. (2005). *The Hermeneutics of the Subject: Lectures at the Collége de France, 1981-1982* (G. Burchell, Trans. F. Gros Ed.). New York: Palgrave Macmillan.

François, K., & Van Kerkhove, B. (2010). Ethnomathematics and the philosophy of mathematics (education). In B. Lowe & T. Muller (Eds.), *Philosophy of Mathematics: Sociological Aspects and Mathematical Practice* (Vol. 11, pp. 121-154). London: College Publications.

Gartner IT Gossary (2016). Retrieved 09/05/16 from: http://www.gartner.com/it-glossary/big-data/

Graham, M. (2016). Historicizing Big Data and Geo-information [Blog post]. Retrieved 09/05/16 from http://cii.oii.ox.ac.uk/2016/03/30/historicizing-big-data-and-geo-information/

Hacking, I. (1990). *The Taming of Chance*. Cambridge, MA: Cambridge University Press.

Hattam, R., & Baker, B. (2015). Technologies of self and the cultivation of virtues. *Journal of the Philosophy of Education, 49*(2), 255-273.

Kei Daniel, B. (Ed.) (2017). *Big Data and Learning Analytics in Higher Education: Current Theory and Practice*. Switzerland: Springer.

Khanna, P. (2016). *Connectography: Mapping the Future of Global Civilization*. New York: Random House.

Koro-Ljungberg, M., & Maclure, M. (2013). Provocations, re-un-visions, death, and other possibilities of "Data". *Cultural Studies ++ Critical Methodologies, 13*(4), 219-222.

Latour, B. (2013). *An Inquiry into Modes of Existence: An Anthropology of the moderns* Cambridge MA: Harvard University Press.

Lohr, S. (2013). The origins of 'Big Data': An etymological detective story [Blog post]. Retrieved 12/05/16 from http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/?_r=0) [A version of this article appears in print on 02/04/2013, on page B4 of the NewYork edition with the headline: Searching for Origins Of the Term 'Big Data'].

Long, P., & Siemens, G. (2011). Penetrating the fog: Learning analytics and education *EDUCAUSE Review, 32*, 31-39.

Luhmann, N. (1985/1995). *Social systems* (J. John Bednarz, with Dirk Baecker, Trans.). Stanford, CA: Stanford University Press.

Morabito, V. (2015). *Big Data and Analytics: Strategic and Organizational Impacts*. Switzerland: Cham: Springer International Publishing

Morozov, E. (2013). *To Save Everything, Click Here*. London: Penguin Books.

Myer-Schoenberger, V., & Cukier, K. (2014). *Learning with big data: The future of education*. New York, NY: Houghton Mifflin Harcourt

Nakata, M. (2007). *Disciplining the Savage, Savaging the Disciplines*. Canberra: Aboriginal Studies Press

O'Neill, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Osberg, D., Biesta, G., & Cilliers, P. (2008). From representation to emergence: Complexity's challenge to the epistemology of schooling. *Educational Philosophy and Theory, 40*(1), 213-227.

Paige, M. A. (2016a). A new book about VAMs "on trial." [blog post]. Retrieved 12/05/16 from http://vamboozled.com/a-new-book-about-vams-on-trial/

Paige, M. A. (2016b). *Building a Better Teacher: Understanding Value-Added Models in the Law of Teacher Evaluation*. Lanham, MA: Rowman & Littlefield.

Poovey, M. (1998). *A history of the modern fact. Problems of knowledge in the sciences of wealth and society*. Chicago: University of Chicago Press.

Porter, T. (1986). *Rise of statistical thinking, 1820-1900*. Princeton, NJ: Princeton University Press.

Prinsloo, P., Archer, E., Barnes, G., Chetty, Y., & Dion, v. Z. (2015). Big(ger) Data as better data in open distance learning. *The International Review of Research in Open and Distributed Learning, 16*(1).

Richardson, J. (2009). Institutional sequences and curriculum history: Classical versus scientific knowledge and the formation of a new nation. In B. Baker (Ed.), *New Curriculum History* (pp. 141-168). Rotterdam: Sense Publishing.

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, 40*(6), 601-618.

Sells, M. (1994). *Mystical languages of unsaying*. Chicago, IL: University of Chicago Press.

Selwyn, N. (2014). *Distrusting educational technology: Critical questions for changing times*. New York: Routledge.

Siemens, G. (2011). Critiques of learning analytics? [blog post]. Retrieved 12/05/16 from https://scope.bccampus.ca/mod/forum/discuss.php?d=16477

Siemens, G., & Baker, R. (2012). *Learning analytics and educational data mining: Towards communication and collaboration* Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, New York, NY, USA.

Stiegler, B. (1998). *Technics and Time: The Fault of Epimetheus* (Vol. 1). Palo Alto, USA: Stanford University Press.

Stigler, S. M. (1999). *Statistics on the Table: The History of Statistical Concepts and Methods*. Cambridge, MA: Harvard University Press.

Taylor, C. (1992). *Sources of the Self: The Making of the Modern Identity*. Cambridge, MA: Harvard University Press.

Watkins, W. (1994). Black curriculum orientations: A preliminary inquiry. *Harvard Educational Review, 63*(Fall), 321-338.

Notes

[1] I use quotes here to refer to the disagreement around such a construct. Michael Sells' (1994) rendition of the complexity in naming West points to the intersection of Ptolemaic cosmology with Abrahamic traditions around belief in a single transcendental principle of reality.

[2] See Nakata (2007) which takes on all three seeming bastions – idea of religion (over and above the magical and less than the scientific), of nationhood, and of individual.

[3] I am not referring to Heidegger's multiple and well-known critiques of being here but rather to the classical Greek reduction of being to notions of self that will be elaborated in later sections via Foucault. Readers interested in such distinctions can consult Taylor's (1989) seminal text *Sources of the Self*.

[4] I have discussed elsewhere (Baker, 2001) Niklas Luhmann's (1995) conception of systems theory, autopoiesis and sociology, amid his recognition of the arbitrary distinction between system and environment, his dislike of the term "the subject" and of unreflective notions of agency. I deploy system here to refer to interlinked processes that congeal with temporary identifiable boundaries such as systems of thought, self-care, teaching, performance, evaluation etc. This includes the Luhmannian recognition that one "system" may constitute an "environment" for another system, etc.

[5] There is a certain irony and reentry in making reference to dictionaries and etymologies as though compendia of data themselves. I have recently discussed data's etymology alongside the other key terms with which it has become allied. See Baker (2016).

[6] Diebold goes on to add: "The appropriate allocation is open to debate, however, as there are issues of Big Data interpretation and context, and things get murkier if one

includes unpublished and/or non-academic references" (Diebold, 2012, pp. 2–3). http://www.ssc.upenn.edu/~fdiebold/papers/paper112/Diebold_Big_Data.pdf

[7] One of the core problems in delimiting and naming a field such as ethnomathematics, however, are the cultural assumptions, even when shifting or revised, about what is related to mathematics and what not. The lens is set in advance at a second-order level. The logocentrism is not resolved by lists of predefined distilled skills such as comparing, inferring, counting etc and calling them mathematical as opposed to something else.

[8] Amrein-Beardsley notes in regard to context here: "Before the passage of the Every Student Succeeds Act (ESSA), 40 states had written into their state statutes, as incentivized by the federal government, to use growth in student achievement growth for annual principal evaluation purposes. More states had written growth/value-added models (VAMs) for teacher evaluation purposes,… but this [entry] pertains only to school and/or principal evaluation purposes. Now since the passage of ESSA, and the reduction in the federal government's control over state-level policies, states now have much more liberty to more freely decide whether to continue using student achievement growth for either purposes (http://vamboozled.com/u-s-department-of-education-value-added-not-good-for-evaluating-schools-and-principals/). Retrieved Nov 9, 2016. No page numbers available for quoted material.

[9] I thank Peter O'Brien for alerting me to the fecundity of Foucault's scribbles around the biopower lectures.