

ORIGINAL ARTICLE

The Structural Consequences of Big Data-Driven Education

Elana Zeide^{1–3,*}

Abstract

Educators and commenters who evaluate big data-driven learning environments focus on specific questions: whether automated education platforms improve learning outcomes, invade student privacy, and promote equality. This article puts aside separate unresolved—and perhaps unresolvable—issues regarding the concrete effects of specific technologies. It instead examines how big data-driven tools alter the structure of schools' pedagogical decision-making, and, in doing so, change fundamental aspects of America's education enterprise. Technological mediation and data-driven decision-making have a particularly significant impact in learning environments because the education process primarily consists of dynamic information exchange. In this overview, I highlight three significant structural shifts that accompany school reliance on data-driven instructional platforms that perform core school functions: teaching, assessment, and credentialing. First, virtual learning environments create information technology infrastructures featuring constant data collection, continuous algorithmic assessment, and possibly infinite record retention. This undermines the traditional intellectual privacy and safety of classrooms. Second, these systems displace pedagogical decision-making from educators serving public interests to private, often for-profit, technology providers. They constrain teachers' academic autonomy, obscure student evaluation, and reduce parents' and students' ability to participate or challenge education decision-making. Third, big data-driven tools define what "counts" as education by mapping the concepts, creating the content, determining the metrics, and setting desired learning outcomes of instruction. These shifts cede important decision-making to private entities without public scrutiny or pedagogical examination. In contrast to the public and heated debates that accompany textbook choices, schools often adopt education technologies ad hoc. Given education's crucial impact on individual and collective success, educators and policymakers must consider the implications of data-driven education proactively and explicitly.

Keywords: big data; personalized learning; competency-based education; smart tutors; learning analytics; MOOCs

Introduction

We are entering a brave new world of data-driven education. Teachers, administrators, and policymakers increasingly rely upon automated technologies for pedagogical decision-making.^{1–3} I use "data-driven" to describe systems that operate without human intervention. In this article, I focus on automated instructional technologies that perform core educational functions: delivering instructional material, assessing student progress, and documenting attainment.⁴ Big data analytics diagnose and predict student progress to inform both instructional and institutional choices. Smart learning environments create new ways to communicate academic attainment. This article considers big data-driven

education in public schools, although many of the points mentioned hereunder also apply to publicly-funded private education institutions.

Big Data-Driven Education

As background, I describe different ways schools use data.⁵ In physical schools, teachers and administrators use student information to inform teaching decisions, give grades, award credits, and create transcripts. Today, faster internet speeds and cloud computing supports similarly interactive virtual learning environments.⁶ Learners select among modular videos, practice problem sets, and explore supplemental reference material at their convenience. Schools also incorporate these platforms into

¹Center for Information Technology Policy, Princeton University, Princeton, New Jersey.

²Information Society Project, Yale Law School, New Haven, Connecticut.

³Data & Society Research Institute, New York, New York.

*Address correspondence to: Elana Zeide, Center for Information Technology Policy, 317 Sherrend Hall, Princeton University, Princeton, NJ 08544, E-mail: elana@nyu.edu

their curricula to provide “blended” learning experiences that contain both online and physical components or “flip” instruction, so that students watch lectures as homework and discuss their content in classroom.⁷

Massive Open Online Courses (MOOCs) are perhaps the most-hyped virtual learning environments. At first, MOOCs offered content from elite universities for free to anyone with an internet connection. They promoted their platforms as a means to democratize high quality education, prompting the New York Times to declare 2012 “The Year of MOOC.”⁸ Instructional education technologies have evolved since then. The most prominent MOOC providers, Coursera, edX, and Udacity, have all narrowed their focus to vocational and professional subjects and charge users for participation or certification.⁹ New providers continuously create innovative education experiences through adaptive e-textbooks, mobile apps, and education games.^{10,11,12}

Technology-mediated instruction

Technologically-mediated education technologies generate a continuous stream of information as learners interact with digital platforms.^{13–16} Students and teachers input information such as usernames, emails, and grade level to set up student accounts. Learners also provide typical academic information through emails, online discussions, assignments, and tests. Digital education technologies collect an unprecedented amount of information about students’ behavior and performance during the learning process—details that could not be recorded or analyzed at scale in physical classrooms. This includes metadata such as time stamps, device identifiers, and even geolocation information.^{13,17} These systems track the parts of a video students actually watch, when they log in, and how they paused before answering a question.¹⁸

The Internet of Things will expand the possible sources for student data exponentially.^{19,20} Online proctoring platforms use video, facial recognition, audio, and biometric information to verify student identity and detect cheating.²¹ “Smart” college campuses collect information from radio-frequency identification (RFID) cards that record when students go to the gym and what they bought for lunch.²² Oral Roberts University, for example, requires students to wear Fitbits.²³ The most ambitious education reformers and researchers envision a future wherein sensors track students’ eye or breathing patterns to determine their level of engagement or anxiety.^{24,25}

Data-informed decisions

Big data-driven education technologies incorporate and analyze this wealth of information to inform classroom and institutional decision-making.^{26,27} Learning analytics can provide a more precise diagnosis than harried human instructors.^{28,29} They track student progress using knowledge maps that break down the relevant subject matter into concepts and “competencies.”³⁰ Platforms, for example, can determine that a student’s poor chemistry grade is the result of failure to grasp a specific algebraic concept the prior year, rather than any difficulty with the scientific concepts themselves.³¹ These cognitive models may include or infer emotional and cognitive states as well as academic progress.^{32,33}

Most of today’s school technologies interpret and present this information to educators on digital dashboards. The level of interpretation and inference involved in these systems varies widely. Some dashboards show “skill meters” that visually graph learners’ mastery of specific concepts.^{34,35} Others reduce a complex array of information to sort learners into simple categories.^{36,37} One early warning tool, for example, uses red, yellow, or green indicators to show teachers students’ likelihood of passing.

Educators can incorporate data-generated student assessments and predictions to support their independent decision making. As Ryan Baker memorably writes, “stupid” tutoring systems can be crafted to inform, rather than replace, “intelligent” human decision-making.³⁸ Many platforms currently in use are primarily oriented at detecting and presenting these patterns to educators. Alt-schools, for example, constantly monitor classrooms to collect digital, audio, and visual information about student interactions and behavior.³⁹ Teachers rely heavily on computer analytics to make sense of this data, still decide what students need next.

Personalized platforms

In contrast to data-informed decision-making, data-driven education systems do not support, but supplant human decisions. Instead, computers “personalize” learning automatically by evaluating instructional options in light of students’ profiles and delivering content accordingly.⁴⁰ In doing so, they try to mimic the way that teachers adapt to student needs in physical settings.

Computers can customize instruction in several ways.^{32,41,42} Some adaptive systems deliver different material, for example a review of concepts as opposed to practice problems, based on student responses. Others

let students advance through subject matter at their own pace.^{32,41,43–46} At Summit schools, for example, students work on a specific concept until they master it, independent of their classmates' progress.³⁹ At their most sophisticated, "intelligent tutoring" platforms do more than lead students through pre-determined pathways.

As already discussed above, big data analytics often use cognitive models to track and assess student progress. "Smart" learning systems perform a similar process using "tutoring" models that capture different instructional options.^{47,48} The software uses historical data about how students with similar profiles fared to determine the choice most likely to lead to student success.³³

To use a greatly simplified example, say that a student tries three times before entering a correct answer to practice problem. The system updates a student profile with this information, which creates a data pattern I shall call "ABCD." The tutoring model includes three different instructional options: it can show a video or have students review the relevant part of earlier lectures. The platform analyzes data created by prior students with ABCD profile patterns which show that 70% of students who watch new videos answer the next set of questions correctly, compared to 55% of those who review old material. The data accordingly "predicts" how the two options will affect the student currently using the system, and, in this case, plays the new video. Smart tutoring systems have yet to migrate into the mainstream, but they are poised to do so.^{2,3,49}

Visionary Benefits

Smart learning systems promise to promote both equality and efficacy.⁵⁰ The U.S. Department of Education and philanthropists including Bill and Melinda Gates and Mark Zuckerberg promote data-driven learning technologies as a means to a more effective, cost-efficient, and equitable education system.^{51–53}

Moving beyond the factory model

Proponents present personalized learning systems as a way to move past the one-size-fits-all factory model of education.^{44,54,55} Reformers and providers see automated instruction and assessment as a way to improve education quality, particularly in underserved and overcrowded schools.⁵⁶ In doing so, they also hope to address disparities in educational achievement and attainment across racial, ethnic, gender, and class categories and create more equitable access to opportunities.⁵⁷

Competency-based credit

Many reformers, including the U.S. Department of Education, want to use data to change how schools measure and document academic success. Algorithmic profiles "embed" assessment seamlessly into instruction instead of periodic, high stakes tests.⁵¹ Schools can award credit and document attainment based on technology-defined competencies, rather than traditional courses, grades, and credit hours.^{44,58} Students can streamline their education acquiring the specific skills they need.^{59–65} Competency-based credentials might also help students from less prestigious schools compete with their peers at more elite institutions.^{59–62,64–66}

Independent credentialing

Reformers promote mastery-based assessment and credentials as a means to capture student skills more accurately across institutions. However, competency-based credentials will only be valuable if admissions boards and employers trust in the accreditor.⁶⁷ The most ambitious CBE visions seek to employ distributed ledger technology like the block chain to create immutable, self-verifying records.⁶⁸ The decentralized nature of these record-keeping systems would also allow anyone to contribute to students' credentials, so that learners can accumulate recognized credit for informal learning and life experience outside classrooms.^{69,70}

Efficacy, Privacy, and Equity Concerns

The public conversation surrounding new big data-driven education technologies focuses on their effects, considering efficacy, privacy, and, equity concerns.^{71,51,72,73,14,37,74} The following section provides a brief overview of relevant issues as a contrast to the structural shifts discussed below.

Efficacy

Most reformers and schools evaluate big data-driven education technologies in terms of their efficacy in achieving defined learning outcomes.⁷⁵ In doing so, they overlook the problematic aspects of implementing big data driven education technologies at scale. The experimental nature of ed tech innovation means it is inevitable that some big data-driven education ventures will fail.^{41,76,77,78,75} The accuracy of outcomes will depend on the representativeness, accuracy, and relevance of the data incorporated into systems as well as the technologies themselves.⁷⁹ The size and complexity of data-driven systems makes it difficult to detect errors—and to implement the correct

adjustments upon doing so.^{38,80,81} How long will it take to discover that a particular tool is not effective or has unequal effect on different populations? Who will be responsible for tracking these and making sure?

Adjusting to flaws and failure will be increasingly complicated, given the highly politicized, bureaucratized, and decentralized structure of the U.S. public education system. It's one thing to have agile development for software that provides a relatively narrow array of services directly to users, like apps. It is another to have to change fundamental aspects of systems deployed nation-wide and with clients—schools and districts—who may not have the money or resources to keep track of and implement changes.^{82,83} We do not yet have the institutional, ethical, or governance mechanisms set up to grapple with beta education at scale.

Privacy

Increased data collection and school reliance on outside technology providers raise student privacy concerns about access to and commercial use of student information.^{82,84–86} Big data-driven education systems capture vast amounts of personal and personally identifiable information about students and teachers.^{87–89} Schools share this information with education technology providers who are predominantly private, for-profit entities.^{90,91} This data holds considerable commercial value outside the school context and apart from education purposes.⁹¹ Parents fear that companies will prioritize short-term profits over cautious information use and disclosure.^{92–95}

Traditional student privacy regulation, like the Family Educational Rights and Privacy Act (FERPA) don't address these issues sufficiently.^{86,96} Although FERPA theoretically provides parents with control over school sharing of personally identifiable student information, its exceptions delegate most data-related decisions to educators.^{68–71} Even in cases wherein parents do have the choice to opt-out of specific classroom technologies, they often do not feel they can do so in practice without putting their children at a significant social and academic disadvantage.⁸⁶

Newer state student privacy regulations try to ensure that vendors use student information appropriately through purpose limitations.^{97–102} Whether regulated directly or indirectly through school disclosure requirements, ed tech providers can only use student information for “school” purposes. As a result, these rules do not impose specific rules regarding school use of big data-driven education tools. Purpose limitations, how-

ever, provide minimal protection against problematic aspects of big data-driven education tools used in schools. These laws operate under the assumption that school purposes serve educational interests.¹⁰³ They do not account for institutional pressures that may put schools' interests at odds with students'.¹⁰⁴ Recently, for example, a university president tried to use predictive analytics to determine which freshmen to encourage to drop out in order to improve reported retention rates.¹⁰⁵ In addition, many state laws have an explicit exemption for personalized learning platforms.^{97–102} SOPIPIA allows education technology vendors to use covered student information for adaptive or customized services.

Equity

Big data-driven education also offers the promise of more equitable education outcomes, but may inadvertently have the opposite effect in the long run. Data-driven systems promise to be more precise and consistent than humans, which often leads to the presumption that they will be more accurate and objective as a result.⁷² Educators, even with the best of intentions, may rely on irrelevant or inappropriate factors in pedagogical decision-making.^{107,108} This can be because of bias toward certain groups or cognitive tendencies that may inadvertently shape decision-making of their conclusions, such as, whether a teacher grades student articles before or after lunch. Machine analysis may offer consistency, but that is not the same as objectivity. Algorithmic analysis can be just as biased as human decisions.^{109–111} Big data systems may incorporate input or create predictive models based on historical patterns of inequity.^{108,111–115}

Big data's predictive tools can similarly help or hinder socioeconomic mobility.^{41,76} Schools can monitor students to identify those at risk of dropping out in time to intervene.^{37,117,118} At the same time, predictions may be based on historical patterns that reflect existing inequities and discrimination.^{110,111,117,119} Predictions can also create self-fulfilling prophecies that unfairly limit future opportunities based on early performance.^{114,117,119} Long term predictions are particularly problematic in education spaces, which are explicitly environments dedicated to student development. Learning; and life trajectories, are rarely linear. Predictive analytics cannot “literally predict [student] life outcomes” because they cannot incorporate the impact of outside circumstances and student agency.^{116,120}

Structural Shifts

With all the focus on whether a specific technology “works” and who can access student information, educators often overlook the important structural shifts that occur even if technology or policy resolves the above-mentioned efficacy, privacy, and equity issues. These occur even in an impossibly perfect world where big data-driven tools perform as intended, technology providers only use student data to serve school purposes, and data analysis does not reflect hidden bias.

Monitored and memorialized learning environments

Smart learning platforms fundamentally alter learning environments by imposing new information infrastructures. Constant collection, scoring, and memorialization reduces the intellectual privacy characteristic of physical classrooms.^{13,17,121,122} The current approach to big data analytics presumes more data are better, leading to the expansion of types of information collected about students and increasingly creating spaces of pervasive surveillance.¹²³ This constant monitoring has documented chilling effects on student expression, risk-taking, and diversity of opinion.^{124,125} Data-driven education environments accordingly undermine the traditional safety that supports the learning process. Research suggests that students’ sense of vulnerability impedes academic promise and disproportionately affects minorities.^{100,101}

As already discussed above, data-driven learning platforms continuously assess student progress. While teachers do the same in physical classrooms, technologically-mediated education environments collect and capture of students’ experiments and mistakes during the learning process.¹²⁶ This collapse of formative feedback, summative assessment, and credentialing raises the stakes of every mistake or misstep.⁸⁶ Students’ every action might be incorporated into the digital equivalent of transcripts.^{127–129} Students’ early mistakes can be preserved for later scrutiny and mined for new algorithmic inferences.¹³⁰

The prospect of preserving these records using block chain technology ratchets up the stakes even more. The open nature of the block chain makes these truly public permanent records. This runs counter to the consistent theme in U.S. society, economic policy, and political rhetoric that past should not unduly limit future opportunities. Like the expungement of juvenile criminal records or old bankruptcy proceedings, the practical obscurity of classroom proceedings promotes what Andrew Tutt refers to as “revisability.”^{130,131} The surveil-

lance and memorialization in virtual learning environments has the potential to discourage the intellectual experimentation, free expression, and creativity commonly promoted as goals of big data-driven education and America’s education system at large.^{82,85,103}

Displaced pedagogical decision-making

Schools outsourcing pedagogical functions to companies outsource important decisions to them as well. The networked flow of information replaces the transparency, autonomy, and accountability expected in education spaces with standardized and decontextualized decision-making.^{124,125} Changes in educational evaluation and credentialing shifts the power dynamic to the entities creating competency models and evaluative systems.¹³² In performing these seemingly mundane processes, technologies—and their corporate providers—in fact exercise significant authority over fundamental aspects of education that have previously been invested in teachers, school administrators, and (often local) policy makers.¹³²

Automated education tools also reduce the autonomy of on-site educators. While personalized learning tools may be more “customized,” they still create standardized systems.^{133,134} Because algorithmic analysis relies on probabilities, data-driven instruction, evaluation, and credentials cannot reflect or react to the unique aspects of specific circumstances.¹³⁵ Resulting path dependencies limit educators’ ability to deviate based on highly contextual circumstances to the degree that educators and institutions defer to algorithmic determinations.^{124,134,136} This cuts against the highly contextualized decision-making characteristic of physical classroom settings. It goes against the idealized education values espoused by big data-oriented reformers as well as their critics, who seek to treat students equally regardless of their group affiliations.¹³⁷

By relocating the site of pedagogical functions, data-driven education technologies make it more difficult for students, parents, and communities to exercise agency and demand accountability.^{114,138} Instead of readily available teachers and administrators, stakeholder must navigate remote corporate communication structures. Those who do obtain information about decision-making may not be able to make sense of the complex algorithms and probabilistic decisions driving personalized learning.^{81,139} The lack of transparency and obvious sources to exercise agency or ensure accountability may exacerbate the alienation of students and parents who already feel disconnected and disempowered in the traditional system.^{133,140}

Inconsistent implementation of big data-driven education technologies creates the risk of a two-tier system. Underserved schools may lack the resources to provide more flexible instruction and assessment compared with better funded counterparts, and rely almost exclusively on the determinations of smart education technologies. Students enrolled in schools which can afford to allow teachers to deviate from algorithmic recommendations or supplement automated assessment with personal evaluation will have the opportunity for accommodation based on individual circumstances. Less fortunate students may receive automatically differentiated instruction without the flexibility of contextualized assessment.

Computable competencies and priorities

Digital mediation of student—instructor communication changes more than the mode of delivery. It changes the content, metrics, and goals of education itself. This makes the stakes much higher than, say, the information practices governing a basic commercial transaction. In virtual retail environments, for example, information infrastructures that shape customers' shopping experiences and perhaps their choice of widget, but do not alter the nature of the widget itself. In learning environments, information technologies perform crucial functions, and, accordingly, shape education's content, metrics, and values.

Competency maps and measurement systems are not the result of a transparent reflection of reality.^{95,105} Traditional grading involves human interpretation—whether through informal mental processes or through explicit rubrics. Automated commensuration of students' educational experience is similarly interpretive. The categorization of tasks and knowledge, the criteria for reaching “competency,” and the metrics used to track and measure students' achievement all involve value-laden decisions about relevant information, learning processes, and desired outcomes.^{41,82,107,108,136,142} These choices end up creating the epistemology that then defines education.

The hardware and software used to collect, capture, analyze, interpret, and store student data also limits the content measures, and format of the student scores and credentials.^{108,143} These tools only measure and, accordingly, can only respond to the features or variables factored into the algorithmic process. This gives short shrift to the psychosocial skills research that increasingly shows to be essential for education attainment.^{125,133,134} Big data-driven education platforms

focus on student mastery of skill and knowledge acquisition. In doing so, they implicitly define education as a collection of expertise and demonstrable abilities.^{89,107} Although some technologies attempt to measure metacognitive abilities, computer platforms cannot currently capture the “soft” skills, like teamwork, crucial for long-term success.^{125,133}

Data-driven education systems also have the potential to inadvertently shift education away from paradigms that promote unquantifiable values to solely instrumental ones.^{144–148} Schools in America have historically served a plurality of purposes, including cultivating civic participation, promoting socioeconomic mobility, and encouraging intellectual fulfillment.^{4,149} Overreliance on big data discrimination education technologies risks reducing these more abstract goals to an afterthought.

Conclusion

As pervasive data collection and mining to feed learning analytics creates ubiquitous surveillance, these consequences will impact more and more of everyday life outside of formal education. They change the data used, the evaluation mechanisms, and the ultimate format or records used to assess and represent student achievement, academic credit, and intellectual mastery. The value judgments and commensuration inherent in these systems are often opaque and may be inadvertent, but have important consequences on what “counts” as and toward education and achievement in academic and employment environments.

Each shift in pedagogical decision-making has the potential for unintended consequences because of inaccurate or unrepresentative data, algorithmic bias or disparate impact, scientism replacing more holistic and contextualized personal evaluation, and the exclusion of noncomputable variables and nonquantifiable learning outcomes. Examining big data-driven education in light of structural dynamics teases out the agendas that are being advanced—intentionally or otherwise—when adopting data-driven education technologies.

It is important that the changes wrought by big data in education are not made unknowingly and inadvertently by thoughtlessly implementing new technologies. They should instead be the result of considered choices sufficiently transparent to permit public scrutiny. This may mean requiring more transparency, accountability, or precautionary approaches to information and privacy practices in learning environments. Just as we adopt new approaches based on technological

affordances, schools must also implement accompanying oversight and governance structures that match them.

Author Disclosure Statement

The author received support from Microsoft for research related to privacy in early massive open online courses as part of her prior affiliation with New York University's Information Law Institute.

References

- Price, T. in *Encyclopedia of Strategic Leadership and Management* 1314–1325 (IGI Global, 2017).
- Adams Becker, S., Freeman, A., Giesinger Hall, C., Cummins, M. & Yuhnke, B. *NMC/CoSN Horizon Report: 2016 K-12 Edition*. (2016).
- Adams Becker, S., Estrada, V., Freeman, A. & Johnson, L. *NMC horizon report: 2017 Higher education edition*. Austin, Texas: The New Media Consortium (2017). at <http://academedia.org/2017_NMC_horizon.pdf>
- Labaree, D. F. Public Goods, Private Goods: The American Struggle Over Educational Goals. *American Educational Research Journal* 34, 39–81 (1997).
- Bates, T. Computer Assisted Learning or Communications: Which Way for Information Technology in Distance Education? *International Journal of E-Learning & Distance Education* 1, 41–57 (1986).
- Zhang G, Li J, Hao L. Research on Cloud Computing and Its Application in Big Data Processing of Distance Higher Education. *International Journal of Emerging Technologies in Learning* 10, 55–58 (2015).
- Hirsch, B. & Ng, J. W. P. Education Beyond the Cloud: Anytime-Anywhere Learning in a Smart Campus Environment. in *Internet Technology and Secured Transactions (ICITST), 2011 International Conference for* 718–723 (IEEE, 2011).
- Pappano, L. The Year of the MOOC. *The New York Times* 2, 2012 (2012).
- Shah, D. MOOC Trends in 2015: Big MOOC Providers Find their Business Models. *Class Central's MOOC Report* (2015). at <<https://www.class-central.com/report/mooc-business-model/>>
- Stinson, L. Tinybop makes kids apps that are so strange and beautiful, even adults love to play them. *WIRED* (2015). at <<https://www.wired.com/2015/11/tinybop-dials-up-the-weird-to-craft-stellar-kids-apps/all/1/>>
- Young, J. R. The New Transcript. *The Chronicle of Higher Education* (2016). at <<http://chronicle.com/article/The-New-Transcript/235450>>
- Vance, A. & Tucker, J. W. *School Surveillance: The Consequences for Equity and Privacy*. (2016).
- Mayer-Schönberger, V. & Cukier, K. *Learning With Big Data*. (Eamon Dolan/Houghton Mifflin Harcourt, 2014).
- Boninger, F. & Molnar, A. *Learning to Be Watched: Surveillance Culture at School*. National Education Policy Center (National Education Policy Center, 2016). at <<http://nepc.colorado.edu/files/publications/RB%20Boninger-Molnar%20Trends.pdf>>
- Hirsch, B. & Ng, J. W. P. Education Beyond the Cloud: Anytime-Anywhere Learning in a Smart Campus Environment. in 718–723 (IEEE, 2011).
- Junco, R. & Clem, C. Predicting course outcomes with digital textbook usage data. *The Internet and Higher Education* 27, 54–63 (2015).
- Atif, Y., Mathew, S. S. & Lakas, A. Building a smart campus to support ubiquitous learning. *Journal of Ambient Intelligence and Humanized Computing* 6, 223–238 (2015).
- Horowitz, E. A Peek at a 'Smart' Classroom Powered by the Internet of Things. *EdSurge* (8/11/15). at <<https://www.edsurge.com/news/2015-08-11-a-peek-at-a-smart-classroom-powered-by-the-internet-of-things>>
- Singer, N. Online Test-Takers Feel Anti-Cheating Software's Uneasy Glare. *The New York Times* (2015). at <<http://www.nytimes.com/2015/04/06/technology/online-test-takers-feel-anti-cheating-software-uneasy-glare.html>>
- Kravets, D. Tracking School Children With RFID Tags? It's All About the Benjamins. *WIRED* (2012). at <<http://www.wired.com/2012/09/rfid-chip-student-monitoring/>>
- <http://www.washingtontimes.com>, T. W. T. Oral Roberts University requires freshmen to wear Fitbit, track 10K steps per day. *The Washington Times* (2016). at <<http://www.washingtontimes.com/news/2016/jan/11/oklahoma-university-requires-freshmen-to-wear-fitbit/>>
- Was, C. *Eye-Tracking Technology Applications in Educational Research*. (IGI Global, 2016). at <<http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-1005-5>>
- Eyes Don't Lie: A New Way to Use Gaze Tracking in the Classroom (EdSurge News). *EdSurge* at <<https://www.edsurge.com/n/2015-04-27-eyes-don-t-lie-a-new-way-to-use-gaze-tracking-in-the-classroom>>
- Cope, B. & Kalantzis, M. Big Data Comes to School: Reconceptualizing Evidence and Research in the Era of Technology-mediated Learning. (2014). at <http://kc-test.cg-creator.com/_uploads/Big_Data_Comes_to_School.pdf>
- Dede, C. in *Towards Ubiquitous Learning* (Kloos, C. D., Gillet, D., García, R. M. C., Wild, F. & Wolpers, M.) 1–8 (Springer Berlin Heidelberg, 2011). doi:10.1007/978-3-642-23985-4_1
- Gibson, D. C., Webb, M. & Ifenthaler, D. Challenges of Big Data in Educational Assessment. *Proceedings of the IADIS International Conference on Cognition & Exploratory Learning in Digital Age* 92–100 (2015).
- Tempelaar, D. T., Heck, A., Cuypers, H., van der Kooij, H. & van de Vrie, E. Formative Assessment and Learning Analytics. in *Proceedings of the Third International Conference on Learning Analytics and Knowledge* 205–209 (ACM, 2013).
- Nkambou, R., Mizoguchi, R. & Bourdeau, J. *Advances in intelligent tutoring systems*. 308, (Springer Science & Business Media, 2010).
- Baker, R. Using Learning Analytics in Personalized Learning. *Handbook on personalized learning for states, districts, and schools*. Philadelphia, PA: Center on Innovations in Learning (2016).
- Baker, R. in *Handbook on personalized learning for states, districts, and schools* (Center on Innovations in Learning, 2016). at <http://www.centeril.org/2016handbook/resources/Cover_Baker_web.pdf>
- Boulanger, D. et al. in *Emerging Issues in Smart Learning* (Chen, G., Kumar, V., Kinshuk, Huang, R. & Kong, S. C.) 289–296 (Springer Berlin Heidelberg, 2015). at <http://link.springer.com/10.1007/978-3-662-44188-6_39>
- Arnold, K. E. & Pistilli, M. D. Course signals at Purdue: using learning analytics to increase student success. in 267 (ACM Press, 2012). doi:10.1145/2330601.2330666
- Baker, R. S. Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education* 26, 600–614 (2016).
- Mead, R. Learn Different. *The New Yorker* (2016). at <<http://www.newyorker.com/magazine/2016/03/07/altschools-disrupted-education>>
- New, J. *Building a Data-Driven Education System in the United States*. (2016). at <<http://dataserver.lrp.com/DATA/servlet/DataServlet?fname=2016-data-driven-education.pdf>>
- Bulger, M. *Personalized Learning: The Conversations We're Not Having*. (2106). at <http://www.datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf>
- Samwel, E. Toward the Development and Implementation of Personalized, Adaptive, and Comprehensive E-learning Systems. (2016). at <http://nsuworks.nova.edu/gscis_etd/373/>
- Vander Ark, T. *The Future of Learning: Personalized Adaptive and Competency-Based*. *DreamBox Learning* 1–13 (2013). at <<http://www2.dreambox.com/14872/2014-10-10/3q28b5>>
- Ark, T. V. *Educational Data Analysis With MOOCs and Beyond*.
- Redding, S. in *Handbook on personalized learning for states, districts, and schools* (Center on Innovations in Learning, 2016). at <http://www.centeril.org/2016handbook/resources/Redding_chapter_web.pdf>
- Aleven, V., McLaughlin, E. A., Glenn, R. A. & Koedinger, K. R. Instruction based on adaptive learning technologies. *Handbook of research on learning and instruction*. Routledge (2016). at <http://www.cs.cmu.edu/~aleven/Papers/2016/Aleven_et_al_Handbook2016_AdaptiveLearningTechnologies.pdf>
- Koedinger, K. R., Anderson, J. R., Hadley, W. H. & Mark, M. A. Intelligent tutoring goes to school in the big city. (1997).
- Nunn, S., Avella, J. T., Kanai, T. & Kebritchi, M. Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning* 20, (2016).

45. Clerkin, K. Three Ways To Fix the Workforce Readiness Gap Revealed By the Gallup/Lumina Survey of Employers. *College For America* (2014). at <<http://collegeforamerica.org/three-ways-to-fix-the-workforce-readiness-gap/>>
46. *Teachers Know Best: Making Data Work for Teachers and Students*. (2015). at <<http://k12education.gatesfoundation.org/learning/teachers-know-best-making-data-work/>>
47. Bienkowski, M. & Feng, M. *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief. Proceedings of conference on advanced technology for education* (2012). at <<http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf>>
48. Herold, B. Facebook's Zuckerberg to Bet Big on Personalized Learning. *Education Week* (2016). at <<http://www.edweek.org/ew/articles/2016/03/07/facebook-zuckerberg-to-bet-big-on-personalized.html>>
49. McIntire, M. E. 'Machine Teaching' Is Seen as Way to Develop Personalized Curricula. *The Chronicle of Higher Education* (2015). at <<http://ezproxy.library.nyu.edu:3715/blogs/wiredcampus/machine-teaching-is-seen-as-way-to-develop-personalized-curricula/57209>>
50. Andreou, S. N. & Koutsampelas, C. in *Using Data to Improve Higher Education* (Menon, M. E., Terkla, D. G. & Gibbs, P.) 181–193 (Sense-Publishers, 2014). at <http://link.springer.com/chapter/10.1007/978-94-6209-794-0_11>
51. Alli, N., Rajan, R. & Ratliff, G. How Personalized Learning Unlocks Student Success. (2016). at <<https://er.educause.edu/~media/files/articles/2016/3/erm1621.pdf>>
52. 'Machine teaching' holds the power to illuminate human learning. *EurekAlert!* (2015). at <http://www.eurekalert.org/pub_releases/2015-08/uow-th081115.php>
53. Webley, K. The Adaptive Learning Revolution. *Time* (2013).
54. DeMark, S. Western Governors University's assessment quality rubrics: high standards for assessments in CBE programs. *The Journal of Competency-Based Education* 1, 85–89 (2016).
55. Eaton, J. S. Accreditation and competency-based education. *The Journal of Competency-Based Education* 1, 12–16 (2016).
56. Everhart, D., Bushway, D. & Schejbal, D. *Communicating the Value of Competencies*. (2016). at <<https://www.acenet.edu/news-room/Pages/Communicating-the-Value-of-Competencies.aspx>>
57. Twyman, J. S. *Competency-Based Education: Supporting Personalized Learning. Connect: Making Learning Personal*. (Center on Innovations in Learning, Temple University, 2014). at <<http://eric.ed.gov/?id=ED558055>>
58. Neal, S. Competency-based education will redefine how future degrees are attained. *Enrollment Management Report* 18, 3–3 (2015).
59. Paquette, G., Mariño, O., Rogozan, D. & Léonard, M. Competency-based personalization for massive online learning. *Smart Learn. Environ.* 2, 4 (2015).
60. Garrett, R. & Lurie, H. *Deconstructing CBE: An Assessment of Institutional Activity, Goals, and Challenges in Higher Education*. (2016).
61. Neal, S. Competency-based education will redefine how future degrees are attained. *Enrollment Management Report* 18, 3 (2015).
62. Craig, R. Coursera, Udacity And The Future of Credentials. *Forbes* (2015). at <<http://www.forbes.com/sites/ryanecraig/2015/09/30/coursera-udacity-and-the-future-of-credentials/>>
63. Jagers, C. Verifiable Credentials on the Blockchain: Disclaimer, Convenience and Necessity, Lifelong Learning, Lifting the Ecosystem. *Medium* (2016). at <https://medium.com/learning-machine-blog/blockchain-credentials-b4cf5d02bbb7?source=reading_list> 62-3>
64. Vian, K. Own Your Achievements: Three Ways Blockchain Tech is Disrupting Education. *Blockchain Futures Lab* (2016). at <<http://blockchainfutureslab.wordpress.com/2016/03/16/own-your-achievements-three-ways-blockchain-tech-is-disrupting-education/>>
65. Rath, D. How Blockchain Will Disrupt the Higher Education. *Campus Technology* (5/16/16). at <<https://campustechnology.com/articles/2016/05/16/how-blockchain-will-disrupt-the-higher-education-transcript.aspx>>
66. Macfadyen, L. P. & Dawson, S. Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Educational Technology & Society* 15, 149–163 (2012).
67. Klein, C., Lester, J., Rangwala, H. & Johri, A. Using predictive analytics in teaching and advising: Promising practices and practical considerations. in *Innovations in Teaching & Learning Conference Proceedings* 8, 103 (2016).
68. Papamitsiou, Z. & Economides, A. A. Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Journal of Educational Technology & Society* 17, 49–64 (2014).
69. Pasquale, F. *The black box society: the secret algorithms that control money and information*. (Harvard University Press, 2015).
70. Citron, D. K. & Pasquale, F. A. The scored society: due process for automated predictions. *Washington Law Review* 89, (2014).
71. Cumbley, R. & Church, P. Is 'Big Data' creepy? *Computer Law & Security Review* 29, 601–609 (2013).
72. Zeide, E. Moving Beyond FERPA and FIPPs: Student Privacy Principles for the Age of Big Data. *Drexel Law Review* (2016).
73. Rubel, A. & Jones, K. M. L. Student Privacy in Learning Analytics: An Information Ethics Perspective. *Forthcoming in The Information Society* (2015) (2014).
74. Slade, S. & Prinsloo, P. Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist* 57, 1–20 (2013).
75. Armerding, T. Big Brother goes to school. *CSO Online* (2014). at <<http://www.csoonline.com/article/2841969/big-data-security/schools-keep-track-of-students-online-behavior-but-do-parents-even-know.html>>
76. Herold, B. Analytics in K-12 Schools: Big Data, or Big Brother? *Education Week* (2016). at <<http://www.edweek.org/ew/articles/2016/01/13/analytics-in-k-12-schools-big-data-or.html>>
77. Simon, S. Big Brother: Meet the Parents. *Politico* (6/6/14). at <<http://www.politico.com/story/2014/06/internet-data-mining-children-107461.html>>
78. Watters, A. Top Ed-Tech Trends of 2016: For Profits. *Hack Education* at <<http://2016trends.hackeducation.com/for-profits>>
79. Matsons, M. R. & Sandronsky, S. How Big Companies Like Microsoft, Pearson Are Gearing to Profit from Our Children's Successes and Failures. *AlterNet* (2014). at <http://www.alternet.org/education/how-big-companies-microsoft-pearson-are-gearing-profit-our-childrens-successes-and?paging=off¤t_page=1#bookmark>
80. Nicholson, C. Are You Securing My Kids' Data? *EdSurge* (2013). at <<https://www.edsurge.com/n/2013-07-15-are-you-securing-my-kids-data>>
81. Reidenberg, J. R. et al. *Privacy and Cloud Computing in Public Schools*. (2013).
82. Zeide, E. Parsing Student Privacy. *Technology | Academics | Policy* (2015). at <<http://www.techpolicy.com/Blog/Featured-Blog-Post/Parsing-Student-Privacy.aspx>>
83. Singer, N. InBloom Student Data Repository to Close. *N.Y. Times* (4/21/14). at <<http://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close/>>
84. Zeide, E. Student Privacy Principles for the Age of Big Data: Moving Beyond FERPA and FIPPs. *Drexel Law Review* 8, 339 (2016).
85. Solove, D. J. FERPA and the Cloud: Why FERPA Desperately Needs Reform. *LinkedIn* (2012). at <<https://www.linkedin.com/pulse/20121211124311-2259773-ferpa-and-the-cloud-why-ferpa-desperately-needs-reform>>
86. Steinberg, D. *Student Online Personal Information Protection Act*. *Cal. Bus. & Prof. Code* (9/30/14).
87. *The Student Online Personal Information Protection Act*. *Arkansas* 6-18-109, (2015).
88. *Student Data Privacy Act*. *Md. Code. Ann. Educ.* (2015). at <http://mgaleg.maryland.gov/2015RS/chapters_noln/Ch_413_hb0298E.pdf>
89. *Student Data Privacy Protection Act*. *Del. Code. Ann.* 14, (2016).
90. *Oregon Student Information Protection Act*. *Or. Rev. Stat.* 20, (2015).
91. *The Student Information Privacy Act*. *Me. Rev. Stat.* 20-A, (2016).
92. Zeide, E. The Limits of Education Purpose Limitations. *U. Miami L. Rev.* 71, 493 (2017).
93. Sykes, S. Drowning Bunnies or Saving Lives? Putting Data In The Hands Of Academic Advisors. *The EvOLLution* (2016). at <<http://evollution.com/technology/tech-tools-and-resources/drowning-bunnies-or-saving-lives-putting-data-in-the-hands-of-academic-advisors/>>
94. Sawchuk, S. Research Detects Bias in Classroom Observations. *Education Week* (2014). at <<http://www.edweek.org/ew/articles/2014/05/13/32observe.h33.html?tkn=NMLFiGKAHohBYILX0avcll9hOI3zOW5LKWbq&intc=es&intc=mes>>
95. Espeland, W. N. & Stevens, M. L. A sociology of quantification. *European Journal of Sociology* 49, 401–436 (2008).

96. Selwyn, N. *Distrusting educational technology: critical questions for changing times*. (Routledge, Taylor & Francis Group, 2014).
97. Crawford, K. The Hidden Biases in Big Data. *Harvard Business Review* (4/13). at <<http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>>
98. Calders, T. & Žliobaitė, I. in *Discrimination and Privacy in the Information Society* (Custers, B., Calders, T., Schermer, B. & Zarsky, T.) 43–57 (Springer Berlin Heidelberg, 2013). doi:10.1007/978-3-642-30487-3_3
99. boyd, danah & Crawford, K. Six Provocations for Big Data. (2011). at <<http://papers.ssrn.com/abstract=1926431>>
100. Lerman, J. Big Data and Its Exclusions. *Stanford Law Review Online* 66, 55 (2013).
101. Prinsloo, P. & Slade, S. in *Big Data and Learning Analytics in Higher Education* (Daniel, B. K.) 109–124 (Springer International Publishing, 2017). doi:10.1007/978-3-319-06520-5_8
102. Ekowo, M. & Palmer, I. *The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis*. (2016).
103. Har Carmel, Y. & Harel Ben-Shahar, T. Reshaping Ability Grouping Through Big Data. (2017). at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2944743>
104. Subotzky, G. & Prinsloo, P. Turning the tide: a socio-critical model and framework for improving student success in open distance learning at the University of South Africa. *Distance Education* 32, 177–193 (2011).
105. Williamson, B. in *Surveillance Futures: Social and Ethical Implications of New Technologies for Children and Young People* 50 (2016). at <<https://books.google.com/books?hl=en&lr=&id=S83LDAAAQBAJ&oi=fnd&pg=PT64&ots=W16amByPd-&sig=elzP7xpG2lTsGXRjz2CM23f2wcE>>
106. Cuban, L. Algorithms, Accountability, and Professional Judgment (Part 3). *Larry Cuban on School Reform and Classroom Practice* (2013). at <<http://larrycuban.wordpress.com/2013/01/15/algorithms-accountability-and-professional-judgment-part-3/>>
107. Cuban, L. The School of the Future: Automated Classrooms? (Part 1). *Larry Cuban on School Reform and Classroom Practice* (2015). at <<https://larrycuban.wordpress.com/2015/01/15/the-school-of-the-future-automated-classrooms-part-1/>>
108. stevenmsinger. Standardized Tests Every Day: the Competency Based Education Scam. *gadflyonthewallblog* (2015). at <<https://gadflyonthewallblog.wordpress.com/2015/11/30/standardized-tests-every-day-the-competency-based-education-scam/>>
109. Cody, A. Will the Data Warehouse Become Every Student and Teacher's 'Permanent Record'? *Education Week* (2013). at <http://blogs.edweek.org/teachers/living-in-dialogue/2013/05/will_the_data_warehouse_become.html?cmp=SOC-SHR-FB>
110. Sirota, D. Big Data Means Kids' 'Permanent Records' Might Never Be Erased. *Motherboard* (2013). at <<http://motherboard.vice.com/blog/permanent-records-are-hurting-kids>>
111. Zeide, E. The Proverbial Permanent Record. (2014). at <<http://papers.ssrn.com/abstract=2507326>>
112. Watters, A. The Ideology of the Blockchain (for Education). *Hack Education* (2016). at <<http://hackededucation.com/2016/04/14/blockchain-ideology>>
113. Big Data Means Kids' 'Permanent Records' Might Never Be Erased. (0). at <<http://motherboard.vice.com/blog/permanent-records-are-hurting-kids>>
114. Roberts, L. D., Chang, V. & Gibson, D. in *Big Data and Learning Analytics in Higher Education* (Daniel, B. K.) 89–108 (Springer International Publishing, 2017). doi:10.1007/978-3-319-06520-5_7
115. Kohn, A. Four Reasons to Worry About 'Personalized Learning'. *Alfie Kohn* (2015). at <<http://www.alfiejohn.org/blogs/personalized/>>
116. Watters, A. The problem with 'personalisation'. *Australian Educational Leader* 36, 55 (2014).
117. Kohn, A. Four Reasons to Worry About 'Personalized Learning'. (2015).
118. Roberts-Mahoney, H., Means, A. J. & Garrison, M. J. Netfixing human capital development: personalized learning technology and the corporatization of K-12 education. *Journal of Education Policy* 31, 405–420 (2016).
119. Hartley, R. Conceptualising and supporting the learning process by conceptual mapping. *Smart Learn. Environ.* 1, 7 (2014).
120. Lessig, L. *Code: And Other Laws of Cyberspace, Version 2.0*. (Basic Books, 2006).
121. Wilson, M. Personalization: It's Anything but Personal. *Educational Leadership* 71, 73–77 (2014).
122. Selwyn, N. *Distrusting educational technology: critical questions for changing times*. (Routledge, Taylor & Francis Group, 2014).
123. Slade, S. & Prinsloo, P. Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist* 57, 1510–1529 (2013).
124. Macfadyen, L. P. & Dawson, S. Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Educational Technology & Society* 15, 149–163 (2012).

Cite this article as: Zeide E (2017) The structural consequences of big data-driven education. *Big Data* 5:2, 164–172, DOI: 10.1089/big.2016.0061.

Abbreviations Used

CBE = competency-based education
FERPA = Family Educational Rights and Privacy Act
MOOCs = massive open online courses