23

Inclusion and Equity as a Paradigm Shift for Artificial Intelligence in Education

Rod D. Roscoe, Shima Salehi, Nia Dowell, Marcelo Worsley, Chris Piech, and Rose Luckin

CONTENTS

23.1 Inclusion and Equity as a Paradigm Shift for Artificial Intelligence in Education

Artificial intelligence (AI) methods allow researchers and educators to assess complex patterns among diverse variables (e.g., learner backgrounds, behaviors, outcomes, learning context, and outcomes) to generate inferences and predictions for supporting learners and teachers (Baker & Inventado, 2014; Roll & Wylie, 2016). For example, various intelligent tutoring systems (ITSs, e.g., Kulik & Fletcher, 2016; Nye, 2015) have modeled learners based on factors such as task performance, behaviors, interaction patterns (Baker et al., 2010), natural language (Nye et al., 2014), and signals of affect (D'Mello et al., 2009; D'Mello & Graesser, 2012) and then used these AI-driven models to guide instruction and feedback. Similar approaches have contributed to automated scoring (Yan et al., 2020) and writing evaluation (AWE) technologies (McNamara et al., 2015; Shermis et al., 2016; Wilson & Roscoe, 2020), game-based learning and assessment (Shute, 2011; Shute et al., 2021), and social and collaborative learning (Schneider et al., 2021; Walker & Ogan, 2016). Collectively, the applications of AI in education (AIED) have enabled broad classes of adaptive and personalized educational technologies that facilitate students' learning.

As a paradigm shift, AI and AIED experts are increasingly attending to questions of diversity, inclusion, equity, ethics, belonging, and justice within their efforts (Blanchard, 2015; Holmes et al., 2021; Joyce et al., 2021). For brevity, we collectively and inclusively refer to these sweeping issues using the acronym 'DEI' (i.e., diversity, equity, and inclusion). Although AI applications for education are powerful and beneficial,

DOI: 10.1201/9781003181187-28

their development and implementation may have neglected social and societal factors related to bias, injustice, and how learners' identities and experiences affect their learning processes and environments. There is growing awareness of algorithmic bias, such that algorithms and automated systems can recreate or exacerbate discriminatory or oppressive outcomes. For instance, recent research has investigated biases in devices that use biometric sensors and measures (Drozdowski et al., 2020), algorithms used to inform decisions about criminal sentencing and recidivism (Miron et al., 2020; Wisser, 2019), and algorithms for guiding diagnosis and treatment in healthcare (Obermeyer et al., 2019; Panch et al., 2019; Walsh et al., 2020; Wien et al., 2020). Given that AI-driven technologies are increasingly ubiquitous in such everyday (e.g., phones and cars) and high-stakes (e.g., criminal justice and medicine) environments, there are substantial dangers associated with tools that do not work correctly, safely, or fairly for certain groups of people.

Educational contexts and technologies are not immune from bias. For example, in writing instruction and assessment, human ratings of writing can be biased based on presentation, dialect, content, perceived errors, and other aspects of linguistic diversity (Canz et al., 2020; Hammond, 2019; Johnson & VanBrackle, 2012; Johnson et al., 2017; Reaser et al., 2017). Although computer-based writing assessment may be perceived as 'objective' or 'fair' (i.e., the algorithms don't have feelings or personal agendas), this perspective ignores that algorithms are typically trained based on human annotations and ratings. Biases in training data are not automatically removed when developing computational algorithms. On the contrary, biases may become reified, reinforced, and even harder to inspect (Mayfield et al., 2019). A related challenge is training data that are sourced from exclusive or non-representative samples (see Roscoe, 2021) and thus fail to capture the true range or variation in student writers. Algorithms derived from limited samples can only be validly accurate or predictive within those limited samples. On a broader scale, there is increasing awareness that conclusions based on statistical averages can be misleading or exclusionary for learners who do not conform to 'average' or majority demographics (Rose, 2016).

To address these and related concerns, AIED scholars must carefully consider DEI challenges and alternative approaches to studying educational phenomena, analyzing data, and drawing meaningful educational conclusions without biases against a particular group(s). For instance, models may need to be disaggregated to include more nuanced variables and effects related to demographic factors and social identities (Kauh et al., 2021; Else-Quest & Hyde, 2016; Nichols & Stahl, 2019). Simultaneously, intersectional approaches (see Bauer et al., 2021; Bowleg, 2008; Cole, 2009; McKay et al., 2018; Rosenthal, 2016) are needed to represent learners' multiple identities (and associated power, privilege, and history) and to interpret these effects in findings and models.

Fortunately, AI methods have significant potential for investigating complex relationships among variables (Kizilcec et al., 2013; Piech et al., 2012) and characterizing learners at differing scales (e.g., from districts to individuals) with accuracy (Wang et al., 2021) if we are cautiously and mindfully inclusive throughout all stages. Therefore, AI methods can enhance DEI efforts in education through their power to carefully identify learners and their learning progression and needs. Hence, this paradigm shift in AIED is poised to empower personalized and effective educational outcomes for a much greater diversity of learners.

This chapter will discuss the *bidirectional relationship* between AI methods and DEI approaches. DEI approaches offer a valuable and necessary lens for conceptualizing, implementing, and interpreting AI while avoiding unintended but consequential biases. Synergistically, AI approaches and methods offer valuable ways for exploring complex data and nuanced relations – to enhance DEI in education. Together, this bidirectional relationship represents an important 'paradigm shift' for AIED as a field.

23.2 AI and DEI: A Bidirectional Relationship

The AI and DEI relationship is *bidirectional*: the analytical power of AI can enhance the DEI research through closer examination of learners, learning contexts, and learning outcomes (AI for DEI), and DEI lenses are necessary to improve AI approaches and avoid bias (DEI for AI).

23.2.1 AI for DEI: How Can the Principles and Methods of Artificial Intelligence Support Diversity, Equity, and Inclusion?

AI principles and methods can deepen our understanding of DEI and empower us to design and test interventions that address DEI issues and challenges. In the following, we present example research strands in which AI approaches have enhanced studying DEI challenges and generated new insights for interventions that improve DEI in learning environments.

AU: Please confirm if the shortened running head is fine.

23.2.2 Inclusivity in STEM Introductory Courses

23.2.2.1 Performance in Introductory STEM Courses

Women and underrepresented racial minority (URM) students remain underrepresented in STEM majors and careers (Dasgupta & Stout, 2014). The relative scarcity of women and URM students entering and persisting in STEM majors constrains opportunities to access high-demand STEM jobs and socioeconomic mobility. Performance in introductory STEM courses is also a key predictor and target for retention (Seymour & Hunter, 2019). However, it has been repeatedly observed that historically marginalized learners seem to be disadvantaged by such courses and underperform compared to peers from more privileged backgrounds (Chen, 2013). However, most of these analyses have been limited to descriptive statistics and comparisons of aggregate performance measures across different demographic groups. This approach arguably leads to further stigmatization of marginalized groups rather than insights about how to help. To address this challenge, Salehi and collaborators have employed large and longitudinal data sets across vastly different institutions to move beyond descriptive analysis to examine factors that impact the performance of students in large introductory physics courses (Salehi et al., 2019a, 2020).

The authors employed more nuanced quantitative approaches to discover that although marginalized groups (e.g. women, first-generation, and URM students) received lower grades in this physics course compared to their peers, almost all of these apparent performance gaps could be explained by variations in incoming STEM preparation (see Figure 23.1). Marginalized students received lower grades across all three institutions. However, when researchers controlled for incoming preparation, these performance gaps were no longer statistically significant – marginalized and non-marginalized students with the same level of STEM incoming preparation *performed the same* in their introductory physics courses.

Unfortunately, due to inequities in the United States' societal structure, along with systems of neglect and discrimination, marginalized students tend to receive reduced incoming preparation for STEM courses. Such students are more likely to attend under-resourced high schools and thus receive fewer opportunities for STEM exposure and preparation (Fahle et al., 2020). Moreover, typical STEM introductory courses often ignore variations in STEM incoming preparation and are tailored mostly to well-prepared students. Thus, underserved students who attend college and aspire to STEM careers may continue to be underserved, and perhaps underperform, compared to their peers.

23.2.2.2 Active Learning in Introductory STEM Courses

Previous analyses suggested that students from demographically marginalized groups in STEM (e.g., first generation, URMs, and women) were likely underserved with regard to STEM preparation (i.e., an indicator of inequity in social structures and educational infrastructure). Given that such preparation is an important predictor of performance in introductory STEM courses, and the importance of these courses

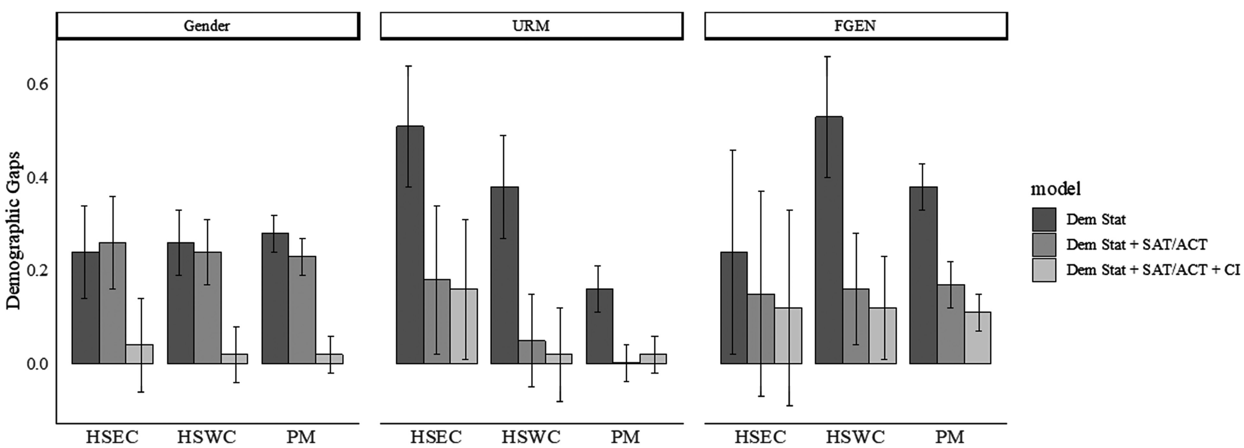

FIGURE 23.1
Size of coefficients for demographic variables in regression models that predict course performance at each institution. In each plot, the left-most bar (darkest) indicates the coefficient when *only the indicated demographic variable* (e.g., URM identity, first-generation status, or gender) is included. The central bar indicates the demographic coefficient when *math SAT or ACT scores are added predictors*. The right-most bar (lightest) indicates the coefficient when *concept inventory (CI) pretest scores are added as a predictor along with math SAT or ACT scores*. Error bars represent the standard error of the coefficients. Regression models that include only demographic status have *R*-squared values of 0.03 or less, but these increase to 0.2–0.3 when measures of incoming preparation are added to the model.

for retention, we need to identify equitable instructional practices that provide students with divergent levels of STEM preparation an equal opportunity to excel.

Introductory college STEM courses remain primarily lecture-based and grounded in foundational knowledge, and thus students' performance depends heavily on pre-college STEM preparation (Salehi et al., 2019a, 2020). This scenario further reinforces inequities in learning opportunities that result from disparities in STEM preparation quality and opportunities (Card & Rothstein, 2007; Fahle et al., 2020; Reardon & Owens, 2014). Furthermore, the effectiveness of lecture-based instruction has been challenged for *all* students. Previous meta-analyses have found that replacing lecture-based teaching with interactive, learner-centered instruction broadly improves the average performance of all students (Freeman et al., 2014; Haak et al., 2011). One caveat, however, is that many prior studies have relied on aggregate measures of students' performance that ignore demographic categories or variance. Thus, it remains unclear whether 'active learning' methods are effective in creating more equitable learning environments or addressing disparities. In a meta-analysis of more than 250 studies on the effects of active learning on academic performance, Theobald et al. (2020) found that only 15 studies reported results disaggregated by demographic group. Encouragingly, those few studies seemed to show that active learning instructional approaches particularly benefit marginalized students in STEM.

There are additional shortcomings in this literature (Theobald et al., 2020). First, although there is a disaggregation of performance outcomes across demographic groups, operationalization of instructional practices remains fairly coarse. Specifically, instructional approaches are broadly labeled as 'lecture-based' versus 'learner-centered' or 'active learning', which do not necessarily specify *how* or *which* instructional components of active learning methods actually benefit marginalized students. Consequently, it is not clear how to implement active learning (i.e., specific methods or activities) in ways that create an authentically inclusive environment.

To explore how active learning benefits marginalized students, Ballen et al. (2017) conducted structural equation modeling analyses using a large data set to explore potential mediating variables. The researchers found that active learning particularly benefits marginalized students by improving a sense of science self-efficacy, which in turn improves course performance. To further explore the specific components of active learning that benefit marginalized students, Ballen et al. (2017) also examined gender disparities across assessment methods in an introductory biology course. High-stakes exams were the most prone to gender disparities due to a disproportionately negative influence of test anxiety on performance of women. The researchers later replicated these results across a larger data set from 15 introductory STEM courses (Salehi et al., 2019b). These findings suggest that active learning improves equity in STEM courses through less reliance on inequitable high-stakes exam assessments.

Another important instructional component of active learning is frequent group activities. In the following sections we discuss how AI approaches help us better understand barriers and interventions for equitable collaborative learning.

23.2.3 Collaboration and Discourse

Collaboration is an essential aspect of learning, research, and modern work. This is particularly true in STEM where team science is responsible for highly impactful discoveries and multidisciplinary team research is the future of solving complex problems. Both educational and professional contexts require bringing together people with varying expertise to share knowledge, learn from each other, solve problems, and create products and ideas. Traditionally, teams have collaborated face-to-face, but contemporary collaboration increasingly occurs via virtual (i.e., online) platforms. Although digitally mediated collaborative problem-solving (CPS) environments hold the potential for creating more equitable and inclusive peer interactions, they are typically not characterized as such (Dasgupta et al., 2015; Du et al., 2015; Huang et al., 2014; Ke & Kwak, 2013). During technology-mediated STEM interactions, women and URMs face unique barriers such as feeling unwelcome to participate, having limited opportunities to contribute when conversations are dominated by a few members, and lacking perceived interpersonal power when attempts to engage are ignored. In each of these circumstances, complex and unique challenges are presented that can result in a detrimental impact on students' sense of belonging in the STEM milieu (Eddy et al., 2015).

To address such challenges, Dowell and colleagues have developed *Group Communication Analysis* (GCA), an innovative artificial intelligence-based methodology for quantifying and characterizing the discourse dynamics between learners in online multi-party interactions (Dowell et al., 2020; Dowell, Nixon, et al., 2019; Dowell & Poquet, 2021; Schneider et al., 2021). GCA applies automated computational linguistic analysis to the sequential interactions of participants in online group communication. GCA captures the structure of the group discussion and quantifies the complex semantic cohesion relationships between learners' contributions as they unfold over time, revealing intra- and interpersonal processes in group communication.

Dowell has used GCA and related NLP methods to study communication dynamics in online team interactions across gender and race to understand inclusivity in collaborative problem-solving (Dowell, 2019; Dowell, Lin, et al., 2019; Dowell et al., 2021; Lin et al., 2019; Lin & Dowell, 2019; Lin et al., 2020). Across several studies, Dowell discovered substantial intra- and interpersonal differences in women and URM's engagement that could influence their sense of belonging in online STEM environments. For example (see Figure 23.2), the difference between women and men in online STEM teams was not in how often they spoke. Instead, differences were evident in the extent to which they engage in productive discourse that responded to what other learners said previously (overall responsivity), provided meaningful contributions that warranted follow-up by peers (social impact), and monitored and built on their own previous contributions over the course of interaction (internal cohesion). Women's conversations showed greater overall responsivity, social impact, and internal cohesion than men's. In another study, Dowell, Lin, et al. (2019) examined how variations in team gender composition (female-minority, sex-parity, and female-majority) impacted socio-cognitive conversation patterns among team members using GCA. Results showed that the behavioral impact of men-dominated teams was more specific than simply gender differences in speaking up. Both men and women engaged in less productive collaborative problem-solving behaviors in men-dominated teams.

Across these illustrative examples and other studies, Dowell's team has revealed substantial intra- and interpersonal communication differences between women and men during CPS interactions. Moving forward, Dowell's team will be directing their efforts towards two important issues: (a) documenting the implications of observed differences for students' learning and psychological experience (e.g., sense of belonging, self-efficacy, propensity to remain in STEM majors) in teams, and (b) how to build sensitive, real-time feedback systems to best mitigate the detrimental impacts of certain team dynamics for marginalized populations.

23.2.4 Learning Assessments

Finally, an important consideration for research that centers on marginalized identities is how we assess and evaluate those learning experiences. This concern has been one focal aspect of the CrossMMLA (Multimodal Learning Analytics Across Spaces)

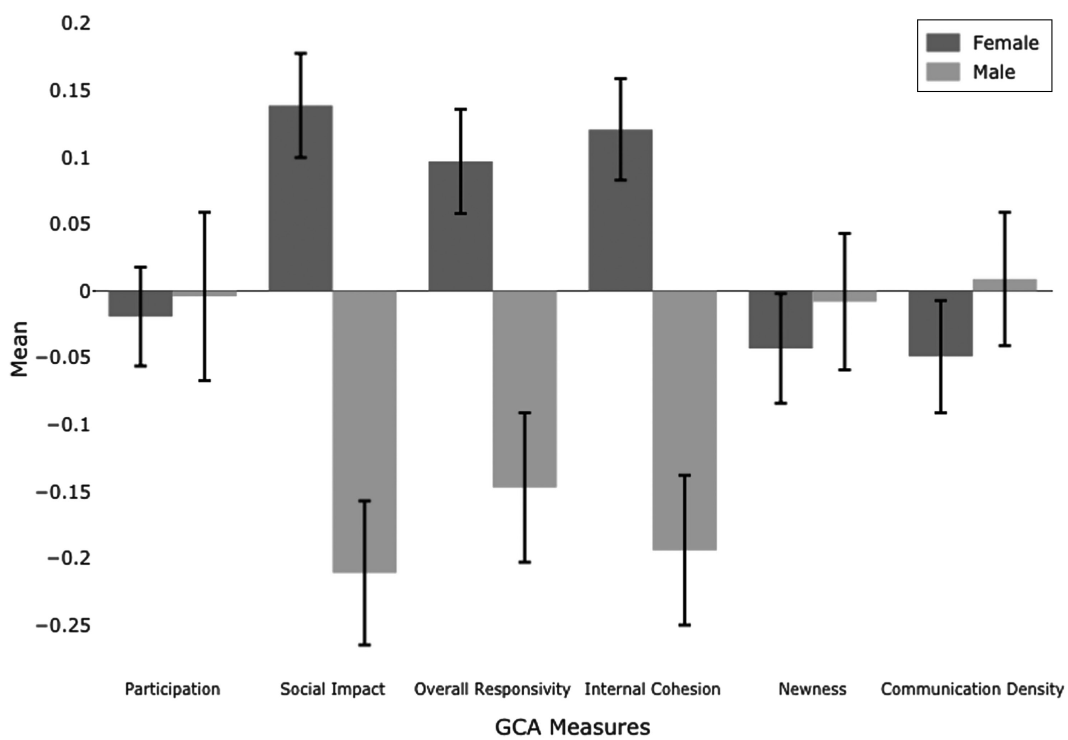

FIGURE 23.2
Collaborative group behaviors across participant's gender.

special interest group within SOLAR (Society of Learning Analytics Research). CrossMMLA highlights two important aspects of supporting and evaluating learning. First, learning takes place across innumerable locations. Although we often compartmentalize learning as primarily taking place within physical school buildings, we know that learning extends beyond school and other organized learning spaces. CrossMMLA thus emphasizes the need to develop systems and analytic approaches that can intelligibly chronicle learning as it unfolds across a variety of spaces (Blikstein & Worsley, 2016; Spikol et al., 2021). Doing so means that we open the door for learners to practice and engage in meaningful learning outside of the confines of schools, and potentially with a broader set of learning partners.

CrossMMLA also emphasizes that learners may demonstrate proficiency or knowledge growth in many ways. Multimodal sensors and analyses provide additional ways to surface student learning. This becomes increasingly imperative as we work to support learning across different spaces and use learning analytics (e.g., Sports Sense, Jones et al., 2020) that might not align with traditional instructional design paradigms. Hence, AI can provide crucial tools in not only opening the door for new patterns of engagement but also towards an appropriate set of metrics to honor the diverse ways that students demonstrate their learning. In turn, this approach can support inclusive learning environments for learners who may not resonate with traditional classroom activities or who carry a primary discourse that does not fully align with schooling practices.

23.2.5 DEI for AI: How Can the Principles of Diversity, Equity, and Inclusion Transform Artificial Intelligence in Education?

Principles of DEI can strengthen and expand AIED by informing research agendas and questions, operationalization and interpretation of variables, and revealing and mitigating biases in AI applications. These pursuits are challenging and require a deep understanding of inequitable outcomes along with underlying local and systemic causes and correlates. In turn, such efforts can inspire interventions to rectify harmful practices and environments, which must then be carefully evaluated for efficacy. Any of these goals – understanding, intervention, and evaluation – can become concrete research agendas for AIED scholarship. In other words, diversity, equity, inclusion, and related constructs can be a valid focus for research. Several of these commitments are specific to multimodal learning analytics, but others have applicability across research that bridges artificial intelligence and education (see Table 23.1).

23.2.6 Person-Centered Variables, Outcomes, and Ownership

DEI approaches enable more authentic consideration of rich, person-centered data. Such data extend beyond classic 'individual differences' like self-efficacy (Ballen et al., 2017) to encompass demographic, cultural, and contextual factors (e.g., race and gender, stereotypes and social norms, and power imbalances). Moreover, DEI conceptualizations emphasize the overlapping and contingent ways that such variables influence each other. For instance, the needs

TABLE 23.1

Twelve Commitments for Centering DEI in AIED Research

Data Collection	Analysis and Inference	Feedback and Dissemination
Multimodality: recognize that learning is a multimodal process.	*Multimodal data and human inference*: triangulate among different data sources and help inform interpretation of learner actions	*Transparency and benefit*: ensure that the research process is transparent to participants and that the experiences provide obvious benefits
Expansive learning experiences: advance opportunities to transcend traditional classroom activities.	*Limitations in prediction from multimodal data*: predictions should be about learner actions and not about assigning decontextualized and static labels to learners	*Multimodal feedback*: move beyond dashboards and consider ways to provide multimodal feedback to participants
Make learners' complexity visible: utilize sensors that can reveal hard to see interactions, actions, and states	*Participatory interpretation of multimodal data*: include participants within data analysis and inference processes	*Meaningful, usable feedback*: develop feedback that is both usable and understandable to people outside of the research community
Learning across spaces: people learn in a variety of contexts.	*Representation and multimodal data analysis*: the ways that data are represented and analyzed plays a major role in the inferences that we draw	
Multimodal data control: learners should have control of their data and how it is used		

and experiences of a first-generation, Autistic, White woman graduate student may be very different from those of a first-generation, Black woman graduate student. Even though they share 'first-generation' and 'woman' identities, both race and neurodiversity exert further mediating and moderating influences. Any analysis that focuses solely on one component of their identities would 'miss the mark'. A general algorithm can help to surface salient features, but the parameters or weights for those features may need to be updated to reflect individuals.

Critical and intersectional frameworks may provide a lens for describing and operationalizing DEI variables and their effects (Bauer et al., 2021; Bowleg, 2008; Cole, 2009; Else-Quest & Hyde, 2016). These approaches articulate that social identity categories (e.g., 'Black' and 'White', or 'man' and 'woman') entail substantial within-category variance stemming from the *intersection* of multiple categories. Every person embodies a multiplicative combination of identities. Intersectional frameworks also address how and why social identities are related to outcomes via differences in power and privilege. Thus, these perspectives offer the explanatory potential for predicting and testing relationships within our data. Instead of merely documenting disparities (e.g., a 'race gap'), we can study factors that might generate those gaps (e.g., disparities in academic preparation opportunities and resources) (Pierson et al., 2020).

It is important to note that the data alluded to above are not just 'person-centered' but also *personal*. This is especially true as we endeavor to make algorithms as individualized as possible to minimize bias. In some cases, AI algorithms might include log or clickstream data. In other instances, algorithms might include audio and video data, which might contain facial expressions, body poses, or information about shared joint attention. Given these challenges, it is essential that researchers carefully reflect on who should own the data within these systems and the implications that this has on participants. By and large, we advocate for student ownership of their data but recognize that this introduces additional challenges in terms of data analysis methodology and introduces an additional division between researchers and the data that can be used to support AIED research.

23.2.7 Revealing, Mitigating, and Preventing Biases in Analysis and Interpretation

Attention to DEI brings awareness of how people are historically ignored, neglected, or excluded, including the effects of systemic biases built into our technologies and human-technology interactions (Chen et al., 2020; Raji et al., 2020). Consequently, DEI frameworks provide a lens for inspecting the production and maintenance of inequities. Existing data might be (re)analyzed or (re)interpreted with respect to DEI principles regardless of whether DEI was the specific focus of the original research. If demographic data are collected, disaggregation allows researchers to explore differences, similarities, and disparities between groups along with crucial within-group variance. Likewise, equity-based approaches may help to account for observed findings such as explaining how and why individuals may perform differently as a result of experiences of exclusion.

A related concern is how we talk about the individuals who participate in research. This is particularly important for marginalized communities whose values, practices, and identities may often be treated as inferior (Williams & Gilbert, 2019). We need to avoid classifying or labeling individuals based on their experience or performance within a given environment. For instance, learners who receive lower grades (i.e., an event) should not be labeled as 'low achievers' (i.e., a trait). Doing so reiterates many of the approaches that educational institutions have used to exclude and oppress marginalized groups. This practice also fails to acknowledge the contextual nature of the data and is counter to the belief that people can learn and improve. A commitment to inclusive language also translates into the terms that we use to refer to different minoritized groups. Diverse groups have varying preferences, and researchers and practitioners should commit to and invest in genuinely learning how individuals prefer to be described (Dunn & Andrews, 2015).

23.2.8 Transparency in Feedback, and Dissemination

Research should provide meaningful benefits to the participants, and the processes and practices should be transparent. There is a long history of exploiting minoritized and marginalized groups to 'advance research' at significant risk to the participants. Although education research does not tend to produce noticeable physical or medical harm to participants, the AIED community must transcend merely avoiding harm and embrace practices that offer substantive benefits (i.e., beyond instruction or financial compensation). We must commit to sharing findings with participants in ways that they can reasonably interpret and proactively invite their feedback. Researchers also need to distill their findings into representations that can be understood by people outside of their discipline. This commitment increases the need for researchers to gather reflections and corroborate interpretations of data from participants, something that rarely happens within many AIED research projects.

In the following, we will present several examples in which DEI lenses can impact AI approaches. It should be mentioned that the implications of AIED-DEI extend beyond just research and how we do research, but also into how we diversify, educate, train, and prepare future scholars and practitioners in the field.

23.2.9 Considering Who Will Use the AI

One way to engage with DEI in AI is to intentionally and respectfully center marginalized identities within the design and implementation of the research. As a precursor to doing this work, researchers should reflect on their own positionality relative to the individuals that they will partner with and approach the setting from a place of bidirectional learning and value.

For example, Worsley and Bar-El (2020), and Bar-El and Worsley (2021), have described university courses that centered on disability both as the focus of the design space and as an important community for critiquing and improving practices within the AI community. One course engaged students in developing strategies for engaging with local organizations that serve people with disabilities and leveraging the capabilities of AI and digital fabrication in ways that are generative for their constituents. Example student-created designs included multimodal AI-based interfaces for navigating new physical spaces and for working with digital fabrication technologies. In other instances, students designed prototypes that used AI to support activities such as utilizing music apps within the Deaf and Hard of Hearing community. Others have explored ways to instrument loom technology with sensors and gesture detection algorithms to support blind and low-vision weavers.

All of the above examples were implemented in conjunction with marginalized communities, and the goals of the community were the driving focus of the work. Simultaneously, the class also challenged students to question underlying assumptions that may contribute to ongoing marginalization within the computer science communities and to recognize the various contributions that disabled people make to computer science. These contributions span a vast range of applications and prototypes, as well as important perspectives and critiques of the existing practices that are often used within the design community.

23.2.10 Educational Opportunities for AI

Another example of centering marginalized identities relates to working on putting the tools of artificial intelligence in the hands of younger people and youth of color as a means to equip and empower them as designers. Several researchers within the learning sciences and computer science community have been embarking on work within this space (Lee et al., 2015; Payne et al., 2021; Zimmermann-Niefield et al., 2019).

Jones et al. (2020) described one instantiation of this type of work that specifically sought to engage youth of color who participated in sports or other physical activities. Whereas many youth are socialized to see sports participation as distinct from (or even detrimental to) academic endeavors, the Sports Sense program (previously Data in Motion) positioned sports and sports participation assets for learning about artificial intelligence. Importantly, it did so from a bidirectional perspective, where sports could be a generative space to learn AI and AI could positively contribute to improving athletic performance. Within this program, youth were introduced to existing sports wearables and applications that utilized various types of artificial intelligence (e.g., machine learning, computer vision, gesture detection, and more). Students were supported as they explored these technologies and subsequently used low-cost tools to design their own AI-enabled sports wearables and applications.

Data from the first implementation of this program suggested that students possessed a number of compelling ideas for creating the next generation of sports wearables, and youth found this design space to be new, exciting, and something that they would be interested in pursuing long term. Arriving at this program design required the researchers to center the interests and motivations of youth from the start, and to consider ways that the learning experience could simultaneously teach them about artificial intelligence and contribute to meeting their goals.

23.3 Ethics and Challenges

Ensuring an effective bidirectional relationship between AI and DEI is complex, and there are many challenges to be addressed as well as significant benefits to be gained. Two examples of these challenges are (a) the ethical implications that are inherent in the analytical power that AI can provide and (b) the lack of understanding about AI among the vast majority of those whose data is and will be processed.

The work of the Institute for Ethical AI and Education[1] tackled the first of these challenges by developing a framework for practitioners and educational leaders to use when procuring AI for use in education. A key motivation for this work was to enable all learners to benefit optimally from AI and to be protected against known risks. In February 2020, an interim report, 'Towards a Shared Vision of Ethical AI in Education', was published by the IEAIED. This report outlined the risks and benefits posed by AI's use and suggested ways in which some of the tensions between the risks and the benefits might be ethically addressed. Suggestions from the report were used to drive a wide consultation with a cross-section of stakeholders through expert interviews and a series of roundtables – including three dedicated to participation by young people – and a Global Summit that brought together over 200 experts and authorities. The aim was to agree on a shared understanding of the ethical implications of using AI in education and to agree on a set of recommendations for how AI could be ethically designed and applied in education. The result from the consultation process was a four-page framework for educators that was organized around a set of nine objectives:

1. *Achieving Educational Goals.* AI should be used to achieve well-defined educational goals based on strong societal, educational, or scientific evidence that this is for the benefit of learners.
2. *Forms of Assessment.* AI should be used to assess and recognize a broader range of learners' talents.
3. *Administration and Workload.* AI should increase the capacity of organizations while respecting human relationships.
4. *Equity.* AI systems should be used in ways that promote equity between different groups of learners and not in ways that discriminate against any group of learners.
5. *Autonomy.* AI systems should be used to increase the level of control that learners have over their learning and development.
6. *Privacy.* A balance should be struck between privacy and the legitimate use of data for achieving well-defined and desirable educational goals.
7. *Transparency and Accountability.* Humans are ultimately responsible for educational outcomes and should therefore have an appropriate level of oversight of how AI systems operate.
8. *Informed Participation.* Learners, educators, and other relevant practitioners should have a reasonable understanding of artificial intelligence and its implications.
9. *Ethical Design.* AI resources should be designed by people who understand the impacts these resources will have.

Each objective is associated with criteria and questions that educators can pose to companies that are marketing an AI product or service. For example, the Equity objective has three criteria including, 'Develop and implement a strategy to reduce the digital divide among the cohort of learners for whom you have responsibility'. The associated question checklist item asked users to consider, 'Will the implementation of this strategy ensure that all learners for whom you are responsible are able to access and benefit from AI? (Pre-procurement)'. One important conclusion from the work completed by the Institute was the following theme:

> Only if well-intentioned people from diverse backgrounds continue to work together with the interests of learners in mind, especially the most disadvantaged, will we ensure that AI is truly going to find its optimal use, which maximize its potential and minimizes its downsides.

The clear need for input and engagement from a diverse population in the development of AI speaks to the second challenge identified at the start of this section: the lack of understanding about AI among the vast majority of those whose data is and will be processed.

The urgent need for people, and particularly educators, to better understand AI and the associated benefits of the participatory design were articulated by Luckin and Cukurova (2019). However, for educators and learners to confidently contribute to AI design, they need to understand basic AI concepts. For instance, why are data so important for machine learning AI? What data might be useful? How can data be accessed? How are new data collected? They need to understand that data has to be prepared by people before an AI can process it and need to understand the implications for the privacy of each individual and the security of their data at each step (i.e., from sourcing data to processing the data to outputting results). Most importantly, they need to understand the importance of the *imperative* for using AI: is the imperative one that will increase diversity, inclusion, and equity, or is it an imperative that will exacerbate existing inequalities?

In parallel, it is crucial for people to understand that the same data set or the same algorithms can produce dramatically different results due to the imperative of the AI application. For example, an algorithm might be deployed on data that has been harvested from public sources such as Twitter feeds, Facebook streams, LinkedIn profiles, or Instagram posts of a group of students taking an online course. The algorithm attempts to find patterns that are associated with cultural features in the students' data and produces a set of student profiles based on such cultural features. One imperative for the use of AI in this case might be to ensure that the course is sensitive to students' cultural context and that adaptations of the course material are guided by the AI-produced profiles. Alternatively, the imperative for the application of AI might be to ensure that students who are members of a particular profile are only allowed to interact with students from that same profile. Or, the aim might be to separate students who belong to a particular profile and provide them with a narrow course that will not broaden their horizons. These last two examples are obviously of concern, and yet the data and algorithm are remarkably similar to the first imperative case.

A possible tool for tackling this second challenge emerges from *AI Readiness*, which is a framework for providing educators with a contextualized practical experience of what AI is and what it can do. This framework is the basis for an AI Readiness course that helps educators and their leaders 'get inside' a machine learning algorithm – to explore what it can do with the type of data that the educators may have access to (Luckin & George, in press).

23.4 Conclusion

In this chapter, we discussed how AI methods and DEI approaches mutually benefit one another and can advance work in each space. AI methods expand the tools that researchers can use to identify barriers to DEI and then design, implement, and evaluate interventions to address these challenges (i.e., AI for DEI). Similarly, DEI approaches are not a superficial addition to AI research but instead offer a core perspective that ensures AI asks meaningful questions and employs equitable methods across diverse populations (i.e., DEI for AI).

AI and DEI can synergize in numerous ways, and the research documented in the preceding sections highlights several relevant approaches. Existing techniques remain relevant and can generate important insights, and there is also ample opportunity for innovation in this nascent field. AI methods have exciting potential for investigating complex relationships among demographic, performance, and behavioral variables (Kizilcec et al., 2013; Piech et al., 2012). For instance, several studies have shown that classic multiple regression and structural equation modeling approaches when they incorporate demographic variables informed by DEI principles, can challenge our assumptions about 'performance gaps' and pathways for different populations (Salehi et al., 2020; Ballen et al., 2017). Other studies have pioneered new methods such as Group Communication Analysis (GCA) (Dowell & Poquet, 2021; Schneider et al., 2021) and Multimodal Learning Analytics Across Spaces (CrossMMLA) (Blikstein & Worsley, 2016). The former combines natural language data, sequence mining, dynamic models, and demographic variables to understand (in)equitable discourse patterns. The latter combines multimodal sensors, gestures, natural language, computer vision, and additional data streams over time via machine learning to develop rich models of learners and to improve accessibility.

An important consideration for bridging AI and DEI is considering how demographic data are appropriately and ethically included in analytic models. A growing literature is investigating how quantitative analyses can be authentically intersectional to respect multiple and overlapping demographic 'categories' (Bauer et al., 2021; Else-Quest & Hyde, 2016), and the ways such categories affect learners' lives and performance. Likewise, learning analytics models typically treat demographic factors as distinct categories (e.g., 'Black' or 'White' or 'Asian') but can also further explore more nuanced *intra*categorical and *inter*categorical variance (McKay et al., 2018) that better captures the range of human experience.

Finally, although this bidirectional relationship is very promising, there are many questions open to further exploration and optimization. For instance, data in equity-oriented research paradigms can be both a *treasure* and a *terror*. Specifically, rich data enable deeper and more contextualized characterization of learners, their needs, and their journeys. However, these data also entail privacy challenges (e.g., revealing personal and identifying information) and have the potential to further stigmatize learners if misused (e.g., uncritically interpreting a 'performance gap' as 'evidence' of inferiority). Similarly, we must contend with and challenge methodological assumptions about statistical power and sampling. In the era of 'big data', techniques have been developed to address thousands and millions of data points. However, increasingly personalized, contextualized, and intersectional analyses drive analytical methods in the opposite direction. Equitable analyses need to be feasible, valid, and reliable even for 'small' samples and populations (e.g., Hispanic, LGBTQ+, Autistic, first-generation college students). There is no clear guideline for how much data is necessary to conduct appropriate analyses.

As always, careful and creative power analyses can be conducted to estimate sufficient sample sizes for a given study, but the field would benefit from new and innovative methods that work at broader scales. Data collection will also need to pursue more representative sampling such that data sets include authentic diversity. Recruiting diverse participants and building inclusive corpora will require more proactive and strategic sampling strategies than mere convenience sampling (Roscoe, 2021).

Overall, as the synergy between AI and DEI continues to develop, we can be guided by heuristic questions such as 'What are the best practices for collecting data?'; 'Who should own the data or have access to it once collected?'; 'How can we ensure privacy and confidentiality of participants when triangulated data facilitates identification?'; and 'How can results and findings be disseminated in ways that are precise, insightful, and beneficial to participants while forestalling misrepresentation or harmful conclusions?' The authors hope this chapter can promote discussion about linking DEI and AI approaches and inspire researchers and practitioners to answer the above questions in their work.

Acknowledgments

This chapter summarizes and expands upon a discussion panel held at the online 2021 Artificial Intelligence in Education (AIED) conference. Drs Roscoe and Salehi served as the organizers and moderators, and Drs Dowell, Worsley, Piech, and Luckin contributed as expert panelists. The authors are grateful to Drs Ido Roll and Danielle S. McNamara, who made the AIED conference panel possible. The authors also appreciate feedback from the book editors and reviewers. Portions of this work were supported by grants to Dr Roscoe (Gates Foundation INV-006123), Dr Dowell (Gates Foundation INV-000752), and Dr Salehi (Stanford's IDEAL initiative). All opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of any funding sources or agencies.

Note

1. https://www.buckingham.ac.uk/research-the-institute-for-ethical-ai-in-education/.

References

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–Affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241. https://doi.org/10.1016/j.ijhcs.2009.12.003

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Larusson & B. White (Eds.), *Learning analytics* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE—Life Sciences Education*, *16*(4), ar56. https://doi.org/10.1187/cbe.16-12-0344

Bar-El, D., & Worsley, M. (2021). Making the maker movement more inclusive: Lessons learned from a course on accessibility in making. *International Journal of Child-Computer Interaction*, *29*, 100285. https://doi.org/10.1016/j.ijcci.2021.100285

Bauer, G. R., Churchill, S. M., Mahendran, M., Walwyn, C., Lizotte, D., & Villa-Rueda, A. A. (2021). Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods. *SSM – Population Health*, *14*, 100798. https://doi.org/10.1016/j.ssmph.2021.100798

Blanchard, E. G. (2015). Socio-cultural imbalances in AIED research: Investigations, implications and opportunities. *International Journal of Artificial Intelligence in Education*, *25*(2), 204–228. https://doi.org/10.1007/s40593-014-0027-7

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex tasks. *Journal of Learning Analytics*, *3*(2), 220–238. https://doi.org/10.18608/jla.2016.32.11

Bowleg, L. (2008). When Black+ lesbian+ woman≠ Black lesbian woman: The methodological challenges of qualitative and quantitative intersectionality research. *Sex Roles*, *59*(5), 312–325. https://doi.org/10.1007/s11199-008-9400-z

Canz, T., Hoffmann, L., & Kania, R. (2020). Presentation-mode effects in large-scale writing assessments. *Assessing Writing*, *45*, 100470. https://doi.org/10.1016/j.asw.2020.100470

Card, D., & Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics*, *91*(11–12), 2158–2184. https://doi.org/10.1016/j.jpubeco.2007.03.006

Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2020). Ethical machine learning in healthcare. *Annual Review of Biomedical Data Science*, *4*, 123–144. https://doi.org/10.1146/annurev-biodatasci-092820-114757

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014–001. National Center for Education Statistics. https://eric.ed.gov/?id=ED544470

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist*, *64*(3), 170. https://doi.org/10.1037/a0014564

D'Mello, S. K., Dowell, N. M., & Graesser, A. C. (2009). Cohesion relationships in tutorial dialogue as predictors of affective states. In V. Dimitrova, R. Mizoguchi, B. Du Boulay, & A. C. Graesser (Eds.), *Artificial intelligence in education: Building learning systems that care: From knowledge representation to affective modeling* (pp. 9–16). IOS Press.

D'Mello, S., & Graesser, A. C. (2012). Language and discourse are powerful signals of student emotions during tutoring. *IEEE Transactions on Learning Technologies*, *5*(4), 304–317. https://doi.org/10.1109/TLT.2012.10

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences*, *112*(16), 4988–4993. https://doi.org/10.1073/pnas.1422822112

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29. https://doi.org/10.1177/2372732214549471

Dowell, N. M. (2019). Preparing for the future: Group communication analysis as a tool to facilitate adaptive support during digitally-mediated team interactions. *International Conference on Artificial Intelligence + Adaptive Education*. Beijing, China.

Dowell, N. M. M., McKay, T. A., & Perrett, G. (2021). It's not *that* you said it, it's *how* you said it: Exploring the linguistic mechanisms underlying values affirmation interventions at scale. *AERA Open*, *7*(1), 1–19. https://doi.org/10.1177/23328584211011611

Dowell, N. M. M., & Poquet, O. (2021). SCIP: Combining group communication and interpersonal positioning to identify emergent roles in scaled digital environments. *Computers in Human Behavior*, *119*, 106709. https://doi.org/10.1016/j.chb.2021.106709

Dowell, N. M., Lin, Y., Godfrey, A., & Brooks, C. (2019). Promoting inclusivity through time-dynamic discourse analysis in digitally-mediated collaborative learning. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. M. McLaren, & R. Luckin (Eds.), *Proceedings of the 20th international conference on artificial intelligence in education* (pp. 207–219). ACM. https://doi.org/10.1007/978-3-030-23204-7_18

Dowell, N. M., Lin, Y., Godfrey, A., & Brooks, C. (2020). Exploring the relationship between emergent sociocognitive roles, collaborative problem-solving skills and outcomes: A group communication analysis. *Journal of Learning Analytics*, *7*(1), 38–57. https://doi.org/10.18608/jla.2020.71.4

Dowell, N. M., Nixon, T., & Graesser, A. C. (2019). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multi-party interactions. *Behavior Research Methods*, *51*(3), 1007–1041. https://doi.org/10.3758/s13428-018-1102-z

Drozdowski, P., Rathgeb, C., Dantcheva, A., Damer, N., & Busch, C. (2020). Demographic bias in biometrics: A survey on an emerging challenge. *IEEE Transactions on Technology and Society*, *1*(2), 89–103. https://doi.org/10.1109/TTS.2020.2992344

Du, J., Ge, X., & Xu, J. (2015). Online collaborative learning activities: The perspectives of African American female students. *Computers and Education*, *82*, 152–161. https://doi.org/10.1016/j.compedu.2014.11.014

Dunn, D. S., & Andrews, E. E. (2015). Person-first and identity-first language: Developing psychologists' cultural competence using disability language. *American Psychologist*, *70*(3), 255–264. https://doi.org/10.1037/a0038636

Eddy, S. L., Brownell, S. E., Thummaphan, P., Lan, M.-C., & Wenderoth, M. P. (2015). Caution, student experience may vary: Social identities impact a student's experience in peer discussions. *CBE Life Sciences Education*, *14*(4), ar45. https://doi.org/10.1187/cbe.15-05-0108

Else-Quest, N. M., & Hyde, J. S. (2016). Intersectionality in quantitative psychological research: II. Methods and techniques. *Psychology of Women Quarterly*, *40*(3), 319–336. https://doi.org/10.1177/0361684316647953

Fahle, E. M., Reardon, S. F., Kalogrides, D., Weathers, E. S., & Jang, H. W. (2020). Racial segregation and school poverty in the United States, 1999–2016. *Race and Social Problems*, *12*(1), 42–56. https://doi.org/10.1007/s12552-019-09277-w

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

Hammond, J. W. (2019). Making our invisible racial agendas visible: Race talk in assessing writing, 1994–2018. *Assessing Writing*, *42*, 100425. https://doi.org/10.1016/j.asw.2019.100425

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00239-1

Huang, J., Dasgupta, A., Ghosh, A., Manning, J., & Sanders, M. (2014, March). Superposter behavior in MOOC forums. In M. Sahami, A. Fox, M. A. Hearst, & M. T.

H. Chi (Eds.), *Proceedings of the first ACM conference on learning @ scale conference* (pp. 117–126). https://doi.org/10.1145/2556325.2566249

Johnson, A. C., Wilson, J., & Roscoe, R. D. (2017). College student perceptions of writing errors, text quality, and author characteristics. *Assessing Writing*, *34*, 72–87. https://doi.org/10.1016/j.asw.2017.10.002

Johnson, D., & VanBrackle, L. (2012). Linguistic discrimination in writing assessment: How raters react to African American "errors," ESL errors, and standard English errors on a state-mandated writing exam. *Assessing Writing*, *17*(1), 35–54. https://doi.org/10.1016/j.asw.2011.10.001

Jones, S. T., Thompson, J., & Worsley, M. (2020). Data in motion: Sports as a site for expansive learning. *Computer Science Education*, *30*(3), 279–312. https://doi.org/10.1080/08993408.2020.1805287

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius: Sociological Research for a Dynamic World*, *7*. https://doi.org/10.1177/2378023121999581

Kauh, T. J., Read, J. G., & Scheitler, A. J. (2021). The critical role of racial/ethnic data disaggregation for health equity. *Population Research and Policy Review*, *40*(1), 1–7. https://doi.org/10.1007/s11113-020-09631-6

Ke, F., & Kwak, D. (2013). Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction. *Computers and Education*, *61*, 43–51. https://doi.org/10.1016/j.compedu.2012.09.003

Kizilcec, R. F., Piech, C., & Schneider, E. (2013, April). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). ACM. https://doi.org/10.1145/2460296.2460330

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, *86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Lee, V. R., Drake, J., & Williamson, K. (2015). Let's get physical: K-12 students using wearable devices to obtain and learn about data from physical activities. *TechTrends*, *59*(4), 46–53. https://doi.org/10.1007/s11528-015-0870-x

Lin, Y., & Dowell, N. M. (2019, July). Does gender really matter?: Exploring differences in emerging discourse styles during digitally-mediated collaborative interactions. Paper presented at the *29th Annual Meeting of the Society for Text and Discourse*.

Lin, Y., Dowell, N. M., Godfrey, A., Cho, H., & Brooks, C. (2019). Modeling gender differences in intra- and interpersonal dynamics during online learning collaborative interactions. In D. Azcona & R. Chung (Eds.), *Proceedings of the 9th international conference for learning analytics & knowledge* (pp. 431–435). ACM. https://doi.org/10.1145/3303772.3303837

Lin, Y., Yu, R., & Dowell, N. (2020, July). LIWCs the same, not the same: Gendered linguistic signals of performance and experience in online STEM courses. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Proceedings of the 21st artificial intelligence in education conference* (Vol. 12163, pp. 333–345). Springer. https://doi.org/10.1007/978-3-030-52237-7_27

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, *50*(6), 1–20. https://doi.org/10.1111/bjet.12861

Luckin, R., & George, K. (in press). *AI for school teachers*. Routledge.

Mayfield, E., Madaio, M., Prabhumoye, S., Gerritsen, D., McLaughlin, B., Dixon-Roman, E., & Black, A. W. (2019, August). Equity beyond bias in language technologies for education. In *Proceedings of the fourteenth workshop on innovative use of NLP for building educational applications* (pp. 444–460). Association for Computational Linguistics.

McKay, T., Grom, G., & Koester, B. (2018, March). Categorization, intersectionality, and learning analytics. In A. Pardo et al. (Eds.), *Proceedings of the 8th international conference on learning analytics and knowledge*. SOLAR. http://bit.ly/lak18-companion-proceedings

McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., & Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing*, *23*, 35–59. https://doi.org/10.1016/j.asw.2014.09.002

Miron, M., Tolan, S., Gómez, E., & Castillo, C. (2020). Evaluating causes of algorithmic bias in juvenile criminal recidivism. *Artificial Intelligence and Law*, 1–37. https://doi.org/10.1007/s10506-020-09268-y

Nichols, S., & Stahl, G. (2019). Intersectionality in higher education research: A systematic literature review. *Higher Education Research and Development*, *38*(6), 1255–1268. https://doi.org/10.1080/07294360.2019.1638348

Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, *25*(2), 177–203. https://doi.org/10.1007/s40593-014-0028-6

Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, *24*(4), 427–469. https://doi.org/10.1007/s40593-014-0029-5

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: Implications for health systems. *Journal of Global Health*, *9*(2), 020318, https://doi.org/10.7189%2Fjogh.09.020318

Payne, W. C., Bergner, Y., West, M. E., Charp, C., Shapiro, R. B. B., Szafir, D. A., Taylor, E. V., & DesPortes, K. (2021, May). danceON: Culturally responsive creative computing. In P. Bjørn & S. Drucker (Eds.), *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–16). https://doi.org/10.1145/3411764.3445149

Piech, C., Sahami, M., Koller, D., Cooper, S., & Blikstein, P. (2012, February). Modeling how students learn to program. In L. S. King, D. R. Musicant, T. Camp, & P. Tymann (Eds.), *Proceedings of the 43rd ACM technical symposium on computer science education* (pp. 153–160). https://doi.org/10.1145/2157136.2157182

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., Ramachandran, V., Barghouty, P., Phillips, C., Shroff, R., & Goel, S. (2020). A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour*, *4*(7), 736–745. https://doi.org/10.1038/s41562-020-0858-1

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020, February). Saving face: Investigating the ethical concerns of facial recognition auditing. In A. Markham, J. Powles, T. Walsh, & A. L. Washington (Eds.), *Proceedings of the AAAI/ACM conference on AI, ethics, and society* (pp. 145–151). https://doi.org/10.1145/3375627.3375820

Reardon, S. F., & Owens, A. (2014). 60 years after Brown: Trends and consequences of school segregation. *Annual Review of Sociology*, *40*(1), 199–218. https://doi.org/10.1146/annurev-soc-071913-043152

Reaser, J., Adger, C. T., Wolfram, W., & Christian, D. (2017). *Dialects at school: Educating linguistically diverse students*. Routledge. https://doi.org/10.4324/9781315772622

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/0.1007/s40593-016-0110-3

Roscoe, R. D. (2021). Designing for diversity: Inclusive sampling. *Ergodesign & HCI*, *9*(1), 67–81. http://doi.org/10.22570/ergodesignhci.v9i1.1502

Rose, T. (2016). *The end of average: How to succeed in a world that values sameness*. Penguin UK.

Rosenthal, L. (2016). Incorporating intersectionality into psychology: An opportunity to promote social justice and equity. *American Psychologist*, *71*(6), 474–485. https://psycnet.apa.org/doi/10.1037/a0040323

Salehi, S., Burkholder, E., Lepage, G. P., Pollock, S., & Wieman, C. (2019a). Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics. *Physical Review Physics Education Research*, *15*(2), 020114. https://doi.org/10.1103/PhysRevPhysEducRes.15.020114

Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., Harcombe, W., McGaugh, S., Wassenberg, D., Yonas, A., & Ballen, C. J. (2019b). Gender performance gaps across different assessment methods and the underlying mechanisms: The case of incoming preparation and test anxiety. *Frontiers in Education*, *4*, Article 107. https://doi.org/10.3389/feduc.2019.00107

Salehi, S., Cotner, S., & Ballen, C. J. (2020). Variation in incoming academic preparation: Consequences for minority and first-generation students. *Frontiers in Education*, *5*, Article 552364. https://doi.org/10.3389/feduc.2020.552364

Schneider, B., Dowell, N., & Thompson, K. (2021). Collaboration analytics—Current state and potential futures. *Journal of Learning Analytics*, *8*(1), 1–12. https://doi.org/10.18608/jla.2021.7447

Seymour, E., & Hunter, A. B. (2019). *Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education*. Springer Nature. https://doi.org/10.1007/978-3-030-25304-2

Shermis, M. D., Burstein, J., Elliot, N., Miel, S., & Foltz, P. W. (2016). Automated writing evaluation: An expanding body of knowledge. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 395–409). Guilford Press.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Information Age Publishing.

Shute, V., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C. P., Kuba, R., Liu, Z., Yang, X., & Sun, C. (2021). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *Journal of Computer Assisted Learning*, *37*(1), 127–141. https://doi.org/10.1111/jcal.12473

Spikol, D., Ochoa, X., Worsley, M., Di Mitri, D., Cukurova, M., Martinez-Maldonado, R., & Schneider, J. (2021, April). CROSSMMLA futures: Collecting and analysing multimodal data across the physical and virtual. In M. Scheffel, N. Dowell, S. Joksimovic, & G. Siemens (Eds.), *Proceedings of the 11th international conference on learning analytics and knowledge* (pp. 381–385). ACM.

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L. 2nd, Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., & Grummer, J. A. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483. https://doi.org/10.1073/pnas.1916903117

Walker, E., & Ogan, A. (2016). We're in this together: Intentional design of social relationships with AIED systems. *International Journal of Artificial Intelligence in Education*, *26*(2), 713–729. https://doi.org/10.1007/s40593-016-0100-5

Walsh, C. G., Chaudhry, B., Dua, P., Goodman, K. W., Kaplan, B., Kavuluru, R., Solomonides, A., & Subbian, V. (2020). Stigma, biomarkers, and algorithmic bias: Recommendations for precision behavioral health with artificial intelligence. *JAMIA Open*, *3*(1), 9–15. https://doi.org/10.1093/jamiaopen/ooz054

Wang, K. D., Salehi, S., Arseneault, M., Nair, K., & Wieman, C. (2021, June). Automating the assessment of problem-solving practices using log data and data mining techniques. In C. Meinel, M. Pérez-Sanagustín, M. Specht, & A. Ogan (Eds.), *Proceedings of the eighth ACM conference on learning@ scale* (pp. 69–76). ACM. https://doi.org/10.1145/3430895.3460127

Wiens, J., Price, W. N., & Sjoding, M. W. (2020). Diagnosing bias in data-driven algorithms for healthcare. *Nature Medicine*, *26*(1), 25–26. https://doi.org/10.1038/s41591-019-0726-6

Williams, R. M., & Gilbert, J. C. (2019). "Nothing about us without us": Transforming participatory research and ethics in human systems engineering. In R. D. Roscoe, E. K. Chiou, & A. R. Wooldridge (Eds.), *Advancing diversity, inclusion, and social justice through human systems engineering* (pp. 113–134). CRC Press. https://doi.org/10.1201/9780429425905

Wilson, J., & Roscoe, R. D. (2020). Automated writing evaluation and feedback: Multiple metrics of efficacy. *Journal of Educational Computing Research*, *58*(1), 87–125. https://doi.org/10.1177/0735633119830764

Wisser, L. (2019). Pandora's algorithmic black box: The challenges of using algorithmic risk assessments in sentencing. *American Criminal Law Review*, *56*, 1811–1832.

Worsley, M., & Bar-El, D. (2020). Inclusive making: Designing tools and experiences to promote accessibility and redefine making. *Computer Science Education*. https://doi.org/10.1080/08993408.2020.1863705

Yan, D., Rupp, A. A., & Foltz, P. W. (Eds.). (2020). *Handbook of automated scoring: Theory into practice* (1st ed.). CRC Press. https://doi.org/10.1201/9781351264808

Zimmermann-Niefield, A., Turner, M., Murphy, B., Kane, S. K., & Shapiro, R. B. (2019, June). Youth learning machine learning through building models of athletic moves. In J. A. Fails (Ed.), *Proceedings of the 18th ACM international conference on interaction design and children* (pp. 121–132). ACM. https://doi.org/10.1145/3311927.3323139