

UC Irvine

UC Irvine Previously Published Works

Title

Representationalism at work: dashboards and data analytics in urban education

Permalink

<https://escholarship.org/uc/item/9wr2m0rt>

Journal

Educational Media International, 54(4)

ISSN

0952-3987

Author

Crooks, Roderic

Publication Date

2017-10-02

DOI

10.1080/09523987.2017.1408267

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at <https://creativecommons.org/licenses/by-nc-sa/4.0/>

Peer reviewed

Representationalism at Work: Dashboards and Data Analytics in Urban Education

Abstract

This paper explores data analytics applied to urban education, focusing in particular on issues of representationalism, the view that representations (here, digital data) stand in mimetic relation to some external reality from which they are ontologically distinct. Based on interviews conducted over the 2016 – 2017 school year with a team of data professionals employed by a Charter Management Organization (CMO) that operates one dozen schools in South and East Los Angeles, this paper shows that respondents endorse a range of views about how data can stand in for things, events, and people. Data professionals charged with the creation and management of student-level data expressed views consistent with a representationalist understanding of digital data; by contrast, those charged chiefly with aggregating and analyzing heterogeneous streams of data expressed greater skepticism toward representationalist commitments. Despite this difference in viewpoint, all of the data professionals interviewed relied on the same medium to communicate their interpretations of data to other members of the organization, including school personnel: a bespoke platform that consists of 125 dashboards. Critically, graphical means of visualization expressed objectivity, certainty, and actuarial foresight, even in cases where data professionals expressed ambivalence about the representational power of a given source of data.

Introduction

Conflicts abound where the language, techniques, models, and aesthetics of Big Data have been employed to address the needs of public education in general and urban schools in particular, those high-poverty institutions that serve minoritized¹ communities. In urban schools,

¹ As opposed to terms like *underrepresented*, *minority*, or *underserved*, the term *minoritized* draws attention to the historical specificity of American racial and sexual hierarchy and the potential multiplicity of identities of individual persons. *Minoritized* is a way to name social location in a pervasive atmosphere of white supremacy, heteronormativity, and misogyny (Muñoz, 1999).

technologically sophisticated modes of delivering and administering public education have promised to improve the quality of learning in minoritized communities through concepts used in both the tech sector (e.g., innovation and disruption) and in scientific research (e.g., quantification and statistical analysis) (Sims, 2017; Kelly, 2017). This promise, one made by technologists and their adherents to the racially segregated and economically marginalized schools of America's cities, rests in part on the auratic appeal of computation, on its presumed neutrality, objectivity, and, perhaps most of all, its certainty (Golumbia, 2009). For some of its boosters, data promises to remove bias and uncertainty from administrative and pedagogical decisions. From the perspective of those who work with data produced in the routine, everyday activities of urban schools, however, dealing with uncertainty is part of the job. In addition to their quantitative and programming skills, data professionals must also speak on behalf of data and produce authoritative interpretations. As this paper will demonstrate, this interpretive labor must balance incommensurable, conflictual imperatives about how data stand in for real-world phenomena. Given the growing power and prestige of data in public education, one can read the motto of CMO-LA — a federated charter management organization (CMO), the field site of this project — as both hopeful and troubling: “Promoting A Data-Driven Culture.” Conspicuously, this motto is also used by the publisher of Tableau, a data visualization platform that figures prominently in the analysis that follows.

This paper reports early results of a larger project about the application of data analytics to urban education. Despite the presumed certainty of data and associated analytic techniques, I note an important distinction in the work of different kinds of data professionals employed at

CMO-LA. Professionals tasked with analyzing data (i.e., the Analytics Team) tended to view the relationship between data and the things data stand in for more flexibly than those workers tasked primarily with capturing data (i.e., SI, the Student Information Team). Despite this marked contrast in approach, all data workers interviewed relied on the same tools to communicate the results of their work back to the rest of the organization, including to teachers, executives, and other data professionals: graphical means of presentation, primarily dashboards. A dashboard is a visual presentation of statistical data, where “proximity, grouping, orientation, apparent movement, and other graphical effects” produce semantic meaning (Drucker, 2014, p. 131). In this graphical mode of interpretation, visualization necessarily obscured ambiguities or contradictions inherent to the various contexts in which educational technology – including platforms, tools, apps, tests, software, and user behavior —generated digital data. In effect, the tools that made data actionable did so by producing an image of unified and orderly measurement out of messy, irregular data of differing quality, fidelity, and granularity. These variations are not merely questions of cleaning or fixing particular datasets: to data professionals, they speak to the fundamental “messiness” of data.

This paper contributes empirical specificity to the study of the emergent forms of managerial control in education via data-intensive techniques, particularly the creation of visual analytics to direct organizational behavior. It also explores the rhetorical and economic power of large-scale, commercial data manipulation and visualization technologies in the context of the partially privatized public service sector and in the specific case of urban education in South and East Los Angeles. Data-intensive management is ascendant in many domains of public life, including government administration, higher education, law enforcement, and public health; its uptake in urban education might prove illustrative of its movement elsewhere. The next section

has two goals: to trace the representationalist commitments of data-intensive modes of public education and to link urban schools to expectations about the putative educational and ameliorative potentials of data analytics.

Big Data, abstraction, and interpretation

Digital data represent things in the world via abstraction and thereby subject various complex social processes (including public education) to regimes of computation (Evens, 2012). At the same time, digital data are never “raw” and cannot speak for themselves: they may only produce meaning by passing through some interpretive faculty (Gitelman, 2013; Kitchin 2014). These two complementary aspects of representation — interpretation and abstraction — run through much of the scholarly work that seeks to characterize the technological, cultural, and economic phenomenon called Big Data (boyd & Crawford, 2012; Bowker, 2014). Big Data is often defined in terms of increased velocity, volume, and variety: data have become superabundant, produced continuously by new sensors, devices, and platforms (Kitchin, 2014; Matzner, 2016). Big Data encompasses a set of machine learning techniques developed in computer science, including textual mining, logistic regression, support vector machines, and anomaly detection, among others (Burrell, 2016). The power of data to stand in for things — the capacity of representation — is vital to the success of Big Data (van Dijck, 2014).

Critiques in critical data studies have pursued several consistent lines in the past half-decade, responses that tend to undercut characterizations of Big Data as epochal. These critiques point to the persistent centrality of human interpretation in data work; the role of context in understanding data; the significance of materiality in technosocial systems; and the embeddedness of values in technical artifacts (Neff et al., 2017; Bowker, 2014; Boellstorff, 2013). These critiques attack abstraction in the representational sense, pointing to conflicts in the

relationship between data and the things for which they serve as proxy. Strong versions of this argument have directly attacked the whole notion that data record preexisting attributes, arguing instead that digital data are constructed and given meaning in the course of their creation and do not merely record inherent qualities of things (Kirschenbaum, 2008; Drucker, 2014). Likewise, feminist technoscience has criticized the representationalist paradigm, the belief “in the ontological distinction between representations and that which they purport to represent” (Barad, 2007, p. 46). As an alternative, theories of performativity hold that instead of observer-independent objects out in the world waiting to be represented, phenomena are actively shaped by the capacities of observation and measurement that capture them (ibid., p. 109). Perrota and Williamson (2016) employ theories of performativity to explain the emergence “a new form of unquestioned educational consensus around educational data science,” one that serves to occlude critical civic concerns (p. 2).

The consultation of data in specific professional, institutional, or scientific realms is the work of analytics, those transformations that aim to shape data into description, explanation, prediction, and prescription (Minelli, Chambers, & Dhiraj, 2013). Analytics then is the interpretation of data by humans and algorithmic routines and consists of various types of manipulation: data mining and pattern recognition; data visualization and visual analytics; statistical analysis; and prediction, simulation, and optimization (Kitchin, 2014, p. 101). It is important to note that each set of techniques is based on particular assumptions about what data are, how they can be used, and to what particular reality they refer. The production of analytics entails a “professional vision” that allows would-be analysts to “straddle the competing demands of formal abstraction and empirical contingency,” that is to say, between abstraction and interpretation (Passi & Jackson, 2017, p. 2436). The software, platforms, and architectures used

to make data tractable to human workers and algorithmic agents likewise conceal various kinds of subjective judgements (Kitchin, 2014, p. 112). For example, the aggregation and manipulation of various forms of education-related data rely on the “considered judgement” of data professionals (Barocas et al., 2017).

As these techniques have been implemented in education, data analytics concerns the analysis of various kinds of data captured by networked learning environments; the ongoing consultation of data for rapid feedback and selection of instructional materials (learning analytics); and the incorporation of metrics derived from data for use in learning, administration, and design (education analytics) (Cope & Kalantzis, 2016; Ferguson, 2012). Applying data analytics to the realm of education requires several leaps of abstraction: individual students undergo a process of “datafication,” wherein a certain reductive conceptualization allows authorities to “assume a self-evident relationship between data and people, subsequently interpreting aggregated data to predict individual behavior” (van Dijck, 2014, p. 1999).

Scholarship on analytics applied to education has generally focused on taxonomies of analytics and works or ethical implications for student privacy (Rubel & Jones, 2016; Slade & Prinsloo, 2013, Perrotta & Williamson, 2016). Personalized learning, individualization of instruction achieved through rapid analysis of data, has long been a goal of educational technologists. Although learning analytics have been applied to the accomplishment of this vision, they are costly and unproven (Bulger, 2016; Zeide, 2017). Many scholars and advocates have warned of the potential harm of discrimination based on demographic data captured through education or learning activities (Barocas, 2014; Harcourt, 2007)

Finally, the application of analytics is of particular interest in the field of urban education. In the United States, urban education refers to those high-poverty, public schools that serve

minoritized communities, primarily black and Latino, primarily in central cities previously hollowed out by white flight and now the site of intense housing displacement (Costa Vargas, 2006; Means, 2012; Pearman & Swain, 2017). Schools that serve minoritized communities face pressure to make use of contemporary technologies to address persistent conditions of inequality: in the past two decades, these technological fixes have included one-to-one computer programs, hybrid online/in-person instructional designs (i.e., “blended” learning), and games-based curricula, among many other efforts (Crooks, 2015; Christensen, et al., 2008; Sims, 2017). The current wave of Big Data-inspired innovation comes at a time when the basic principles of American schooling have become unsettled, as evidenced by a policy consensus that blames under-resourced schools themselves for failing to produce equity in low-income, urban, or minoritized communities (Au, 2016; Mirón and St. John, 2003). Urban schools, produced through residential segregation and sustained economic disinvestment, are sites of extreme inequality and precarity (Soyjourner, 2012). Urban schools are highly disciplined environments, where the collection and aggregation of data support restrictive regimes of accountability and surveillance (Taylor, 2015). Regimes of data aggregation and manipulation — produced and expanded in each successive wave of technological intervention in public education — “make schools legible for the market, mark specific schools and school districts as pathological and in need of authoritarian governance, and justify minimalist schools in areas of urban disposability” (Lipman 2013). From this perspective, technologized modes of administering and delivering public education achieve a troubling duality: they stand as both a proposed solution for persistent problems of economic inequality in minoritized communities and one of the means by which such inequality is reproduced.

Method

This paper uses ethnography to explore how data professionals working in the context of urban education understand data. An ethnographic text is a “thick description” of a cultural context, one shaped by the researcher’s situated understanding of how participants themselves understand the world (Geertz, 1973). Applied to professional settings, ethnographic methods capture elements of how organizational structure affects attitudes and behavior, how particular people understand the broader organization, profession, and society in which they work (Hodson, 2004). As part of a larger ethnographic project, this paper makes no statistical claims about the prevalence of the attitudes described here outside the single site studied, nor does it contribute directly to outcomes-based research. This is an exploratory project meant to open up new questions in the space of data analytics applied to the domain of public education.

The primary sources of data for this paper are questionnaires and follow-up interviews completed over the 2016 – 2017 school year by a team of data professionals employed by CMO-LA². CMO-LA is a regional not-for-profit charter management organization that operates under the umbrella of a national foundation: the regional organization and the national foundation share resources and data. CMO-LA operates 13 middle and primary schools in South and East LA, representing some 10,000 students. These urban schools serve minoritized, lower socioeconomic status communities: 99% of students identify as black and/or Latino and 91% of students qualify for free lunch. CMO-LA’s schools, like many charter schools in Southern California, are intensely racially segregated (Orfield & Ee, 2014).

CMO-LA’s official mission — to “teach the academic skills, foster the intellectual habits, and cultivate the character traits needed for our students to thrive in high school, college, and life” — concerns a specific strategy of charter-based education reform. CMO-LA seeks to

² Aliases are used extensively in this text.

achieve the growth of its network by demonstrating superior educational outcomes as compared to traditional public schools in Southern California. CMO-LA then promotes itself via publications, videos, public presentations, and appeals to various government agencies and philanthropies. Arguments supported by various forms of data are central to the mission of the network and its demonstration of its own capabilities. The current research focuses on the work of the Data Team, a staff of nine. The following chart shows the structure of the team. The numbers in parentheses indicates the length of the tenure of the person holding that role at the time of writing.

Director of Data & Analytics (7 years)	
<i>Analytics Team</i>	<i>Student Information Team</i>
Analytics Manager (4 years)	Senior Student Information Manager (3 years)
Data Scientist (2 years)	Student Data Systems Analyst (4 years)
Data Analyst (6 months)	Data and Student Information Associate (3 years)
	Human Resources Systems Analyst (<1 year)
	Human Resources Associate (<1 year) (did not participate in research)

Table 1. Organization structure of CMO-LA’s Data Team.

The Director of Data and Analytics, the most senior staff person in terms of both tenure and authority, supervises the work of the two sub-teams: the Analytics Team and the SI (Student Information) Team. The larger team, SI, consists of a Senior Student Information Manager and her three subordinates. This team primarily works in customizing a commercial platform called Illuminate that holds data about schools, students, families, staff, and educational activity. The Analytics Team consists of one Analytics Manager and her two subordinates: a Data Scientist and a Data Analyst. These workers have greater programming skills and work in the aggregation and analysis of data that come to them from all over the organization. The Analytics Team has

two related duties: to engineer “pipelines” for data via data engineering and to produce actionable insights from these aggregated data, primarily via the creation of dashboards.

I first issued workers a written questionnaire intended to elicit general responses with regard to their attitudes toward their own work and the mission of CMO-LA. The initial questionnaire asked each member for personal information, including a resume a job description, and work samples. Next, I interviewed each member of the team to expand upon the initial written responses. These interviews lasted between 20 and 90 minutes.

Interviews were recorded and transcribed. Data were coded in a manner consistent with grounded theory approaches: initial coding was produced by attaching short descriptions to interview data; subsequent rounds of focused coding sought to develop the theme of representationalism developed in this iterative approach. This approach to coding embraces formal and theoretical flexibility, a “way to make discoveries and gain a deeper understanding of the empirical world” through analyzing and re-analyzing codes and data together (Charmaz, 2014, p. 137). Finally, I worked with the team members to produce a diagram that marked important sources of data and important audiences for the communication of analytics (Table 2, included below); this diagram, and its ongoing editing by the team, served as an elicitive tool, one that provided a useful frame for reflection and captured important details of work (Theron et al., 2011). In addition to interviews and questionnaires, I attended team meetings and reviewed interfaces, systems, and documents (e.g., conference presentations, official reports, and internal communications).

Views of representationalism: “All data is problematic”

Respondents described a range of views on issues related to representationalism, the adherence of the data they worked with to some external reality. As scholars have frequently pointed out,

data-intensive modes of decision-making promise objectivity by transforming phenomena into data via processes of representation (Neff et al., 2017; Drucker, 2014). Some respondents' explanations of their own work with data draw into question any thoroughgoing commitment to this representationalist paradigm. It is my goal in this short work neither to exhaustively classify the opinions of the Data Team nor to definitely refute representationalist views. Instead, I argue that analytics work at CMO-LA accommodates a variety of conflicting views about what data are and what they can be and this has ramifications for the application of technology to urban schooling. For data professionals on the Analytics Team, questioning the representational power of data is part of the job.

Notably, all team members expressed unanimous support for the goals of urban education as practiced by the CMO. They viewed their careers as contributions to the improvement of the minoritized communities the CMO serves: two of the interviewed team members spoke of their work with data as intimately related to the experience of growing up in the same communities served by the CMO. In this way, the politics of the charter school movement forms a background against which team members understand their own efforts. Importantly, they view their work with data as potentially ameliorative and community-focused. Said the Student Data Systems Analyst,

The way I really think about it is, if I can make data easy to understand ...then we really are doing what CMO-LA says: we serve underserved communities, low income, socially disadvantaged communities, and so that's the goal.

In scientific collaboration, organizational structure has been shown to influence the ways that data are produced and the way that they circulate (Vertesi & Dourish, 2011); a similar dynamic is particularly pronounced at CMO-LA, even though, in many cases workers use similar data, tools, platforms and software (e.g., Python, SQL, Excel). Although both the Analytics and SI teams work together closely, their respective approaches to data contrast greatly.

On the SI team, data were generally taken as unproblematic proxy for people, places, or events. If team members were concerned about the limits of what data could represent, they expressed this as concern with the extensiveness or quality of data, not as a limit of its representational power. Here, the SI Manger talked about her vision of what data could accomplish:

My goal in my work is to have the flow of student information happening very naturally, and for all of that to be like a well-oiled machine, so that as many people as possible can focus on kids. I just want getting the information to be part of the everyday process where it's just super clean and smooth.

Here, explicitly, the flow of data about students (“student information”) is taken as a kind of double for “kids” themselves. Data are distinct from the things they represent, but they stand in mimetic relation to their referents. What’s good for data is also good for the students so represented. In further comments, the SI manger elaborated that good data were “clean” and bad data were “messy.” Messiness indicated errors made by school staff, a lack of dutifulness in recordkeeping. Any deficiency in representational capacity was error. While team members frequently questioned specific qualities of data they worked with, its timeliness or its accuracy, they did not dispute its representational status. This resulted in a sort of transference, where working on data meant working on students, which also meant working to improve communities. Here, the SI Analyst made this series of leaps explicit:

I think of it as different layers. The way I conceptualize it is of course you have the data, and the data... gets pushed to the teacher, to the student, to the parent and then that impacts the community, and that's how I see it and I think that's what keeps me at CMO-LA.

On the other sub-team, the three workers whose primary duties included aggregating data from multiple sources, respondents expressed a greater variety of approaches to issues of representationalism.

Both the Data Scientist and the Data Analyst held other positions previously at CMO-LA and were promoted for gaining new technical skills. In terms of experience, all three members of Analytics learned vital programming skills on the job, often on the SI Team. As the Analytics Manager explained, “You have to be able to do the work, but we value more being passionate about education.” In this sense, programming skills were less important than an understanding of CMO-LA’s mission. If the work of data analytics involves a particular approach to problem solving, a way of balancing abstraction and interpretation, these interviews suggest that, at least at CMO-LA, such “professional vision” would also have to be complimented with considerable specific domain knowledge (Passi & Jackson, 2017; Schutt & O’Neil, 2013).

Here, the Data Scientist summed up the team’s approach:

A lot of what I do is trying something out, seeing if it ‘looks right,’ double checking it, modifying something small, then rerunning it to try it again until I know that it's correct. Once it's in Tableau, it can be harder to detect issues because there's often a lot more aggregation or calculation going on that can obscure a faulty data point, so you have to be extra vigilant about what you're doing.

In further interviews, all members of the Analytics Team expressed similar concerns about making sure their outputs were “reasonable,” or that results were “in scope.” As the Analytics Manger put it, “All data is problematic.” This playful comment referred to the fundamental ambiguity of data work, the way that manipulating stores of data could yield results that no longer referred to anything out in the world. The Analytics Team frequently required a reflexive component, a way of working “until I know that it’s correct.” A certain amount of “messiness” was assumed here, a presumption of the occasional “faulty data point.” This comment also indicates a shift to a different interpretive strategy: the creation of graphical analytics in the commercial software platform Tableau. As this respondent made clear, Tableau masks uncertainty, presenting as settled what might be an unresolved question. In the next section, I look at this masking of uncertainty achieved through use of Tableau.

Visualizing Ambiguity

The two sub-teams differed in their overall approach to issues of how data could represent things: on the SI team, where work consisted mostly of supplying and verifying data that described students and schools, data were understood to stand in for things in a long chain of representational relationships. As it has been shown in studies of experimental science, knowledge-making practices can bridge theoretical conflicts: research methods and tools do not require agreement over what the world is like (Hacking, 1982). On the Analytics Team, working with data was understood to include a process of reflexivity, of questioning if data formed a coherent description of the world. In many cases, this work required repeating or refining work until the results resembled some predetermined range or scope. In every case, team members on both sub-teams reported their work to stakeholders in the same way, via a bespoke platform that consists of 125 dashboards produced by the commercial visualization platform Tableau. Critically, because the visual mediation of information imbues a particular communication with the hallmarks of the discipline where that mode was created, dashboards produced by commercial visualization software express objectivity, certainty, and actuarial foresight (Drucker, 2014).

The chart below, produced in cooperation with the Analytics Team, shows some of the important data sources managed by the sub-team. This chart does not represent all flows of data: in some sense, Analytics members treated such tidy charts as aspirational. At left, various measures of student learning activity are collected in the Illuminate platform: these include assessments administered by outside organizations, including Northwest Evaluation Association and Smarter Balanced Assessment Consortium. Next, various sources of data are consolidated with Illuminate data, including data produced by a variety of commercial platforms and vendors.

The diagram indicates those nodes that are themselves constituted by multiple sources of data, such as commercial blended learning platforms or Google Sheets. All of these various sources of data are ‘warehoused’ using SQL server. From there, the Analytics team makes visualizations using Tableau, which includes its own server architecture. At right, all of these various transformations result in a set of dashboards, 125 in total, each prepared with one of four specific audiences in mind: CMO staff, school leaders (principals), teachers, or school administrative staff.

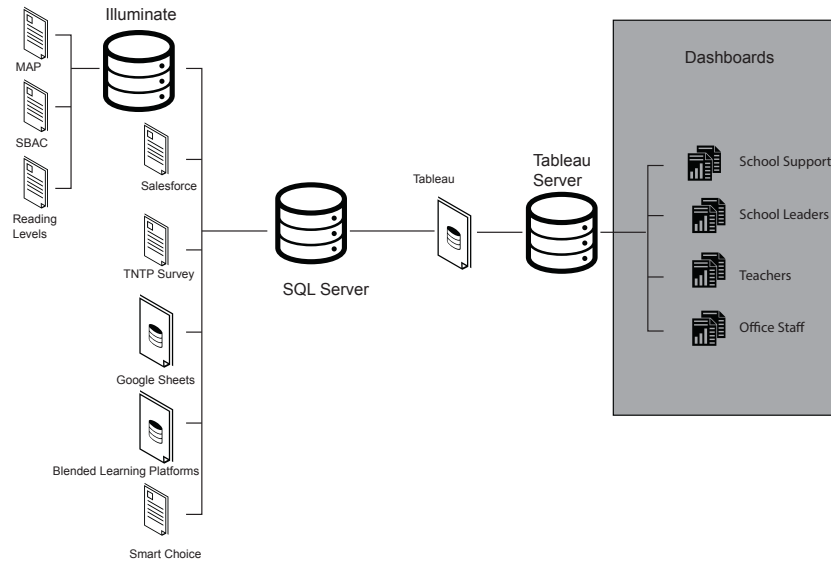


Table 2. Data sources and dashboards.

At its core, the use of data as explained by data professionals at CMO-LA aims for a cybernetic model: various nodes in a network provide data that will be acted upon by agents who can view the action of subservient entities (whatever they are) via feedback (Medina, 2011, p. 37). This vision of control and feedback relies on a fundamentally transitive understanding of semantic meaning. Such an understanding stands in stark contrast to works in science studies and informatics that have shown how the material presentation of information (e.g., the tabular

organization of data in a spreadsheet) carries profound consequence for organizational life (Dourish, 2017). As it applies to the analysis of data in the specific case of urban education in Los Angeles, data visualization tended to obscure important ambiguities in the interpretation of data, presenting as objective, certain, and descriptive what was contextual, provisional, and interpretive.

Educational bureaucracies, particularly those that promote data-intensive modes of organizing public education, manage the fundamental and abiding ambiguity of digital data by producing mediated visual representations of various parameterized phenomena. Data Team members produced dashboards that presented data as trustworthy and definitive, not because they had necessarily made such a determination, but because the visual organization of information carried those associations. As Drucker (2014) argues, it is precisely an over-familiarity with these modes of representation that make them so persuasive. As a corrective, she calls for a theory of humanistic visual epistemology (graphesis) a way of valorizing forms of knowledge and interpretation that are separate from the logo-centric and statistical worlds of science and business. For the present analysis, it is sufficient to point out that the use of dashboards opens up the possibility of bias in what pretends to be an objective, mathematical description. Just as physical spaces support some kinds of behavior and deter others, visual displays and interfaces support or impede certain kinds of thinking (Few, 2006; Schüll, 2014; Farman, 2012).

Conclusion

While there are many potential stances and theoretical commitments that might be imputed from the responses of data professionals, the present work has shown that a variety of nuanced approaches co-exist in this domain. Addressing uncertainty or ambiguity in data is part of the

work of data analytics. In this way, data-intensive modes of organization and management shift interpretive work to certain privileged sites: to the structure of data, to the judgements of data professionals, and to the informational-visual register. Uncertainty, bias, and ambiguity are treated as aspects of metrology, rather than as aspects of human life. Recognizing this, the ambition to rely purely on data analysis to adjudicate decisions about the distribution of public goods and services seems misguided and potentially harmful.

The larger project of which this initial report is a part seeks to better understand the epistemological, techno-social apparatus of Big Data as it spreads to urban schooling. In the setting of the urban schools of South and East Los Angeles, this mapping of computational techniques on to educational institutions makes manifest many troubling assumptions about life in minoritized communities. While data professionals understood themselves to be contributing to the communities served by CMO-LA's schools, many questions about education and the politics of education reform went unasked, particularly as it concerns residential racial segregation that produces urban schools in the first place. This work would suggest that data professionals, committed as they are to service to these communities, could be powerful allies in attempts to address these concerns.

Finally, much has been made of the revolutionary potential of Big Data to reshape knowledge, education, and research. To much fanfare, it is claimed that data analytics obviate theory, because statistical correlations do not require it. Based on interviews with data professionals, it might be more accurate to describe data-intensive inquiry not as *atheoretical*, but as *polytheoretical*, as incorporating multiple ways of thinking about data and how they stand in for things. Research in this space will have to open up the black box of analytics to show how

these conflicts are mitigated and resolved (or not) in data work and how these new kinds of professional identities participate in the ongoing, always unresolved politics of public education.

References

- Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy*, 30(1), 39–62.
<https://doi.org/10.1177/0895904815614916>
- Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Durham: Duke University Press Books.
- Barocas, S. (2014). Data mining and the discourse on discrimination. In *Proceedings of the Data Ethics Workshop, Conference on Knowledge Discovery and Data Mining (KDD)*.
Retrieved from
<https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf>
- Boellstorff, T. (2013). Making big data, in theory. *First Monday*, 18(10). Retrieved from
<http://firstmonday.org/ojs/index.php/fm/article/view/4869>
- Bowker, G. C. (2014). Big Data, big questions: The theory/data thing. *International Journal of Communication*, 8(0), 5.
- boyd, danah, & Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Bulger, M. (2016, July 26). Personalized learning: The conversations we're not having. *Data & Society*. Retrieved from <https://datasociety.net/output/personalized-learning-the-conversations-were-not-having/>

- Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 2053951715622512.
<https://doi.org/10.1177/2053951715622512>
- Charmaz, K. (2014). *Constructing grounded theory*. Thousand Oaks, CA: SAGE Christensen,
- C., Johnson, C. W., & Horn, M. B. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. New York: McGraw-Hill.
- Cope, B., & Kalantzis, M. (2016). Big Data comes to school: Implications for learning, assessment, and research. *AERA Open*, 2(2). <https://doi.org/10.1177/2332858416641907>
- Costa Vargas, J. (2006). *Catching hell in the city of angels: Life and the meaning of blackness in South Central Los Angeles*. Minneapolis: University of Minnesota Press.
- Crooks, R. N. (2015). The coded schoolhouse: Tablet computers and public education. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–1.
<https://doi.org/10.1002/pra2.2015.145052010061>
- Dourish, P. (2017). *The Stuff of Bits: An Essay on the Materialities of Information*. Cambridge, MA: MIT Press.
- Drucker, J. (2014). *Graphesis: Visual forms of knowledge production*. Cambridge, MA: Harvard University Press.
- Evens, A. (2012). Web 2.0 and the ontology of the digital. *Digital Humanities Quarterly*, 6(2). Retrieved from <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>
- Farman, J. (2012). *Mobile interface theory: Embodied space and locative media*. New York: Routledge.

- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304.
<https://doi.org/10.1504/IJTEL.2012.051816>
- Few, S. (2006). *Information dashboard design: The effective visual communication of data*. Sebastopol, CA: O'Reilly.
- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures* (pp. 3–30). New York NY: Basic Books.
- Gitelman, L. (Ed.). (2013). *“Raw Data” Is an Oxymoron*. Cambridge, MA: MIT Press.
- Golumbia, D. (2009). *The cultural logic of computation*. Cambridge, MA: Harvard University Press.
- Hacking, I. (1982). Experimentation and scientific realism. *Philosophical Topics*, 13(1), 71–87.
- Harcourt, B. E. (2007). *Against prediction: profiling, policing, and punishing in an actuarial age*. Chicago: University of Chicago Press.
- Hodson, R. (2004). A Meta-analysis of workplace ethnographies: Race, gender, and employee attitudes and behaviors. *Journal of Contemporary Ethnography*, 33(1), 4–38.
<https://doi.org/10.1177/0891241603259808>
- Kelly, A. E. (2017). Is learning data in the right shape? *Journal of Learning Analytics*, 4(2), 154–159. <https://doi.org/10.18608/jla.2017.42.12>
- Kirschenbaum, M. (2008). *Mechanisms: New media and the forensic imagination*. Cambridge, MA: MIT Press.
- Kitchin, R. (2014). *The data revolution: Big Data, Open Data, data infrastructures and their consequences*. Thousand Oaks, CA: SAGE.

- Lipman, P. (2011). *The new political economy of urban education: Neoliberalism, race, and the right to the city*. New York: Routledge.
- Matzner, T. (2016). Beyond data as representation: The performativity of Big Data in surveillance. *Surveillance & Society*, 14(2), 197–210.
- Means, A. (2013). *Schooling in the age of austerity: Urban education and the struggle for democratic life*. New York: Springer.
- Medina, E. (2011). *Cybernetic revolutionaries: technology and politics in Allende's Chile*. Cambridge, MA: MIT Press.
- Minelli, M., Chambers, M., & Dhiraj, A. (2013). *Big Data, big analytics: Emerging business intelligence and analytic trends for today's businesses*. Hoboken, NJ: John Wiley.
Retrieved from <http://onlinelibrary.wiley.com/book/10.1002/9781118562260>
- Mirón, L., & St. John, E. (Eds.). (2003). *Reinterpreting urban school reform: Have urban schools failed, or has the reform movement failed urban schools?* Albany, NY: State University of New York Press.
- Muñoz, J. E. (1999). *Disidentifications: Queers of color and the performance of politics*. Minneapolis: University of Minnesota Press.
- Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and contribute: A practice-based framework for improving critical data studies and data science. *Big Data*, 5(2), 85–97. <https://doi.org/10.1089/big.2016.0050>
- Orfield, G., & Ee, J. (2014). *Segregating California's future: Inequality and its alternative 60 years after Brown v. Board of Education*. The Civil Rights Project at UCLA. Retrieved from <https://www.civilrightsproject.ucla.edu/research/k-12-education/integration-and->

diversity/segregating-california2019s-future-inequality-and-its-alternative-60-years-after-brown-v.-board-of-education/?searchterm=california%20segregation

Passi, S., & Jackson, S. (2017). Data vision: Learning to see through algorithmic abstraction (pp. 2436–2447). ACM Press. <https://doi.org/10.1145/2998181.2998331>

Pearman, F. A., & Swain, W. A. (2017). School choice, gentrification, and the variable significance of racial stratification in urban neighborhoods. *Sociology of Education*, 38040717710494. <https://doi.org/10.1177/0038040717710494>

Perrotta, C., & Williamson, B. (2016). The social life of Learning Analytics: cluster analysis and the “performance” of algorithmic education. *Learning, Media and Technology*, 1–14. <https://doi.org/10.1080/17439884.2016.1182927>

Rubel, A., & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, 32(2), 143–159. <https://doi.org/10.1080/01972243.2016.1130502>

Schüll, N. D. (2014). *Addiction by design: Machine gambling in Las Vegas*. Princeton, NJ: Princeton University Press.

Schutt, R., & O’Neil, C. (2013). *Doing data science: Straight talk from the frontline*. Sebastopol, CA: O’Reilly.

Sims, C. (2017). *Disruptive fixation: school reform and the pitfalls of techno-idealism*. Princeton, NJ: Princeton University Press.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510–1529. <https://doi.org/10.1177/0002764213479366>

Sojoyner, D. (2012). Criminals, planters and corporate capitalists: The case of public education in Los Angeles. In I. Banks, G. Johnson, G. Lipsitz, U. Taylor, D. Widener, & C. Woods

- (Eds.), *Black California dreamin': The crises of California's African-American communities* (pp. 107–126). Center for Black Studies Research. Retrieved from <http://escholarship.org/uc/item/63g6128j>
- Taylor, E. (2013). *Surveillance schools: Security, discipline and control in contemporary education*. New York: Springer.
- Theron, L., Mitchell, C., & Smith, A. (Eds.). (2011). *Picturing research: Drawing as visual methodology*. Rotterdam: Sense Publishers.
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208.
- Vertesi, J., & Dourish, P. (2011). The value of data: Considering the context of production in data economies (p. 533). *ACM Press*. <https://doi.org/10.1145/1958824.1958906>
- Zeide, E. (2017). The structural consequences of Big Data-driven education. *Big Data*, 5(2), 164–172. <https://doi.org/10.1089/big.2016.0061>