Who's Got the Power? Data Feminism as a Lens for Designing AIED Engagement Systems

Angela E.B. Stewart[1 [0000−0002−6004−9266]], Jaemarie Solyst[2 [0000−0002−0549−0454]], Xinyi Bao[1 [0009−0002−3917−8765]], Paras Sharma[1 [0009−0007−7645−653X]], Amanda Buddemeyer[3 [0000−0002−2277−2402]], Tara Nkrumah[4 [0000−0002−4723−1944]], Amy Ogan[3 [0000−0003−2671−6149]], and Erin Walker[1 [0000−0002−4723−1944]]

[1] University of Pittsburgh, Pittsburgh PA 15260, USA
{angelas,xib22,pas252,eawalker}@pitt.edu
[2] University of Washington, Seattle WA 98195, USA
jaemarie@cs.washington.edu
[3] Carnegie Mellon University, Pittsburgh PA 15213, USA
{amandab,aeo}@andrew.cmu.edu
[4] Arizona State University, Tempe AZ 85281, USA
tnkrumah@asu.edu

Abstract. A goal of the AIED community is to create equitable systems; yet, we lack a cohesive viewpoint on how to do so. In the present work, we propose power as this organizing principle. We utilize the data feminism framework to showcase how we might balance power, focusing on learner engagement. We utilize multimodal data from ten middle school girls in a virtual computer science camp to discuss how the AIED community might create systems of equity that support all learners.

Keywords: Equity · Data Feminism · Engagement · Multimodal Data.

1 Introduction

The AIED community has increasingly grappled with what it means to create equitable AI systems. Equity refers to "fairness or justice in the way people are treated" [33], whereas inequity is the lack of fairness or justice [34]. Within our community, equity has been conceptualized as dealing with equal access to technologies [23], engagement in ethical design practices [6], or unbiased model results [6]. However, these efforts lack a cohesive viewpoint, where we make connections between individual instances of inequity. These connections are crucial because inequity is systemic, or embedded into all parts of a system [5, 13, 14, 38]. Thus, inequity is embedded into how we conceptualize our technologies [6, 14, 37, 38], design and implementation processes [12, 14, 46, 47], analysis and evaluation frames [6, 23, 28], and distribution and usage of technologies [23, 39].

We argue that attending to power is the key organizing viewpoint by which we can view equity in AIED systems. Power refers to control, authority, or influence over others [10, 14]. Power encompasses an individual's experience, systemic policies and authorities, and hegemonic ideas that exert control over people [10]. When power is imbalanced, inequity occurs. For example, power imbalance shows up in disparities in access to AI innovations created by our community [23]. This is an example of power because certain groups, such as people from low-income households, are prevented from accessing technologies due to financial barriers [39]. As another example, researchers have shown race- and gender-based disparities in learners' comfort with data collected by AIED systems [28]. This is an example of power at work, with women and Black students recognizing the lack of control they have over their data [14, 28]. Power is the invisible force with visible effects - shaping our experience as system designers, learners, teachers, and other stakeholders in AIED [13, 14]. Without addressing power, we fail to address the root cause of inequity [5, 10, 12, 14].

We use the data feminism framework for exploring how to incorporate an interdisciplinary understanding of power into AIED systems. Data feminism is a socio-technical framework that presents principles for how we might consider power in data-centric systems [14]. It has previously been used in the Learning Analytics community as a way of posing guiding questions towards more equitable practices [38, 46]. This current work is guided by the following question: **How can the data feminism framework inform what we choose to build and how we choose to build it in AIED?** While there have been previous calls to do this kind of social-technical analysis [6, 23, 37, 38, 46], there is a dearth of studies that show worked examples in authentic contexts. To our knowledge, we contribute the first case study for data feminism in AIED.

We focus on engagement — a long-studied construct in AIED, presenting a particularly rich space for exploring power. Engagement detectors use behavioral traces from learners to predict whether or not they are engaged in a certain task [16, 37]. However, this paradigm is exemplary of a power imbalance, where learners lack the power to make decisions about what counts as an engaged behavior and how a system might respond [38]. This is particularly problematic because learners, especially those at the margins of society, often have their expressions of engagement misunderstood or dismissed [2, 40, 47, 46]. Our data context is rich for this exploration — middle school girls engaged in a virtual computing camp in 2020. Girls' engagement in computing education has long been misunderstood as pedagogical strategies tend to include male-centric curricular activities and behavioral norms [42]. Furthermore, the camp took place during the COVID-19 pandemic, when the learners interacted primarily online, and widespread challenges to cultivating engagement persisted [44].

Our present work is not a traditional paper — instead, we present an illustration of how the data feminism principles might inform AIED. To create equitable systems, we must first **examine power**, which we do in our literature review (Section 2). We then present our case study in Sections 3 and 4. The work of creating equitable AIED systems is vast and requires varied expertise that often goes unrecognized. Therefore, in Section 3, we **make labor visible** related to the varied expertise needed for this work. Section 4 presents a worked example of how we might use multimodal data generated from an online computing camp to **elevate emotion and embodiment**, **rethink binaries and hierarchies**, **consider context**, and **embrace pluralism**. Finally, in Section 5, we form a cohesive vision of how we might **challenge power** in AIED systems.

2 Examine Power - Related Work

In this section, we use the data feminism principle **examine power** to understand how power influences what we choose to build and how we choose to build it, with a focus on engagement. Engagement is considered a prerequisite to learning [2, 16, 40]. However, a precise definition of the phenomenon remains hazy. Researchers consider engagement a multi-faceted phenomenon, relying on behavioral, cognitive, and socio-emotional indicators[16, 40]. For example, affective dimensions of engagement, such as connectedness and belonging, can interact with cognitive factors such as self-regulation and motivation to produce behaviors that indicate engagement (e.g., assignment completion) [40].

The AIED community has a robust scholarship on detecting learners' engaged behaviors [43]. This work focuses on the many aspects of engaged and disengaged learning, such as concentration [21], mind-wandering [43], emotional engagement [7], or cognitive engagement [20]. Further, engagement has been investigated across learning environments, such as in-person classrooms [1], asynchronous courses [15], flipped classrooms [30], or massively open online courses [20].

This scholarship is influenced by the power differential between the technology designer/researcher and the learners at the center of the technology. Implicit in this scholarship is the assumption that the designer of educational technologies has more expertise than users in determining what will lead to learning [12, 18, 23]. An extension of this line of reasoning is that indicators of engagement can be more effectively detected by an AI system than expressed directly by the student [14, 18, 38]. Work in AIED should consider returning power to learners and teachers by confirming detector results with them, rather than relying solely on detectors to provide this determination [26, 38].

Another assumption made by this work is that individuals exhibit engagement in similar ways. However, expressions of engagement are quite diverse, varying by individual learners according to their culture, values, preferences, and abilities [2, 16, 40, 47]. If the technology designer is using the cultural standpoint of a dominant group, learners who are not in said groups are likely to be held to irrelevant or harmful standards [2, 12, 37, 46, 47]. Given that we cannot pinpoint a particular set of expressions that universally indicate engagement, our definitions need to remain flexible [46, 47]. While a limited amount of work has explored how behavioral indicators of engagement vary by the pedagogical context [47], more work should explore the nuances of these individual and cultural differences, and consider the implicit standpoint on which detectors are built.

Most concerning, by emphasizing an individual's engagement, we create a false narrative that we understand engagement, without looking at the external factors that produce or fail to produce it. Naming specific systems of power (e.g., cisheteropatriarchy, racism) can be helpful here, and indeed is an integral part of examining power. Consider cisheteropatriarchy, the interlocking systems of patriarchy, heterosexism, and trans* oppression, that control women, girls, and LGBTQIA+ people [27]. Under cisheteropatriarchy, women's academic achievements are rarely centered in learning spaces [27], marginalizing alternative ideas that do not center men. This lack of representation reflects a power dynamic that marginalizes femme learners by erasing the contributions of people similar to them. Lack of representation is damaging to engagement, for example, by feeding into deficit narratives of who can and cannot succeed in the field, and discouraging them from bringing their unique viewpoints to the learning environment [31]. Pulling from the sociological theory of intersectionality, we know that power overlaps across identity categories, compounding its effects for people at the intersections [11]. For example, learners who are both girls and people of color experience both cisheteropatriarchy and racism, where social, political, economic, and cultural rewards are at least in part allocated along racial lines [5]. Concerning engagement, learners might engage in suppressing their authentic racial or gendered expressions in favor of blending into the dominant norm [27]. This may show up as adopting performative behaviors, such as head nodding, to appear more engaged [27]. This showcases the ways that learners are indeed engaged, but in unexpected and likely counterproductive ways [36].

Broadly, within the AIED community, our response to disengagement is to build an intervention on the learner, for example, detecting if they are engaged and creating supports for them to re-engage, such as personalized stimuli or alerting a teacher for support [16, 43]. However, these in-the-moment interventions do not challenge the surrounding power structures, such as those outlined above, that cause disengagement. However, now that we have explicitly named these systems, we can reference the people, community organizations, and scholars that are doing the work to rebalance power [14]. Our work builds off of this notion by working directly with community partners that center learning experiences that encourage alternate modes of learning engagement.

3 Make Labor Visible - Data Collection and Processing

The work of creating AIED systems involves many hands, yet much of this work goes unrecognized [12, 14]. This is an unequal power dynamic as some work is credited as being more important to designing AIED systems, while other foundational work goes uncredited. Recognizing this invisible labor is important because it now affords us opportunities from which we can learn and build our systems. Throughout this section, we connect our work to the data feminism principle **make labor visible**, by surfacing the varied and extensive work necessary to create AIED systems. In our work, this involved designing activities and tools for an online computing camp, coordinating recruitment and logistics, facilitation and data collection, and data processing. In the following sections, we first describe our methods and the requisite labor that went into the design.

3.1 Online Computing Camp

Participants were ten middle school girls (ages 12-14; four white, one Hispanic/Latina and white, one Asian and white, two Hispanic/Latina, one American Indian/Alaskan Native, and one chose not to report). All participants were recruited through a regional chapter of an international organization focused on empowering girls. Nine learners self-reported prior programming experience (e.g., Scratch [41]). The camp was over three subsequent Saturdays on Zoom [49], with sessions lasting two to three hours a day. For recruitment and logistics, extensive labor was involved in working with a local organization to recruit these learners. This involved discussions to ensure alignment of values and goals for the camp, coordination of recruitment and participant compensation, as well as consent and communication coordination with both participants and their parents.

The camp was designed to provide an introductory, culturally-responsive computing experience. Culturally-responsive computing aims to address both technical literacy alongside social topics, such as community, identity, and power [42]. Camp activities combined computer science concepts and reflective exercises focusing on power and identity in learners' everyday lives.

Three main facilitating instructors led 53 activities that fell into one of ten categories (Table 1). Note, activities could be quite short, leading to an overall high number of activities. For example, we conducted a **Power and Identity** activity where learners were asked to reflect on whether certain words represented power, immediately followed by similar reflections on images. These were counted as separate activities as the stimuli changed between them, even if the individual activities themself were related and brief.

For designing camp activities, we leveraged the expertise of one member of our research team who specializes in culturally-relevant teaching [25]. She worked to create activities to connect with learners and scaffold their reflection on camp topics. This work was done in concert with other members of the team to apply culturally-relevant pedagogy to computing education [25, 42]. Running the camp entailed social and emotional labor [14] to build relationships with learners while supporting their learning of the camp content. Additionally, research assistants sat in on the session, recording field notes, allowing us to improve the camp between subsequent weekends, and providing crucial context to our eventual data analysis.

Programming was done in a custom-built, online, block-based programming interface [17], where the goal was to use code blocks to control a robotic character (similar to [9]). The programming interface consisted of (1) a canvas for learners to build their programs, (2) a robot, which executed commands, and (3) a toolbox of blocks with standard programming constructs (e.g., if statements and loops), as well as custom-built blocks to control the robot (e.g., speak or move). One member of our research team was responsible for designing, building, testing, and maintaining the custom programming environment and generating log data.

Table 1. Descriptions of the categories and number of activities are shown.

Category	#	Description
Active Prompt	2	Respond to prompt via chat
Breakout Room Activity	2	Collaborate in small groups to solve a problem
Programming	6	Individual programming tasks following the lessons
Community Building	6	Get to know other learners
Feedback	3	Give facilitators feedback on how to improve the camp
Lesson	9	Learn computer science concepts and how to implement in programming interface
Movement	3	Move around to increase energy
Power and Identity	11	Reflect on culture and representations of power and identity
Presentation	5	Presentations about robots, programming, and notable women of color in computing
Share Out	6	Share creations from programming assignments

3.2 Data Processing and Metrics

To analyze engagement within the camp, we utilized data from three modalities: chats, speech, and programming logs. We removed irrelevant data, such as logging into the session. In total, 759 chat utterances were included in our analyses. For speech data, we utilized 660 human-transcribed utterances, generated from a third-party service. Two researchers quality-checked the transcriptions and corrected speaker identification, discarding 22 unidentifiable utterances. For both chats and transcripts, we tokenized words using NLTK [3].

To compute quantitative engagement metrics, we summarized the signals at the same granularity. Accordingly, we aggregated the data per category as shown in Table 1. Since activity duration varied, we normalized the number of words chatted or spoken by dividing by the activity duration (minutes). Our final **engagement magnitude** metrics were words chatted and words spoken per minute. We also created a **binary engagement** metric to indicate whether or not a learner engaged in a particular activity. This was done to capture overall engagement patterns, regardless of magnitude (see Section 4.2).

We qualitatively examined engagement by coding the content of what learners programmed during programming tasks. Each task contained a programming component and a Power Identity component, which correspond to the technical literacy and culturally-responsive aims of the camp (see Section 3.1. For example, in the first task, learners were asked to create a program with text that demonstrates a problem they see in their community and how they might solve it. The programming component of this task was to properly use the text block – this was one code for this task. The Power Identity component was to define a community problem and solution. These corresponded to two codes, one defining the problem and one for the solution. One task was omitted from the analysis because we could not meaningfully define Power Identity codes. To account for learners engaging with the programming task in unexpected ways, we added codes for when the program content went beyond the task at hand. For example, in the task given above, a learner went beyond the task at hand in the programming component by using blocks other than a text block. For the Power Identity component, they could go beyond the task by explaining why a problem is important to solve, how it impacts their community, or they could provide multiple solutions. Two coders independently marked each learner's final program for whether or not it satisfied each of the defined codes. We utilized consensus-based coding and resolved disagreements through discussion [19].

To summarize the labor involved in data processing, a team member segmented and labeled the start and end times of activities. Verbal data was extensively cleaned, transcribed, and aggregated to form metrics. Programming logs were extracted, cleaned, and qualitatively coded.

3.3 Learner, Parent, and Community Organization Labor

Labor was also provided by those outside our research team, namely, learners and their parents. Parents enrolled their child in our partner organization, found relevant opportunities, such as our camp, and coordinated schedules and technology to allow their child to participate. Some parents even went so far as to email us feedback related to their children's engagement in the camp compared to school. Finally, the child had to express interest and be present on the day of the camp, using their limited out-of-school time to engage with us.

So often in the AI space, we elevate technical knowledge [14, 18]. Yet, as we elucidate in this section, an interconnected network of social, emotional, knowledge, and time labor is necessary just to collect the rich data upon which we build our technologies. By not attending to the power imbalance in recognized labor, we miss out on opportunities for creating better technologies. For example, we worked directly with a community organization whose mission is to educate learners across many backgrounds. As this is a well-stated goal of AIED [6, 23], attending to relationship-building labor here is important. In our case, we worked alongside our community organization to understand their values as they relate to events for their members, and how they conceptualize important aspects of their members' engagement. Through these conversations, we learned to focus on how learners are thinking about their own identity and relationship to their community. Accordingly, these values are reflected both in our selection of engagement metrics (e.g., in how we qualitatively coded the programming logs) and in activity design (see Section 3.1). Further, our interactions with parents themselves caused us to rethink how we were considering engagement. For example, we experienced difficulties cultivating verbal engagement (Section 4.1); however, we received direct reports from parents on how excited their child was to return to camp. This caused us to rethink how we were imposing our narrow definition of engagement on learners, which is explored more in Section 4.

4 Results

We leverage three data feminism principles as a theoretical lens, demonstrating how they can be incorporated into AIED. Note, this is not a comprehensive analysis, but instead an illustration of analytic methods to address power.

4.1 Elevate Emotion and Embodiment

We use the data feminism principle **elevate emotion and embodiment** as a starting place for our analyses. Learners exist in the world as "living, feeling bodies" [14]. Upending dominant power dynamics means creating AIED systems that emphasize learners' embodied and emotional experience, as it is precisely these experiences that AIED technologies try to support. We selected a short vignette that demonstrates learners' emotional and embodied experience of engagement on the first day of camp. This particular vignette was identified as a core moment of learner engagement and connection through examining field notes, especially, triangulating moments that were notated by multiple researchers.

The camp took place during Fall 2020, which corresponded to COVID-19 lockdowns and in-person school closures for our learners. Because of this, they were spending excessive learning time behind their computers. During a community-building activity, learners were asked to express their feelings by choosing from a set of images and sharing why they chose that image. Several learners expressed positive emotions (e.g., "happy", "ready to learn"), and several others either expressed a negative feeling ("tired") or were slow to respond. A lead facilitator took these expressions as an opportunity to engage learners, breaking from the pre-defined protocol and asking if they were tired of Zoom. Many learners confirmed they were bored with Zoom, tired of sitting in the same chair all day, and struggling to pay attention. The facilitator validated those feelings, which may have been the best she could accomplish in the moment, being unable to relieve the emotional or physical impact of lockdown. The learners' engagement, and indeed any conclusions that we can draw about the camp, will always be colored by the circumstances: a virtual camp rather than face-to-face, a virtual robot rather than a physical one, and learners whose emotional and embodied experiences were shaped by a trying moment in history. Indeed, it was these interactions that indicated a deeper dive into engagement was necessary, utilizing the signals available to us and our discussions with learners.

4.2 Rethink Binaries and Hierarchies, Consider Context

Building on our starting place of emotion and embodiment, we use behavioral traces to rethink how we conceptualize engagement by bringing in two other data feminism principles: **rethink binaries and hierarchies** and **consider context**. AIED systems are fundamentally built on hierarchies and binaries, where learners are categorized into conceptual buckets based on some pre-defined criteria, and some buckets of behavior are considered better than others. In engagement detection work, the binary is whether the learner is engaged or not, and the related hierarchy is that more engagement is better than less. However, by **rethinking binaries and hierarchies**, we can reevaluate the ways our AIED systems might perpetuate dominant views of engagement that exclude learners' diverse expressions [2]. We marry this data feminism principle with **consider context**, which posits that considering the setting in which the data is collected can elucidate more meaning. In this analysis, we consider engagement binaries and hierarchies against two contextual backdrops: the activity at hand (curricular context) and individual differences (learner context).

We classified each learner as overall low or high in chat contributions using a median split per modality [24]. Note, we label learners as high or low to demonstrate the limits of such a label, and not because these learners should be considered as such. We then computed the proportion of low chat contribution learners who were engaged in a particular category (according to our binary engagement metric). This tells us which types of activities most engaged the low chat learners. We repeated this for high chat contributions learners, as well as for low and high speech contributions. Results are shown in Table 2. Regardless of modality, the learners who were overall highly engaged tended to also be engaged in each category (unless no engagement on that modality was required, e.g., speaking out loud during Active Prompt or Movement).

Table 2. For each modality, the proportions of all (A), overall low (L), or overall high (H) engagement learners that were engaged in each category are shown.

	Chats			Speech		
Category	**A**	**L**	**H**	**A**	**L**	**H**
Active Prompt	.8	.6	1	0	0	0
Breakout Room Activity	1	1	1	.9	.8	1
Programming	1	1	1	.7	.4	1
Community Building	1	1	1	.8	.6	1
Feedback	.7	.4	1	.5	.2	.8
Lesson	.8	.6	1	.4	.2	.6
Movement	1	1	1	.3	.2	.4
Power and Identity	1	1	1	.9	.8	1
Presentation	.9	.8	1	.6	.2	1
Share Out	.7	.6	.8	.8	.6	1

This view of the data is particularly insightful for low chat or speech contribution learners because it shows which types of activities best solicited their responses. For example, small group activities were effective, with all low chat learners engaging in Breakout Room and Programming categories. This is similar for speech (e.g., four out of five low speech learners engaged in Breakout Rooms). For the chats, activities that most closely aligned with the culturally-responsive goals of the camp (Community Building, Movement, and Power Identity) successfully solicited contributions from all low chat learners. Taken together, a typical view of these learners is that they are low-engagement, but that turned out not to be true. They were contextually engaged according to the activity, indicating that pedagogical strategies, which are typically outside the learners' control, are important to when and how learners engage. Further, this shows that the "low engagement" label is insufficient, with patterns of engagement emerging only when individual and curricular context were layered in the analysis.

4.3 Embrace Pluralism

The verbal engagement indicators are limited in how they reflect a learner's viewpoint during the camp, as they describe what learners were doing, but not how they were doing it. Data feminism argues that it is necessary to **embrace pluralism** by synthesizing multiple viewpoints, with priority given to the viewpoints of those most affected by the analyses. Computing behavioral metrics alone is not enough to categorize engagement. Therefore, we embrace pluralism by examining learners' actions during the programming task. This gives us an understanding of learner thought processes, thus elevating their perspective.

As there were three programming tasks, we formed groupings related to meeting the task requirements and producing programs beyond the requirements. For the former, for each task, we calculated whether each learner satisfied full, partial, or none of the programming requirements. Similarly, we calculated whether they satisfied full, partial, or none of the Power Identity requirements. The proportion of learners who satisfied each combination of programming and Power Identity (PI) requirements is shown in Table 3. For going beyond the task requirements, we calculated whether learners included extra programming, Power Identity, both, or no content in their programs (also shown in Table 3).

Table 3. The proportion of learners satisfying the full, partial, or no requirements from the programming (P) and Power Identity (PI) aspects of programming tasks is shown. The proportion of learners who went beyond task requirements is also shown.

<table>
<tr><td rowspan="4">Task Requirements</td><td>Partition</td><td>Full</td><td>Full CS
Partial PI</td><td>Full CS
No PI</td><td>Partial CS
Full PI</td><td>Partial CS
Partial PI</td><td>None</td></tr>
<tr><td>Text</td><td>0.5</td><td>0.2</td><td>0.1</td><td>0.0</td><td>0.0</td><td>0.2</td></tr>
<tr><td>Text & Movement</td><td>0.5</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.3</td><td>0.2</td></tr>
<tr><td>Variables</td><td>0.3</td><td>0.0</td><td>0.0</td><td>0.0</td><td>0.4</td><td>0.3</td></tr>
<tr><td rowspan="4">Beyond Task Requirements</td><td colspan="2">Partition</td><td colspan="2">Extra CS & PI</td><td>Extra CS</td><td>Extra PI</td><td>None</td></tr>
<tr><td colspan="2">Text</td><td colspan="2">0.1</td><td>0.1</td><td>0.4</td><td>0.4</td></tr>
<tr><td colspan="2">Text & Movement</td><td colspan="2">0.1</td><td>0.0</td><td>0.5</td><td>0.4</td></tr>
<tr><td colspan="2">Variables</td><td colspan="2">0.0</td><td>0.1</td><td>0.0</td><td>0.9</td></tr>
</table>

We found that no learners completed the full Power Identity requirements at the expense of programming requirements, indicating they chiefly focused on the programming requirements. Across tasks, when learners go beyond the requirements, they are more often added to Power Identity portions than programming. An average of 18% learners went beyond the Power Identity requirements, and only 7% did for programming. This view of learner engagement would have been missed had we focused solely on technical knowledge as a marker of engagement. Instead, as guided by the values of our community partners, we also focused on how learners attended to their identity, showcasing these alternate forms of engagement. Taken together, by embracing pluralism, we synthesize multiple perspectives; for our work, that means not just elevating a narrow set of markers of engagement, but also showcasing thinking processes.

5 Discussion

The present work demonstrates how power imbalances are embedded into existing AIED practices, such as how we conceptualize learning constructs like engagement (Section 2), whose perspective (Sections 4.1, 4.3) and labor (Section 3) is considered valuable, and the binaries and hierarchies implicit in our systems (Section 4.2). We do this all in service of the data feminism principle **challenge power**. Challenging power means that we disrupt the systems that cause power imbalance. This section is structured to synthesize five key insights surrounding our guiding question: how can the data feminism framework inform what we choose to build and how we choose to build it in AIED?

With respect to what we choose to build, our first key insight is to **Make our systems systemic.** The AIED community focuses keenly on the technology intervention. However, our work has demonstrated that thinking about the entire system surrounding learners is crucial to challenging power and creating equitable systems. For example, our system goes beyond the engagement analytics themselves to also include our relationship-building and values-alignment with community partners, the culturally-responsive pedagogical and curricular strategy, and the contextual backdrop of our work being conducted during a particularly difficult moment for learners (COVID-19 pandemic). However, this is not an exhaustive list of levers in a learning system. Indeed, shifting from an intervention approach to a systems approach requires interdisciplinary scholarship with those who are elucidating those levers. For example, we might leverage the work of education researchers who synthesize frameworks for promoting equity through leadership, community, policy, and teaching and learning [13].

Our second key insight for what we choose to build is to **focus on co-creation instead of surveillance.** The surveillance frame, where learners are constantly observed and assessed by AIED technologies, puts the power into the hands of an unseen authority [22]. Instead, we can co-create our technologies by leveraging co-design methodologies that inherently center those most affected by said technology [12, 14, 26]. Authentic co-design involves the reciprocal design of technologies from conception, ensuring that the voices of those using our technologies are engaged at all steps of the process. That said, co-creating our technologies goes beyond the design process itself. It also means co-creating alongside the AI system [29]. In the co-creative AI frame, a learner is in dialogue with the AI technology itself, and the learner is defining what the AI is doing, giving real-time feedback on AI interpretations of learner data, and potential next-steps for intervention. This kind of frame allows for the user of an AI system to hold real power, actively controlling the system and adding their unique knowledge alongside the AI.

Concerning how data feminism can inform how we choose to build AIED systems, we join other scholars in this space by suggesting that researchers **Look to learner assets.** An asset-based approach leverages learners' existing resources for their learning, rather than exclusively focusing on what learners lack (i.e., their deficits) [37, 46]. In our work, an asset-based approach told us far more about learners' engagement than a deficit-based approach. For example, by leveraging young learners' existing preference for text-chat rather than speaking out loud [44], we were able to get a more holistic view of their engagement rather than if we had relied solely on the typical measure of speaking out loud [40]. By focusing on who learners are, their values, and how we can leverage their existing assets, we are able to better support authentic engagement, rather than enforcing our narrow view of engagement on them.

Our second insight for how we build AIED systems is to **look to the margins, rather than focusing on generalizability.** Generalizability is an oft sought-after goal in AIED [45], prioritizing systems that are deployable across task contexts and people. However, this ignores that some people, contexts, and scenarios are inherently at the margins. Those experiences are not elevated if we focus on large-scale patterns [14]. To realize this, we must work closely with small groups to qualitatively understand their unique context and design in partnership. This isn't to say we can't use data mining techniques, but that if we are to understand learners at the margins, we have to shift our analysis techniques to zoom in on those experiences. This will mean that sample sizes become smaller, as is the case in the present work, where we focus on depth and richness of data [32, 35, 48]. Small samples allow researchers to dedicate more time and resources to each participant, resulting in richer and more detailed data.

Our final insight into how to build AIED systems is **don't be a stranger.** Data feminism posits this notion of being a "stranger in the dataset," where the person analyzing the data is so far removed from the data itself that crucial meaning is lost as assumptions and interpretations abstract away local meaning. Indeed, this is a core strength of our work, where we were not "strangers" instead working closely with our community partners, learners, and parents, as well as actually facilitating the camp and data collection. That said, in AIED, we often utilize large existing datasets upon which to build our work. When this is the case, researchers have little control over how the data was collected and the surrounding context. There are still mechanisms to not be a stranger in this case, such as involving those who can support in conceptualizing important constructs and interpreting results. For example, AIED researchers might conduct focus groups or interviews with key stakeholders (e.g., parents, teachers, or even children) [4, 8]. This would entail extra, but necessary, labor from our community to communicate findings clearly to non-technical audiences, and reflect alongside them on how results can be meaningful in real-world contexts. This would also necessitate a shift in labor on how these kinds of studies are reviewed and accepted by the AIED community.

In conclusion, we present a worked example of how data feminism can be used to understand power and power imbalances in an AIED context. We then theorized about how this understanding can inform what we choose to build and how we choose to build it in AIED. This is all in service of the goal of creating more equitable AIED systems.

Acknowledgments. We thank Drs. Kimberly Scott, Leshell Hatley, and Sharon Henderson for their support on this project. This work is funded by NSF DRL-2315042, DRL-2415872, and DRL-2415873.

References

1. Ahuja, K., Kim, D., Xhakaj, F., Varga, V., Xie, A., Zhang, S., Townsend, J.E., Harrison, C., Ogan, A., Agarwal, Y.: Edusense: Practical classroom sensing at scale. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **3**(3), 1–26 (2019)
2. Bingham, G.E., Okagaki, L.: Ethnicity and Student Engagement, pp. 65–95. Springer US, Boston, MA (2012). https://doi.org/10.1007/978-1-4614-2018-7_4
3. Bird, S., Klein, E., Loper, E.: Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. O'Reilly Media, Inc. (2009)
4. Blandford, A., Furniss, D., Makri, S.: Qualitative HCI research: Going behind the scenes. Morgan & Claypool Publishers (2016)
5. Bonilla-Silva, E.: What makes systemic racism systemic? Sociological Inquiry **91**(3), 513–533 (2021). https://doi.org/10.1111/soin.12420
6. Borchers, C., Liu, X., Lee, H.H., Zhang, J.: Ethical AIED and AIED Ethics: Toward Synergy Between AIED Research and Ethical Frameworks. In: Olney, A.M., Chounta, I.A., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds.) Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky. pp. 18–31. Springer Nature Switzerland (2024). https://doi.org/10.1007/978-3-031-64315-6_2
7. Bosch, N., Chen, Y., D'Mello, S.: It's written on your face: Detecting affective states from facial expressions while learning computer programming. In: International Conference on Intelligent Tutoring Systems. pp. 39–44. Springer (2014)
8. Buckingham Shum, S., Ferguson, R., Martinez-Maldonado, R.: Human-centred learning analytics. Journal of Learning Analytics **6**(2), 1–9 (2019)
9. Buddemeyer, A., Hatley, L., Stewart, A., Solyst, J., Ogan, A., Walker, E.: Agentic engagement with a programmable dialog system. In: Proceedings of the 17th ACM Conference on International Computing Education Research. pp. 423–424. ICER 2021, Association for Computing Machinery, New York, NY, USA (2021), https://doi.org/10.1145/3446871.3469782
10. Collins, P.H.: Black feminist thought in the matrix of domination. Black feminist thought: Knowledge, consciousness, and the politics of empowerment **138**(1990), 221–238 (1990)
11. Collins, P.H.: Intersectionality as Critical Social Theory. Duke University Press (2019)
12. Costanza-Chock, S.: Design justice: Community-led practices to build the worlds we need. The MIT Press (2020)
13. Diem, S., Welton, A.D., Brooks, J.S.: Antiracism education activism: A theoretical framework for understanding and promoting racial equity. AERA Open **8**, 23328584221126518 (2022). https://doi.org/10.1177/23328584221126518
14. D'ignazio, C., Klein, L.F.: Data feminism. MIT press (2023)
15. Dixson, M.D.: Measuring Student Engagement in the Online Course: The Online Student Engagement scale (OSE). Online Learning **19**(4), n4 (2015)
16. D'Mello, S.K.: Improving student engagement in and with digital learning technologies. Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots pp. 79–104 (2021)
17. Fraser, N.: Ten things we've learned from blockly. In: 2015 IEEE Blocks and Beyond Workshop (Blocks and Beyond). pp. 49–50 (2015). https://doi.org/10.1109/BLOCKS.2015.7369000

18. Gebru, T.: Dr. timnit gebru: Hierarchy of knowledge in machine learning & related fields and its consequences [YouTube Video]. https://www.youtube.com/watch?v=OL3DowBM9uc
19. Hammer, D., Berland, L.K.: Confusing claims for data: A critique of common practices for presenting qualitative research on learning. Journal of the Learning Sciences **23**(1), 37–46 (2014)
20. Hayati, H., Idrissi, M.K., Bennani, S.: Automatic classification for cognitive engagement in online discussion forums: Text mining and machine learning approach. In: International Conference on Artificial Intelligence in Education. pp. 114–118. Springer (2020)
21. Henderson, N., Rowe, J., Paquette, L., Baker, R.S., Lester, J.: Improving affect detection in game-based learning with multimodal data fusion. In: Bittencourt, I.I., Cukurova, M., Muldner, K., Luckin, R., Millán, E. (eds.) Artificial Intelligence in Education. pp. 228–239. Springer International Publishing, Cham (2020)
22. Holmes, W.: Aied—coming of age? International Journal of Artificial Intelligence in Education **34**(1), 1–11 (2024)
23. 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., et al.: Ethics of ai in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education pp. 1–23 (2022)
24. Iacobucci, D., Posavac, S.S., Kardes, F.R., Schneider, M.J., Popovich, D.L.: The median split: Robust, refined, and revived. Journal of Consumer Psychology **25**(4), 690–704 (2015). https://doi.org/10.1016/j.jcps.2015.06.014
25. Ladson-Billings, G.: Toward a theory of culturally relevant pedagogy. American Educational Research Journal **32**(3), 465–491 (1995)
26. Lawrence, L., Echeverria, V., Yang, K., Aleven, V., Rummel, N.: How teachers conceptualise shared control with an ai co-orchestration tool: A multiyear teacher-centred design process. British Journal of Educational Technology **55**(3), 823–844 (2024). https://doi.org/10.1111/bjet.13372
27. Leyva, L.A., McNeill, R.T., Duran, A.: A queer of color challenge to neutrality in undergraduate stem pedagogy as a white, cisheteropatriarchal space. Journal of Women and Minorities in Science and Engineering **28**(2), 79–94 (2022)
28. Li, W., Sun, K., Schaub, F., Brooks, C.: Disparities in students' propensity to consent to learning analytics. International Journal of Artificial Intelligence in Education **32**(3), 564–608 (2022)
29. Li, Y., Nwogu, J., Buddemeyer, A., Solyst, J., Lee, J., Walker, E., Ogan, A., Stewart, A.E.: "I Want to Be Unique From Other Robots": Positioning Girls as Co-creators of Social Robots in Culturally-Responsive Computing Education. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. CHI '23, Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3544548.3581272
30. Lin, L.C., Hung, I.C., Chen, N.S., et al.: The Impact of Student Engagement on Learning Outcomes in a Cyber-Flipped Course. Educational Technology Research and Development **67**(6), 1573–1591 (2019)
31. Martin, A.E., Fisher-Ari, T.R.: "If We Don't Have Diversity, There's No Future to See": High-school students' perceptions of race and gender representation in STEM. Science Education **105**(6), 1076–1099 (2021). https://doi.org/10.1002/sce.21677
32. Mason, M.: Sample size and saturation in phd studies using qualitative interviews. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research. vol. 11 (08 2010)

33. Merriam-Webster Incorporated: Equity. https://www.merriam-webster.com/dictionary/equity (2025)
34. Merriam-Webster Incorporated: Inequity. https://www.merriam-webster.com/dictionary/inequity (2025)
35. Miles, M.B., Huberman, A.M.: Qualitative data analysis: An expanded sourcebook. Sage (1994)
36. Nkrumah, T.: The Inequities Embedded in Measures of Engagement in Science Education for African American Learners from a Culturally Relevant Science Pedagogy Lens. Education Sciences **13**(7) (2023). https://doi.org/10.3390/educsci13070739
37. Ocumpaugh, J., Roscoe, R.D., Baker, R.S., Hutt, S., Aguilar, S.J.: Toward asset-based instruction and assessment in artificial intelligence in education. International Journal of Artificial Intelligence in Education. pp. 1–40 (2024)
38. Pargman, T.C., McGrath, C., Viberg, O., Knight, S.: New vistas on responsible learning analytics: A data feminist perspective. Journal of Learning Analytics **10**(1), 133–148 (2023)
39. Reich, J.: Failure to disrupt: Why technology alone can't transform education. Harvard University Press (2020)
40. Reschly, A.L., Christenson, S.L.: Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct, pp. 3–19. Springer US, Boston, MA (2012). https://doi.org/10.1007/978-1-4614-2018-7_1
41. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., Kafai, Y.: Scratch: programming for all. Commun. ACM **52**(11), 60–67 (Nov 2009). https://doi.org/10.1145/1592761.1592779
42. Scott, K.A., Sheridan, K.M., Clark, K.: Culturally responsive computing: A theory revisited. Learning, Media and Technology **40**(4), 412–436 (2015)
43. Sidney D'Mello, E.D., Duckworth, A.: Advanced, analytic, automated (aaa) measurement of engagement during learning. Educational Psychologist **52**(2), 104–123 (2017). https://doi.org/10.1080/00461520.2017.1281747, PMID: 29038607
44. Solyst, J., Nkrumah, T., Stewart, A.B., Lee, J., Walker, E., Ogan, A.: Understanding Instructors' Cultivation of Connectedness in K-12 Online Synchronous Culturally Responsive STEM and Computing Education. Proc. ACM Hum.-Comput. Interact. **6**(CSCW2) (Nov 2022). https://doi.org/10.1145/3555759
45. Stewart, A., Bosch, N., D'Mello, S.K.: Generalizability of Face-Based Mind Wandering Detection across Task Contexts. International Educational Data Mining Society (2017)
46. Stewart, A., Mills, C., Hutt, S.: Marrying asset- and deficit-based approaches: A data feminist perspective in learning analytics. CEUR Workshop Proceedings **3667**, 292–294 (2024), publisher Copyright: © 2024 CEUR-WS. All rights reserved.; 2024 Joint of International Conference on Learning Analytics and Knowledge Workshops, LAK-WS 2024 ; Conference date: 18-03-2024 Through 22-03-2024
47. Stewart, A.E.B., Solyst, J., Buddemeyer, A., Hatley, L., Henderson-Singer, S., Scott, K., Walker, E., Ogan, A.: Explaining Engagement: Learner Behaviors in a Virtual Coding Camp. In: Roll, I., McNamara, D., Sosnovsky, S., Luckin, R., Dimitrova, V. (eds.) Artificial Intelligence in Education. pp. 338–343. Springer International Publishing, Cham (2021)
48. Young, D.S., Casey, E.A.: An Examination of the Sufficiency of Small Qualitative Samples. Social Work Research **43**(1), 53–58 (10 2018). https://doi.org/10.1093/swr/svy026
49. Zoom Video Communications, Inc.: Zoom. https://zoom.com/