



The cyclical ethical effects of using artificial intelligence in education

Edward Dieterle¹ · Chris Dede² · Michael Walker³

Received: 30 July 2021 / Accepted: 31 March 2022

© Educational Testing Service, under exclusive license to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract

Our synthetic review of the relevant and related literatures on the ethics and effects of using AI in education reveals five qualitatively distinct and interrelated divides associated with access, representation, algorithms, interpretations, and citizenship. We open our analysis by probing the ethical effects of *algorithms* and how teams of humans can plan for and mitigate bias when using AI tools and techniques to model and inform instructional decisions and predict learning outcomes. We then analyze the upstream divides that feed into and fuel the algorithmic divide, first investigating *access* (who does and does not have access to the hardware, software, and connectivity necessary to engage with AI-enhanced digital learning tools and platforms) and then *representation* (the factors making data either representative of the total population or over-representative of a subpopulation's preferences, thereby preventing objectivity and biasing understandings and outcomes). After that, we analyze the divides that are downstream of the algorithmic divide associated with *interpretation* (how learners, educators, and others understand the outputs of algorithms and use them to make decisions) and *citizenship* (how the other divides accumulate to impact interpretations of data by learners, educators, and others, in turn influencing behaviors and, over time, skills, culture, economic, health, and civic outcomes). At present, lacking ongoing reflection and action by learners, educators, educational leaders, designers, scholars, and policymakers, the five divides collectively create a vicious cycle and perpetuate structural biases in teaching and learning. However, increasing human responsibility and control over these divides can create a virtuous cycle that improves diversity, equity, and inclusion in education. We conclude the article by looking forward and discussing ways to increase educational opportunity and effectiveness for all by mitigating bias through a cycle of progressive improvement.

Keywords Artificial intelligence · Education · Equity · Ethics

1 Introduction

Digital tools and learning platforms are pervasive in education. Educators, students in pre-K–12 and higher education, workers seeking to upskill or reskill, and informal learners of all ages increasingly engage with digital experiences (Decuyper et al. 2021). As a result, they generate enormous

amounts of multimodal data, such as logfiles; audio, video, and text files; and eye tracking data (Giannakos et al. 2019). Analyzing those data through artificial intelligence (AI) techniques, such as machine learning, computer vision, and natural language processing, can answer instructional and administrative questions, discover new and non-obvious relationships and patterns, predict learning outcomes, and automate low-level decisions. Concurrently, complex and interrelated ethical questions about learning and teaching underpin the stages associated with generating, analyzing, and interpreting data with AI.

Our synthetic review of the relevant and related literatures on the ethics and effects of using AI in education reveals five qualitatively distinct and interrelated divides associated with access, representation, algorithms, interpretations, and citizenship. Collectively, the divides have the potential to behave cyclically in a virtuous cycle, as depicted in Fig. 1, to enhance diversity, equity, and inclusion in education. However, unless

✉ Edward Dieterle
edieterle@ets.org

Chris Dede
chris_dede@gse.harvard.edu

Michael Walker
mwalker@ets.org

¹ Educational Testing Service, Washington, DC, USA

² Harvard University, Cambridge, MA, USA

³ Educational Testing Service, Princeton, NJ, USA

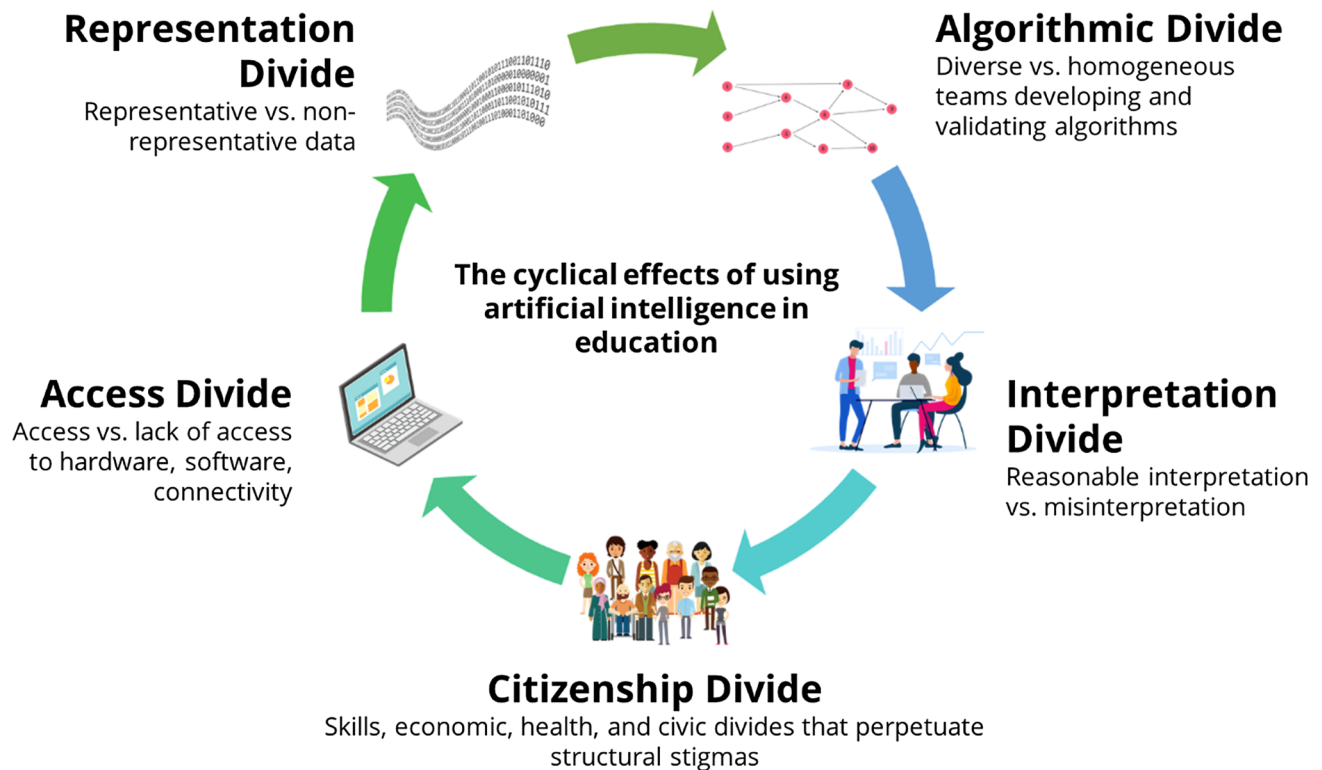


Fig. 1 The cyclical effects of using artificial intelligence in education

we increase reflection and action by stakeholders (i.e., learners, educators, educational leaders, designers, scholars, and policy makers), their current behaviors, choices, and interpretations will deepen these divides and perpetuate structural biases in teaching and learning, thereby furthering inequity and creating a vicious cycle.

The conditions necessary to *do things better* and *do better things* with AI-enhanced digital tools and learning platforms within each divide and across the cycle require ongoing collaborative reflection and improvement by all stakeholders. The COVID-19 pandemic, by accelerating the adoption of technology in education (Schiff 2021), has heightened the implications and urgency of understanding this cycle and its related ethical questions. As learners increasingly rely on AI-enhanced digital tools and online learning platforms, the stakes increase for all involved (Decuypere et al. 2021). In every segment of the cycle, we consider how social, cultural, and other individual differences among various stakeholders shape decisions. We conclude the article by looking forward and discussing ways to increase opportunity and equity while mitigating bias.

2 Five qualitatively distinct and interrelated divides

We open our analysis by probing the ethical effects of *algorithms* and how teams of people can plan for and mitigate bias when using AI tools and techniques to model and inform instructional decisions and predict learning outcomes. We analyze the upstream factors that feed into and fuel the algorithmic divide, first investigating *access* (who does and does not have access to the hardware, software, and connectivity necessary to engage with AI-enhanced digital learning tools and platforms) and then *representation* (the factors making data either representative of the total population or over-representative of a subpopulation's preferences, thereby preventing objectivity and biasing understandings and outcomes). We then analyze the algorithmic divide's downstream consequences associated with *interpretation* (how learners, educators, and others understand the outputs of algorithms and use them to make decisions) and *citizenship* (how the other divides

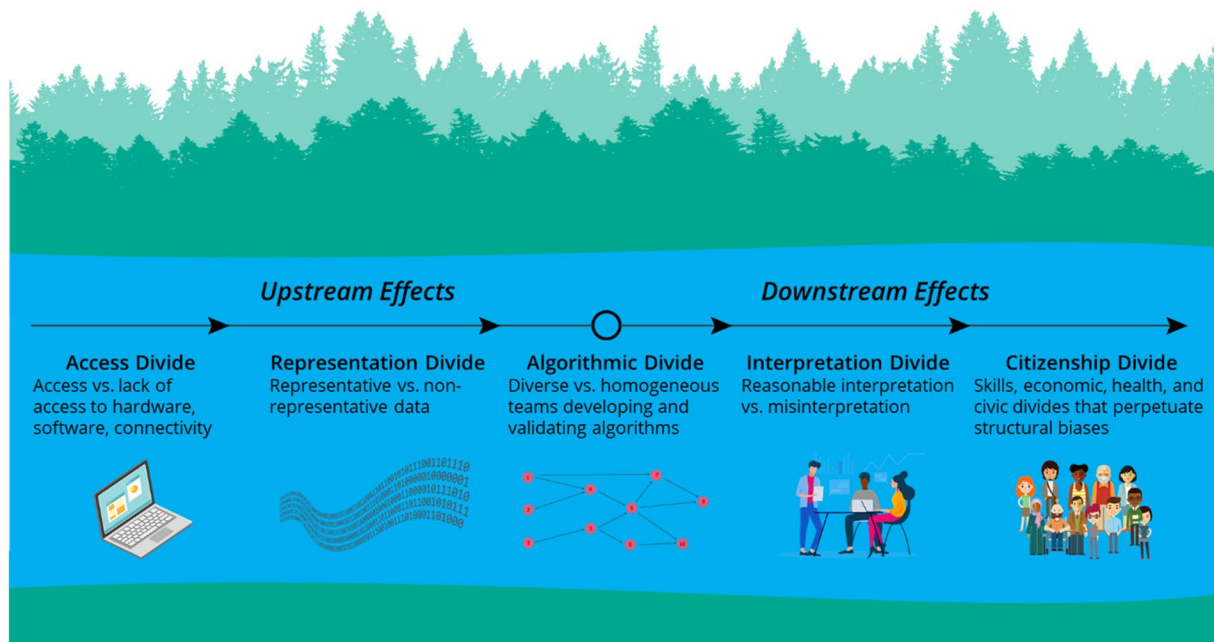


Fig. 2 The upstream and downstream effects of ethical decisions involving education and artificial intelligence

accumulate to impact interpretations of data by learners, educators, and others, which in turn influence behaviors and, over time, skills, culture, economic, health, and civic outcomes). Figure 2 illustrates how these upstream divides influence the algorithmic divide and how the algorithmic divide, in turn, amplifies downstream effects.

3 The algorithm divide

To develop the algorithms that underpin digital tools and learning platforms, researchers and designers from the AI, educational data mining, and learning analytics communities have developed and applied advanced statistical and computational methods to model big educational data (e.g., Levy 2019; Niemi et al. 2018). Such algorithms measure and enhance disciplinary knowledge and understanding (Fischer et al. 2020; Heffernan and Heffernan 2014); predict standardized test scores (Adjei et al. 2017; Pardos et al. 2014) and academic achievement (Jiang et al. 2019; Kostyuk et al. 2018); recommend academic pathways (Shao et al. 2021); and measure student engagement and boredom (D'Mello et al. 2017), creativity (Shute and Ventura 2013), persistence (Wang et al. 2020), inquiry competencies (Sao Pedro et al. 2013), and problem-solving (Shute and Wang 2017). Research has also shown the psychometric value of AI analysis techniques, when applied to big educational data, to detect rapid guessing in assessments (Guo et al. 2016), understand test-taking strategies (Stadler et al. 2019),

improve test design (Lee and Haberman 2016), and improve score reliability (van Rijn and Ali 2017).

Despite its relative brief history compared to other scholarly disciplines, these research communities have built a solid knowledge base and described the kinds of challenges associated with generalizability, interpretability, applicability, transferability, and effectiveness needed to advance the field (Baker 2019). For an education community continually doing more with less, these algorithmic insights can be a welcome information source or sixth sense. However, *is it ethical that the algorithms that inform teaching and learning may not fully represent the learners they are designed to educate? or that the algorithms absorb and reflect human biases during their design, development, and evolution when humans interact with them?*

Algorithms that seek to do things better by increasing efficiencies can embed existing biases and replicate existing conditions. Large social datasets feed systemic bias into algorithms, and unchecked algorithms can result in systemic discrimination that favors certain individuals or groups over others. Education administrators, for example, monitor the behavior of students and school employees to maintain safe learning environments. In a recent review of the empirical evidence, Nance (2019) documents that schools serving primarily white students in the U.S. are less reliant on coercive surveillance measures than schools serving higher concentrations of students of color, despite limited empirical evidence that such measures impact school safety. To traditional surveillance tools, school systems now utilize surveillance-based algorithms that continuously monitor students' social

media accounts, chat messages, email, and schoolwork to predict and then alert school officials to potentially harmful behavior (Gilman 2020). Contrary to desired outcomes, overreliance on extreme security measures in schools can increase drop-out rates and engender distrust and discord among members of the school community in the long term (Gilman 2020; Nance 2019).

Because educational issues are far more complex than a single algorithm can capture, educators and policymakers have a responsibility to avoid overly relying on any single algorithm to make important decisions (Daniel 2019). They also have a responsibility to understand the basics of algorithmic development and how designers mitigate bias (Kirkpatrick 2016; Shah et al. 2020). Although a machine-developed algorithm might be able to make accurate predictions, ethical questions arise about whether decision makers can or should trust a solutions that rely on black-box systems about which stakeholders have no information or insight. By asking critical questions regarding a digital tool or learning platform's underlying data sources, algorithm development, and ongoing testing, educational leaders can elicit the transparency from developers and vendors that students deserve. Three essential questions follow.

First, *how did the developer ensure that the data used to develop algorithms represent user diversity?* Developers need to document and address inclusivity, stakeholder awareness, and potential ethical risks during the design and testing of algorithms to ensure that all populations are represented and protected from harm (Mitchell et al. 2020; Yapo and Weiss 2018).

Second, *how did the developer protect against algorithmic bias in its digital tool or learning platform?* To date, algorithm development has largely occurred with minimal oversight or deep consideration of ethics and bias (Luckin et al. 2016). A recent survey of U.S. corporate executives revealed that although 9 in 10 respondents believe ethical standards in the development and use of emerging technologies can represent a competitive advantage for businesses, about 2 in 3 of those surveyed acknowledged existing bias in AI technologies used by their company (RELEX 2021). AI algorithms are not neutral entities; rather, they are theory-laden and reflect particular world views (Ferrero and Barujel 2019). For organizations that create digital tools and learning platforms, mitigating algorithmic bias begins with the staffing of the team responsible for developing and validating an algorithm. Mitigating bias requires a diversity of disciplines, research questions, life experiences, cultures, races, religions, ages, sexes, sexual orientations, and disabilities (Nielsen et al. 2018). To their work, the team individually and collectively must practice active reflexivity by reflecting on their beliefs, practices, and judgments during and after the research process, acknowledging how these may have influenced the research (Finlay 2016; Soedirgo and Glas

2020). When outside teams incorporate data that they did not collect directly, it can drastically diminish the value of their reflexivity and potentially compromise the validity of their research outcomes (Daniel 2019).

Third, *how does the developer continue to monitor the tool or platform for bias?* The threat of bias requires constant review and audits of equity, quality, and fairness (Educational Testing Service 2014; Shute et al. 2020), as well as policies and markets that incentivize iterative improvements in the accuracy, fairness, reliability, and accountability of these algorithms.

4 Upstream divides

Algorithmic bias and the divide it creates do not occur in isolation. Rather, distinct factors act as its fuel: the *access divide* stems from who does and does not have access to the hardware, software, and connectivity necessary to engage with digital learning tools and platforms; the *representation divide* occurs when data representative of populations or over-representative of a subpopulation's preferences prevents objectivity and influences understandings and outcomes. Both divides predetermine the data that algorithms include in their development, validation, and refinement.

5 The access divide

When employed by well-trained educators, researchers have demonstrated that digital tools and learning platforms have developed and enhanced the following for learners: (a) knowledge within academic disciplines; (b) cognitive skills, such as problem-solving, critical thinking, and systems thinking; (c) interpersonal skills, such as communications, social skills, teamwork, and cultural sensitivity; and (d) intrapersonal skills, such as self-management, time management, self-regulation, adaptability, and executive functioning (Clark et al. 2016; D'Angelo et al. 2014; Díaz et al. 2019; Fishman and Dede 2016; National Academies of Sciences 2000, 2020). Given that the benefits of educational technologies accrue to those learners and educators with regular access to the hardware, software, and connectivity to use digital tools and learning platforms regardless of their physical location, *is it thus ethical that all learners and educators do not have access to the hardware, software, and connectivity necessary to engage with digital tools and learning platforms?*

Universal and equitable access among learners and educators has not yet been realized, falling short of the goal of educational equity in which every learner has access to the resources they need irrespective of race, gender, ethnicity, language, disability, sexual orientation,

family background, or family income (Council of Chief State School Officers 2018). Chandra and Colleagues' (2020) analysis of 2018 American Community Survey data found that about 30% U.S. K–12 public school students and about 10% of K–12 teachers live in households either without an Internet connection or without a device adequate for distance learning at home. Additionally, about 17% of K–12 public school students live in households with neither an adequate connection nor an adequate device for distance learning at home.

Researchers at LearnPlatform (2021) conducted an analysis of daily U.S. K–12 student use of educational technologies used by 2.5 million students in 17 states from February through December 2020. Their analysis, represented in Fig. 3, illustrates the access and usage divide between more affluent districts (i.e., districts with up to 25% free and reduced-price lunch student populations) and less affluent districts (i.e., districts with 25–100% free and reduced-price lunch student populations). The y-axis is LearnPlatform's educational technology usage index, which is based on the number of visits to different tools per 1000 users. The index provides standardization across different user groups, while also adjusting for the breadth of different tools used. After school closures in spring of 2020 due to COVID-19, more affluent districts recovered quickly and increased engagement while less affluent districts did not return to pre-pandemic engagement levels until the fall of 2020.

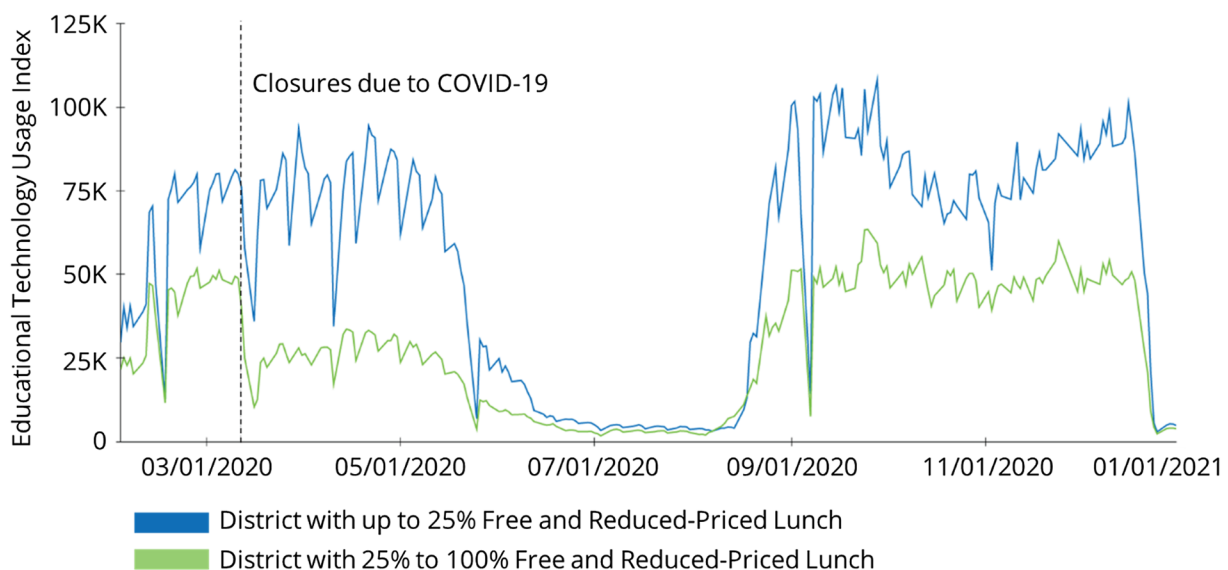


Fig. 3 Daily K–12 student usage of educational technologies from February through December 2020 by more affluent and less affluent U.S. School Districts. Note. LearnPlatform (2021) *EdTech Engage-*

6 The representation divide

One outcome of learners and educators engaging with digital tools and learning platforms is extraordinarily large amounts of micro-, meso-, and macro-data formats (Fischer et al. 2020) that vary by (a) *volume*, the size and scale of data and data sets; (b) *variety*, the production of data from different data sources and in different formats and grain sizes; (c) *velocity*, the speed at which data are created; (d) *veracity*, the noise, bias, and uncertainty in data; and (e) *value*, the administrative, instructional, monetary, and knowledge produced from analysis of data (based on European Economic and Social Committee 2017). The data contain not only summative evaluations of students' knowledge, ability, and skills, but also the processes of their learning and acquisition of relevant skills and knowledge (Ercikan and Pellegrino 2017).

When significant numbers of learners do not have access to the hardware, software, and connectivity necessary to access and engage with digital learning platforms, they are prevented from generating data used to develop and validate algorithms that inform instruction and other decisions. When those learners systematically come from vulnerable populations (e.g., rural learners, learners with special needs, learners in low-income families), the access creates a representation divide. This is an example of the “big data paradox,” a mathematical tendency of big datasets to minimize one type of error due to small sample size, but magnify another that tends to get less attention: flaws linked to systematic biases that make the sample a poor representation of the larger population (Powell 2021). *Is it ethical that a lack of learner*

ment & Digital Learning Equity Gaps. Reprinted with permission by K. Rectanus, September 7, 2021

and educator access means those with access contribute to data sets and are therefore considered in subsequent models, algorithms, and interpretations based on that data, whereas those without access do not contribute?

While no data set is perfect, sets with duplicated, outdated, incomplete, inaccurate, incorrect, inconsistent, or missing records have the potential to create data bias, which is the systemic distortion in data that compromise representativeness (Marco and Larkin 2000; Olteanu et al. 2019). The lack of representation in data plays out in otherwise well-intentioned technologies. For example, investigating the racial, skin type, and gender disparities embedded in commercially available facial recognition technologies, Buolamwini and Gebru (2018) revealed how those systems largely failed to differentiate and classify darker female faces while successfully differentiating and classifying white male faces. The poor classification for darker female faces stemmed from the data sets used to develop the algorithms, which included a disproportionality large number of white males and few Black females. When researchers used a more balanced data set to develop the algorithm, it produced more accurate results across races and genders.

7 Downstream divides

Downstream divides result from algorithms fed by data that disproportionately draw from those learners who have access and representation. The downstream consequence of an *interpretation divide* occurs when learners, educators, and others misinterpret the outputs of algorithms or rely on inherently faulty algorithmic outputs to make decisions. The consequence of a *citizenship divide* represents how the other divides accumulate to impact interpretations by learners, educators, and others, which then influence behaviors and, over time, skills, culture, economic, health, and civic outcomes.

8 The interpretation divide

When we reach the interpretation divide, educators and learners excluded through the access and data divides have not contributed to the data sets used to develop algorithms. As classroom instruction and education policy has increasingly relied on data-based decision-making (DBDM), educators, researchers, and policymakers cite data and its interpretation to justify and guide decisions at the student and classroom level, scaling up to entire populations and subpopulations of students in schools, districts, states, and countries (Schildkamp et al. 2013). Data generated when students and educators engage digital tools and learning platforms provide yet another data source to support

decision-making. However, *is it ethical for educators to use data and the outputs of algorithms to make decisions without having received training on how to interpret or use them?* Users' knowledge of DBDM can be (a) *missing*, they never learned how to interpret it in the first place; (b) *inert*, they know how to interpret the data but do not know when or how to apply it; (c) *routinized*, they apply an interpretation technique without thinking through whether it is the right technique for the given situation; (d) *surface*, they are familiar but not proficient with an interpretive technique; or (e) *perishable*, they knew something at one point but have lost it because they have not applied it recently (based on the fragile knowledge construct developed by Perkins 1992).

An additional risk to interpreting data from learning platforms and other sources comes from *confirmation bias*, which occurs when individuals or groups search for or interpret data in a way that confirms their experiences and preconceptions, leading to biased decisions. In their summary of research on confirmation bias, MacLean and Dror (2016) note that an inaccurate, initial understanding of a situation can compromise an individual's or group's attempts to reach correct decisions. They further note that individuals working alone or in groups tend to seek out and give greater weight to information consistent with their expectations, while ignoring, discrediting, or trivializing information that is inconsistent with their working theory.

To counteract the interpretation divide, Mandinach and Schildkamp (2021) make several recommendations that education decision makers can enact, including prioritizing and supporting data interpretation education for teachers. Additionally, they recommend that those analyzing data refrain from relying on a single data source, such as assessments or that generated from a learning platform, instead considering it along with other classroom data sources. Contextualizing the data is also critical: "Educators need data, such as demographics, attendance, health, transportation, justice, motivation, home circumstances (i.e., homelessness, foster care, potential abuse, poverty), and special designations (i.e., disability, language learners, bullying), to contextualize student performance and behavior" (p. 2).

Educational leaders can also support the design of thoughtful frameworks that take and make sense of data from a variety of sources. For example, Mandinach and Miskell (2018) studied the affordances of technologies used in blended learning environments and how they affected teaching and learning activities. The study used mixed methods to examine whether the blended learning environments provided enhanced access to and more diverse data for teachers and students from which to make educational decisions. The study found that the technologies provided more diverse data to administrators, teachers, and students and allowed for flexible adaptations to virtual and face-to-face learning to meet students' needs. The blended environments helped

to create data cultures within the schools where educators used data to communicate and have an impact on instructional activities.

To do things better, Schildkamp et al. (2019) identify five key building blocks for educational leaders wanting to cultivate effective teams that use data and mitigate confirmation bias in their school: (a) discussing and establishing a vision, norms, and goals of data use with educators; (b) meeting educators where they are by providing individualized technical and emotional support; (c) sharing knowledge across the team and providing autonomy; (d) creating a safe and productive climate within the team that focuses on data use for improvement rather than accountability; and (e) brokering knowledge and creating a network that is committed to data use.

To do better things with AI-enhanced learning technologies, educators need to capitalize on each learner's unique skills and interests to facilitate new learning. The Universal Design for Learning (UDL) Framework is a research-based example of how to plan for a culturally responsive, equity-focused learning environment to improve teaching and learning for all people (CAST 2018). UDL involves providing multiple means of engagement (the *why* of learning), representation (the *what* of learning), and action & expression (the *how* of learning). Using this framework, teachers can help students to access, build on, and internalize information on their way to becoming expert learners who are purposeful and motivated, resourceful and knowledgeable, and strategic and goal-oriented (Chardin and Novak 2020).

In adopting a UDL framework, teachers, students, and other stakeholders become partners in a democratic process in which every individual plays an active role and the perspectives of diverse groups are considered, fostering a culturally responsive learning environment (Chardin and Novak 2020). Culturally responsive teaching requires allowing students to draw upon their own culture to guide the curriculum while constantly reflecting on the teacher's cultural perspective to avoid allowing bias to unwittingly shape their instruction. Further, a culturally responsive perspective demands that stakeholders focus not only on who has access to information and who does not, but also on whose information is valued and whose is not. In the classroom, this entails giving every student a voice and a safe space to explore ideas and reasoning.

As an example, in a class exploring fractions for the first time, a teacher may ask a student to position a given fraction on a number line and justify their decision. A teacher who allows the student to explain their reasoning and to respond to questions from the class regardless whether the answer is correct or incorrect has acknowledged the value of that student's perspective and opened up the entire class to a possibly different way of approaching problems (Shepard 2019). Effective AI to do better things will be based around a similar democratic

process that involves shared power, considers diverse voices, and protects individual rights from powerful institutions (see Shohamy 2016).

9 The citizenship divide

Though many early researchers and scholars associate the digital divide with unequal access to technology hardware and the Internet, the disparity among the haves and have nots not only precludes learners and educators from accessing information and collaborating over distance and time but also diminishes their ability to accumulate social capital and prepare for success in a knowledge-based economy (Culp et al. 2005; Ritzhaupt et al. 2020; Valadez and Duran 2007). Across life outcomes, such as health, wages, and indicators of civic engagement and trust, adults with higher levels of literacy, numeracy, and 21st-century skills, as well as technology access, fare better than their counterparts with lower skill levels and less connectivity (Kirsch et al. 2021; Ramsetty and Adams 2020).

Add to this that education is not alone in its susceptibility to bias; examples of algorithmic bias also associated with life outcomes abound in the criminal justice system (Angwin et al., 2016; Završnik 2019), social services (Eubanks 2017), and job recruitment (Caliskan et al. 2017; Raghavan et al. 2020). There are also interaction effects among algorithms. For example, Hao (2020) illustrates the deleterious interactions between credit-reporting algorithms that affect access to private goods and services (e.g., obtaining a home or student loan, employment, and renting an apartment) and U.S. Government agency algorithms purchased from private vendors that affect access to public benefits and leave citizens without insights or recourse when algorithms generate mistakes. Individually, they disproportionately and systematically affect the poor, but together they have a greater, negative effect on life outcomes than either has on its own.

Is it ethical that without intervention, each cohort of learners is poised to perpetuate structural stigmas associated with access, achievement, identity, and power and thereby preserve the trends of the past, with some subgroups unduly benefiting and others not? As Zwitter (2014) warns, the more that the lives of learners and educators become mirrored in the data they generate through digital tools, learning platforms, and other media, the more their present, past, and future potentially become more transparent and predictable.

10 The cyclical effects of using artificial intelligence in education

Learner and educator engagement with digital platforms and other technologies that generate massive amounts of data could create a virtuous cycle by implementing new

mechanisms for advancing what is known about how people learn, developing new administrative procedures on how best to use scarce resources, discovering novel relationships and patterns, increasing the accuracy of predictions, and improving the automation of low-level tasks and decisions. Coupled with these opportunities for a virtuous cycle are technical, ethical, and logistical challenges that will grow more complex over time and which may lead to a vicious cycle of inequity.

As an illustration, the ability of AI to find patterns in rich datastreams and make predictions about what students do and do not know is central to *learning engineering*, an instructional design strategy that applies a principled set of evidence-based strategies to the continual re-design of educational experiences to optimize their effectiveness and efficiency (Dede et al. 2019). This in turn enables *personalized learning*, which requires four fundamental capabilities (Dede 2019):

1. Developing multimodal experiences and a differentiated curriculum based on universal design for learning principles;
2. Enabling each student's agency in orchestrating the emphasis and process of his or her learning, in concert with the evidence about how learning works best and with mentoring about working toward long-term goals;
3. Providing community and collaboration to aid students in learning by fostering engagement, a growth mindset, self-efficacy, and academic tenacity; and
4. Guiding each student's path through the curriculum based on diagnostic assessments embedded in each educational experience that are formative for further learning and instruction.

Substantial evidence shows that combining these four attributes, which AI helps to enable, leads to learning experiences that provide strong motivation and good educational outcomes for a broad spectrum of students.

Unwise use of AI-enhanced, educational technology for learning and teaching further magnifies the structural disparities already inherent in society. To take an ethical approach to addressing the divides described in the cycle and to increase opportunity and equity while mitigating bias, educators, designers, and policymakers must constantly reflect on the ethical questions raised in this article and a constellation of investments and actions. The following strategy is based on Dede (2015):

1. Empower communities of researchers, funders, policymakers, practitioners, and other stakeholders to use new forms of evidence to puzzle through and answer questions together. Investigations should focus on the purpose of the study, prior research, and research ques-

tions before selecting data and methods for answering those research questions.

2. Infuse evidence-based decision-making continuously with constant reflection on issues of cultural responsiveness, equity, quality, and fairness.
3. Use new and traditional forms of micro-, meso-, and macro-data to develop new ways of measuring learning and impact for formative, summative, and administrative purposes.
4. Work to reconceptualize how data are generated, collected, stored, accessed, analyzed, interpreted, and acted on by different categories of users (e.g., educators vs. research vs. policymaking) for different purposes.
5. Develop new types of analytic methods to enable rich findings from complex forms of educational data and new forms of visualizations to identify useful patterns in educational data that may not be obvious and help educators more easily navigate, interpret, and act on data (Daniel 2019).
6. Build human capacity through professional development and degree programs to better integrate pedagogical and ethical uses data science into the design, development, and use of digital learning platforms and digital tools.

Correspondingly, anticipating the promise and pitfalls of AI, entities have been set up to offer guidance to those developing AI (e.g., IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019; UNESCO 2019) and to monitor the development and deployment of AI tools (Digital Promise n.d.; European Commission n.d.). In the education community, the EdSAFE AI Alliance launched in 2021 as an international group of tech companies, educators, policymakers, and other stakeholders organized to collectively scrutinize the quality of new AI tools and techniques and inform future regulations ("Edsafe AI Alliance to Drive Healthy Ecosystem of AIEd Sector," 2021). Over the coming years, the alliance will define new safety, accountability, fairness, and efficacy (i.e., SAFE) industry benchmarks to address data security, reliability, and equity in learning.

11 Conclusion and future work

By addressing the ethical questions associated with each segment of the cycle and by addressing these recommendations, we can begin to realize the potential of a virtuous cycle: lifelong education and improved life outcomes. Much of this article is about using AI to make conventional educational practices more effective and efficient through analyzing large datasets to promote evidence-based decisions. *Doing things better* is a useful objective, but a more important aspect of AI-based analytics is the ways they can enable *doing better things*. In future, if

work, civic participation, and personal fulfillment are to be equitable and sustainable, today's industrial-era systems of schooling must change. Using AI, instructors can create, evaluate, and improve transformative lifewide, lifelong learning experiences that provide students with the sophisticated knowledge, skills, and dispositions they need in our evolving global digital civilization (Dede 2020). The future will be quite different than the immediate past: We and our children face a world-wide interdependent civilization shaped by economic turbulence from AI and globalization (Osei Bonsu and Song 2020), failure to reach the United Nations (2015) sustainability goals, global climate change (National Academies of Sciences 2020), and rapid shifts driven by world-wide mobile devices equipped with social media (e.g., Mozur 2018). We stand on the brink of an epic half-century, equivalent in its challenges and opportunities to those faced from 1910 to 1960: two world wars, a global pandemic, a long-lasting economic depression, and constant conflicts between capitalism and communism.

To fulfill their responsibilities in preparing learners of all ages for this turbulent, disruptive future, educators at every level are now faced with developing people's capacity for unceasing reinvention to face an uncertain and changing workplace, and for inventing and mastering occupations that do not yet exist. Students must develop personal dispositions for finding opportunity in uncertainty: creating new value, reconciling tensions and dilemmas, and assuming moral/ethical agency on equity and respect for diversity (Organisation for Economic Cooperation and Development 2018). Students must acquire knowledge and skills underemphasized in current curriculum standards and omitted from today's high-stakes summative tests: fluency of ideas, social perceptiveness, systems thinking, originality, and conflict resolution (Bakhshi et al. 2017). Advances in AI, which are accelerating this challenge, can also help us to meet it, if we develop ways to use AI wisely, ethically, and virtuously.

Data availability statement No data were used or analyzed for this manuscript.

Declarations

Conflict of interest The authors whose names are listed on the title page certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

- Adjei S, Ostrow K, Erickson, E, Heffernan N (2017) Clustering students in ASSISTments: exploring system and school-level traits to advance personalization. In: Proceedings of the 10th international conference on Educational Data Mining, Hubei Province, China
- Baker RS (2019) Challenges for the future of educational data mining: the Baker learning analytics prizes. *J Educ Data Mining* 11(1):1–17. <https://doi.org/10.5281/zenodo.3554745>
- Bakhshi H, Downing J, Osborne M, Schneider P (2017). The future of skills: employment in 2030. Pearson and Nesta. <https://futureskills.pearson.com/research/assets/pdfs/technical-report.pdf>
- Buolamwini J, Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Proceedings of the 1st conference on fairness, accountability and transparency, New York, NY. <http://proceedings.mlr.press>
- CAST (2018) Universal design for learning guidelines. Version 2.2. <https://udlguidelines.cast.org/>
- Caliskan A, Bryson JJ, Narayanan A (2017) Semantics derived automatically from language corpora contain human-like biases. *Science* 356(6334):183–186. <https://doi.org/10.1126/science.aal4230>
- Chandra S, Chang A, Day L, Fazlullah A, Liu J, McBride L, Mudalige T, Weiss D (2020) Closing the K-12 digital divide in the age of distance learning. https://www.common sense media.org/sites/default/files/uploads/pdfs/common_sense_media_report_final_7_1_3pm_web.pdf
- Chardin M, Novak K (2020) Equity by design: Delivering on the power and promise of UDL. Corwin Press, California
- Clark DB, Tanner-Smith EE, Killingsworth SS (2016) Digital games, design, and learning: a systematic review and meta-analysis. *Rev Educ Res* 86(1):79–122. <https://doi.org/10.3102/0034654315582065>
- Council of Chief State School Officers (2018) States leading for equity: promising practices advancing the equity commitments. https://ccsso.org/sites/default/files/2018-03/States%20Leading%20for%20Equity%20Online_031418.pdf
- Culp KM, Honey M, Mandinach E (2005) A retrospective on twenty years of education technology policy. *J Educ Comput Res* 32(3):279–307. <https://doi.org/10.2190/7w71-qvt2-pap2-udx7>
- D'Angelo C, Rutstein D, Harris C, Bernard R, Borokhovski E, Haertel G (2014) Simulations for STEM learning: systematic review and meta-analysis. SRI International. <https://www.sri.com/wp-content/uploads/pdf/simulations-for-stem-learning-full-report.pdf>
- Daniel BK (2019) Big data and data science: a critical review of issues for educational research. *Br J Edu Technol* 50(1):101–113. <https://doi.org/10.1111/bjet.12595>
- Decuyper M, Grimaldi E, Landri P (2021) Introduction: critical studies of digital education platforms. *Crit Stud Educ* 62(1):1–16. <https://doi.org/10.1080/17508487.2020.1866050>
- Dede C (2015) Data-intensive research in education: current work and next steps. Computing Research Association. <https://cra.org/wp-content/uploads/2015/10/CRAEducationReport2015.pdf>
- Dede C (2019) Improving efficiency and effectiveness through learning engineering. In: Dede C, Richards J, Saxberg B (eds) Learning engineering for online education: theoretical contexts and design-based examples. Routledge, UK, pp 1–14
- Dede C (2020) Introduction: Reconceptualizing higher education and lifelong learning in the era of the synergistic digital economy. In: Dede C, Richards J (eds) The 60-year curriculum: new models for lifelong learning in the digital economy. Routledge, UK, pp 1–24

- Dede C, Richards J, Saxberg B (eds) (2019) Learning engineering for online education: theoretical contexts and design-based examples. Routledge, UK. <https://doi.org/10.4324/9781351186193>
- Díaz P, Ioannou A, Bhagat KK, Spector JM (eds) (2019) Learning in a digital world: perspective on interactive technologies for formal and informal education. Springer, Berlin. <https://doi.org/10.1007/978-981-13-8265-9>
- Digital Promise (n.d.) Product certifications. Retrieved December 17, 2021. <https://productcertifications.digitalpromise.org/>
- D’Mello S, Dieterle E, Duckworth A (2017) Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educ Psychol* 52(2):104–123. <https://doi.org/10.1080/00461520.2017.1281747>
- Educational Testing Service. (2014) ETS standards for quality and fairness. <https://www.ets.org/s/about/pdf/standards.pdf>
- Ercikan K, Pellegrino JW (eds) (2017) Validation of score meaning for the next generation of assessments: the use of response processes. Routledge, UK
- Eubanks V (2017) Automating in equality: How high-tech tools profile, police, and punish the poor. St. Martin’s Press
- European Commission. (n.d.). AI watch: monitor the development, uptake and impact of Artificial Intelligence for Europe. Retrieved July 25, 2021. https://knowledge4policy.ec.europa.eu/ai-watch_en
- European Economic and Social Committee (2017) The ethics of big data: balancing economic benefits and ethical questions of big data in the EU policy context. European Union. <https://doi.org/10.2864/18937>
- Ferrero F, Barujel AG (2019) Algorithmic driven decision-making systems in education: analyzing bias from the sociocultural perspective. In: 2019 XIV Latin American Conference on Learning Technologies (LACLO), <https://doi.org/10.1109/LACLO49268.2019.00038>
- Finlay L (2016) Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qual Res* 2(2):209–230. <https://doi.org/10.1177/1468794102002002005>
- Fischer C, Pardos ZA, Baker RS, Williams JJ, Smyth P, Yu R, Slater S, Baker R, Warschauer M (2020) Mining big data in education: affordances and challenges. *Rev Res Educ* 44(1):130–160. <https://doi.org/10.3102/0091732x20903304>
- Fishman B, Dede C (2016) Teaching and technology: new tools for new times. In: Gitomer DH, Bell CA (eds) Handbook of research on teaching, 5th edn. American Educational Research Association, Washington, pp 1269–1334
- Giannakos MN, Sharma K, Pappas IO, Kostakos V, Velloso E (2019) Multimodal data as a means to understand the learning experience. *Int J Inf Manage* 48:108–119. <https://doi.org/10.1016/j.ijinfomgt.2019.02.003>
- Gilman M (2020) Poverty lawgorithms: a poverty lawyer’s guide to fighting automated decision-making harms on low-income communities. <https://ssrn.com/abstract=3699650>
- Guo H, Rios JA, Haberman S, Liu OL, Wang J, Paek I (2016) A new procedure for detection of students’ rapid guessing responses using response time. *Appl Meas Educ* 29(3):173–183. <https://doi.org/10.1080/08957347.2016.1171766>
- Hao K (2020) The coming war on the hidden algorithms that trap people in poverty. <https://www.technologyreview.com/2020/12/04/1013068/algorithms-create-a-poverty-trap-lawyers-fight-back/>
- Heffernan NT, Heffernan CL (2014) The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *Int J Artif Intell Educ* 24(4):470–497. <https://doi.org/10.1007/s40593-014-0024-x>
- IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (2019) Ethically aligned design: a vision for prioritizing human well-being with autonomous and intelligent systems, 1st edn. IEEE. <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf>
- Jiang W, Pardos ZA, Wei Q (2019) Goal-based course recommendation. In: Proceedings of the 9th international conference on learning analytics & knowledge, Tempe, AZ. <https://doi.org/10.1145/3303772.3303814>
- Kirkpatrick K (2016) Battling algorithmic bias. *Commun ACM* 59(10):16–17. <https://doi.org/10.1145/2983270>
- Kirsch I, Sands A, Robbins S, Goodman M, Tannenbaum R (2021) Buttressing the middle: a case for reskilling and upskilling America’s middle-skill workers in the 21st century. <https://www.ets.org/s/research/pdf/buttressing-policy-report.pdf>
- Kostyuk V, Almeda MV, Baker RS (2018) Correlating affect and behavior in reasoning mind with state test achievement. In: Proceedings of the 8th international conference on learning analytics and knowledge, Sydney, Australia
- LearnPlatform (2021) Edtech engagement & digital learning equity gaps. Retrieved July 14, 2021. <https://learnplatform.com/insights/infographic/2020-digital-equity-in-review>
- Lee Y-H, Haberman SJ (2016) Investigating test-taking behaviors using timing and process data. *Int J Test* 16(3):240–267. <https://doi.org/10.1080/15305058.2015.1085385>
- Levy R (2019) Dynamic Bayesian network modeling of game-based diagnostic assessments. *Multivar Behav Res* 54(6):771–794. <https://doi.org/10.1080/00273171.2019.1590794>
- Luckin R, Holmes W, Griffiths M, Forcier LB (2016) Intelligence unleashed: an argument for AI in education. Pearson. <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence-Unleashed-v15-Web.pdf>
- MacLean CL, Dror IE (2016) A primer on the psychology of cognitive bias. In: Robertson CT, Kesselheim AS (eds) Blinding as a solution to bias: strengthening biomedical science, forensic science, and law. Academic Press, Cambridge, pp 13–24
- Mandinach EB, Miskell RC (2018) Blended learning and data use in three technology-infused charter schools. *Learn Landsc* 11(1):183–198. <https://doi.org/10.36510/learnland.v11i1.931>
- Mandinach EB, Schildkamp K (2021) Misconceptions about data-based decision making in education: an exploration of the literature. *Stud Educ Eval* 69:1–10. <https://doi.org/10.1016/j.stueduc.2020.100842>
- Marco CA, Larkin GL (2000) Research ethics: ethical issues of data reporting and the quest for authenticity. *Acad Emerg Med* 7(6):691–694. <https://doi.org/10.1111/j.1553-2712.2000.tb02049.x>
- Mitchell M, Baker D, Moorosi N, Denton E, Hutchinson B, Hanna A, Gebru T, Morgenstern J (2020) Diversity and inclusion metrics in subset selection. In: 2020 AAAI/ACM conference on AI, ethics, and society (AIES ’20), New York, NY
- Mozur P (2018) A genocide incited on Facebook, with posts from Myanmar’s military. *The New York Times*. <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>
- Nance JP (2019) Implicit racial bias and students’ fourth amendment rights. University of Florida Levin College of Law Research Paper No. 19-26. <https://ssrn.com/abstract=3140638>
- National Academies of Sciences, Engineering, and Medicine (2000). How people learn: brain, mind, experience, and school (expanded ed.). National Academy Press. <https://doi.org/10.17226/9853>
- National Academies of Sciences, Engineering, and Medicine (2018) How people learn II: learners, contexts, and cultures. The National Academy Press, Washington. <https://doi.org/10.17226/24783>
- National Academies of Sciences, Engineering, and Medicine (2020) Climate change: evidence and causes Update 2020. The National Academies Press, Washington. <https://doi.org/10.17226/25733>

- Nielsen MW, Bloch CW, Schiebinger L (2018) Making gender diversity work for scientific discovery and innovation. *Nat Hum Behav* 2(10):726–734. <https://doi.org/10.1038/s41562-018-0433-1>
- Niemi DM, Pea RD, Saxburg B, Clark RE (eds) (2018) *Learning analytics in education*. Information Age Publishing Inc., USA
- Olteanu A, Castillo C, Diaz F, Kiciman E (2019) Social data: biases, methodological pitfalls, and ethical boundaries. *Front Big Data*. <https://doi.org/10.3389/fdata.2019.00013>
- Organisation for Economic Co-operation and Development (2018) The future of education and skills: education 2030. <http://www.oecd.org/education/2030/oecd-education-2030-position-paper.pdf>
- Osei Bonsu K, Song J (2020) Turbulence on the global economy influenced by artificial intelligence and foreign policy inefficiencies. *J Liberty Int Affairs* 6(2):113–122. <https://doi.org/10.47305/JLIA2020113ob>
- Pardos ZA, Baker RS, San Pedro M, Gowda SM, Gowda SM (2014) Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *J Learn Anal* 1(1):107–128. <https://doi.org/10.18608/jla.2014.11.6>
- Perkins D (1992) *Smart schools: better thinking and learning for every child*. Free Press, New York
- Powell A (2021) Two early vaccination surveys worse than worthless thanks to ‘big data paradox,’ analysts say. <https://news.harvard.edu/gazette/story/2021/12/vaccination-surveys-fell-victim-to-big-data-paradox-harvard-researchers-say/>
- Raghavan M, Barocas S, Kleinberg J, Levy K (2020) Mitigating bias in algorithmic hiring: Evaluating claims and practices. *Proceedings of the 2020 conference on fairness, accountability, and transparency*
- Ramsetty A, Adams C (2020) Impact of the digital divide in the age of COVID-19. *J Am Med Inform Assoc* 27(7):1147–1148. <https://doi.org/10.1093/jamia/ocaa078>
- RELX (2021) RELX emerging tech executive report 2021: accelerated adoption of AI brings new obstacles to light for US business leaders. Retrieved December 7, 2021. <https://stories.relx.com/relx-emerging-tech-2021/index.html>
- Ritzhaupt AD, Cheng L, Luo W, Hohlfield TN (2020) The digital divide in formal educational settings: The past, present, and future relevance. In: Bishop MJ, Boling E, Elen J, Svihla V (eds) *Handbook of research in educational communications and technology*, 5th edn. Springer, New York, pp 483–504. https://doi.org/10.1007/978-3-030-36119-8_23
- Sao Pedro MA, Baker RS, Gobert JD, Montalvo O, Nakama A (2013) Leveraging machine-learned detectors of systematic inquiry behavior to estimate and predict transfer of inquiry skill. *User Model User-Adap Inter* 23(1):1–39. <https://doi.org/10.1007/s11257-011-9101-0>
- Schiff D (2021) Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI Soc* 36(1):331–348. <https://doi.org/10.1007/s00146-020-01033-8>
- Schildkamp K, Lai MK, Earl LM (eds) (2013) *Data-based decision making in education: challenges and opportunities*. Springer, New York
- Schildkamp K, Poortman CL, Ebbeler J, Pieters JM (2019) How school leaders can build effective data teams: five building blocks for a new wave of data-informed decision making. *J Educ Change* 20(3):283–325. <https://doi.org/10.1007/s10833-019-09345-3>
- Shah DS, Schwartz HA, Hovy D (2020) Predictive biases in natural language processing models: a conceptual framework and overview. In: *Proceedings of the 58th annual meeting of the association for computational linguistics*
- Shao E, Guo S, Pardos ZA (2021) Degree planning with PLAN-BERT: multi-semester recommendation using future courses of interest. *Proc AAAI Conf Artif Intell* 35(17):14920–14929
- Shepard LA (2019) Classroom assessment to support teaching and learning. *Ann Am Acad Pol Soc Sci* 683(1):183–200. <https://doi.org/10.1177/0002716219843818>
- Shohamy E (2016) Democratic assessment as an alternative. *Lang Test* 18(4):373–391. <https://doi.org/10.1177/026553220101800404>
- Shute VJ, Ventura M (2013) *Stealth assessment: measuring and supporting learning in video games*. The MIT Press, Cambridge
- Shute V, Wang L (2017) Assessing and supporting hard-to-measure constructs in video games. In: Rupp AA, Leighton JP (eds) *The handbook of cognition and assessment: frameworks, methodologies, and applications*. John Wiley & Sons, New Jersey, pp 535–562
- Shute V, Rahimi S, Smith G, Ke F, Almond R, Dai CP, Kuba R, Liu Z, Yang X, Sun C (2020) Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *J Comput Assist Learn*. <https://doi.org/10.1111/jcal.12473>
- Soedirgo J, Glas A (2020) Toward active reflexivity: positionality and practice in the production of knowledge. *PS Polit Sci Polit* 53(3):527–531. <https://doi.org/10.1017/s1049096519002233>
- Stadler M, Fischer F, Greiff S (2019) Taking a closer look: an exploratory analysis of successful and unsuccessful strategy use in complex problems. *Front Psychol*. <https://doi.org/10.3389/fpsyg.2019.00777>
- UNESCO (2019) *Artificial intelligence in education*. Retrieved July 26, 2021. <https://en.unesco.org/artificial-intelligence/education>
- United Nations. (2015). *Transforming our world: the 2030 agenda for sustainable development*. <https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>
- Valadez JR, Duran R (2007) Redefining the digital divide: beyond access to computers and the internet. *High School J* 90(3):31–44
- van Rijn PW, Ali US (2017) A comparison of item response models for accuracy and speed of item responses with applications to adaptive testing. *Br J Math Stat Psychol* 70(2):317–345. <https://doi.org/10.1111/bmsp.12101>
- Wang Y, Kai S, Baker RS (2020) Early detection of wheel-spinning in ASSISTments. In: Bittencourt II, Cukurova M, Muldner K, Luckin R, Millán E (eds) *Artificial intelligence in education. AIED 2020. Lecture notes in computer science*, vol 12163. Springer International Publishing, New York, pp 574–585
- Yapo A, Weiss J (2018) Ethical implications of bias in machine learning. In: *Hawaii international conference on system sciences 2018*, Waikoloa Village, HI
- Završnik A (2019) Algorithmic justice: Algorithms and big data in criminal justice settings. *Eur J Criminol*. <https://doi.org/10.1177/1477370819876762>
- Zwitter A (2014) Big data ethics. *Big Data Soc* 1(2):1–6. <https://doi.org/10.1177/2053951714559253>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.