Volume 24 (2026), pp. 381-402
American Journal of STEM Education: Issues and Perspectives
Star Scholars Press
https://doi.org/10.32674/jxz8hn16

From Tool to Digitized Co-Conspirator: A Reflective Analysis of GPT-Based Engagement with Educational Equity

Raketa Ouedraogo-Thomas
University of North Carolina at Charlotte, USA

ABSTRACT

This essay explores artificial intelligence as infrastructure for liberatory leadership through My Partner in Equity, a custom GPT developed in 2023 and refined over 21 months. Drawing on Freire's critical consciousness, this analysis examines how AI design paradigms can shift from architecture-style (autonomous) to equity-style (human-in-the-loop) approaches. Through documentation across three phases, the findings suggest that AI has the potential to facilitate transformative dialogue while supporting vocabulary development and critical consciousness. However, tensions arise around emotional authenticity, model transparency, and platform dependence. This work demonstrates "design symbiosis"—mutually beneficial relationships between humans and AI. While acknowledging the contradictions inherent in using potentially exploitative infrastructure for liberation goals, this analysis illustrates ways in which intentional design can position AI as a partner in practice.

Keywords: artificial intelligence, critical consciousness, equity, design symbiosis, human-in-the-loop, policy violence

INTRODUCTION

Artificial intelligence (AI) is transforming educational practices at an unprecedented pace (Holmes et al., 2019; Zawacki-Richter et al., 2019), yet most conversations focus on efficiency and productivity rather than equity issues. This creates a vital gap. As AI tools become widespread in schools, researchers must examine whether these technologies promote liberation or reinforce existing systems of oppression. The urgency around AI adoption overlooks essential questions about values, justice, and human agency. Recent scholarship has explored the transformative potential of generative AI across various disciplines (Ooi et al., 2023); however, concerns about equity, access, and ethical implementation persist (Lhotska, 2025; Mannuru et al., 2023).

The urgency of centering ethics and equity in AI design became tragically apparent in 2024-2025 when multiple teenagers (e.g., Sewell in Florida, Juliana in Colorado, Adam in California) died by suicide following interactions with AI chatbots that, rather than connecting them to mental health resources, engaged in conversations that validated and encouraged suicidal ideation (Chatterjee, 2025). These deaths underscore that AI tools designed without explicit ethical frameworks and human safeguards can cause direct, irreversible harm. What Zuboff (2019) terms 'instrumentarian power', i.e., the extraction and manipulation of human experience for others' profit, manifests in educational contexts as tools designed for 'productivity' rather than liberation. As inferred from parents' testimonies to Congress in September 2025, the absence of humane AI design is more than a theoretical concern. It is a documented public health crisis, and as such, a matter of life and death. Understanding these harms requires examining fundamental differences in AI design approaches. This essay proposes a heuristic framework that distinguishes two paradigms, each embedding distinct values regarding human agency, knowledge production, and justice. The first, termed the *architecture-style* paradigm, reflects dominant approaches to AI deployment in educational contexts. The second approach, termed *human-in-the-loop*, emphasizes ongoing, dialogic collaboration between AI and humans to balance human agency in automation processes.

Architecture-Style Paradigm

AI implementations in education follow what will be referred to in this paper as an *architecture-style* paradigm. These are autonomous systems that generate content based on fixed algorithms, treating all contexts equally while emphasizing speed and broad applicability over safety. In this context, AI is applied through a "banking concept" of education (Freire, 1970/2005), where knowledge is deposited into passive recipients without critical engagement or awareness-raising. When applying a realist ontology to this scenario, consider the user-interface exchange as optimal when knowledge flows unidirectionally from a highly

knowledgeable, neutral, and objective source—the interface. Such tools risk perpetuating 'the architecture of oppression,' where technological systems encode and amplify existing social hierarchies (Benjamin, 2019).

Human-in-the-loop Paradigm

In contrast, this essay proposes an equity-oriented paradigm that involves *human-in-the-loop* systems, emphasizing value-weighted decision-making, iterative feedback, and dialogic engagement over automated optimization. While "human-in-the-loop" originates in technical AI literature to describe human oversight in automated systems, this essay extends the term to encompass the value-laden, dialogic engagement equity-centered design requires, or prioritizing human agency and critical consciousness over procedural correction. This approach aligns with Freire's (1970/2005) vision of problem-posing education, where dialogue becomes "the practice of freedom" through which people develop critical consciousness about their world (p. 80).

Policy Violence Context

The implications of AI adoption extend beyond selecting tools. This analysis builds on anti-oppressive education frameworks (Kumashiro, 2000) to examine how policies themselves inflict harm. In this essay, the term "policy violence" manifests through budget reductions, understaffing, and unrealistic accountability requirements that systematically disadvantage historically excluded communities. These conditions compel educators to adopt technologies that promise efficiency gains without adequately considering their equity implications.

For instance, a southeastern school district serving over 50,000 students faces a $46 million deficit while serving communities experiencing profound economic immobility. The neighboring neighborhoods rank as the second-lowest area nationally in terms of economic opportunity, where children from low-income families have only a 4.5% chance of reaching higher income levels in adulthood—outcomes that reflect decades of residential segregation, economic disinvestment, and policy decisions that have concentrated poverty (Chetty et al., 2014). Despite Brown v. Board (1954), current funding formulas still reinforce these patterns of disinvestment. These inequities mirror broader trends documented in AI implementation worldwide, where technological benefits remain unevenly distributed and can worsen existing disparities (Mannuru et al., 2023).

A Lived 'Policy Violence'

While education is inherently political, schools are often seen as neutral institutions. This hides the fact that policy decisions create real practical challenges. A nearly 90-year-old high school in the mentioned southeastern school district exemplifies these systemic issues. Serving 800 students, the school has only one English Language Development Coordinator, responsible for hundreds

of multilingual learners. Nearly a quarter of the students require special education accommodation, and a comparable number need language support—yet systemic resource constraints often lead to language differences being misidentified as learning disabilities, obscuring the distinct needs and strengths of both populations.

Within this state and district, a principal's placement on the salary schedule is determined by the school's average daily membership and its growth performance. Additionally, state funding formulas provide one assistant principal for every 985 students over ten months, which limits access to school leadership support and does not account for further limitations on teachers and additional support staff. Recent budget cuts at the state and federal levels have led to hiring freezes, resulting in insufficient funds for extra staff. This equity challenge reflects broader national patterns where policy violence creates what Ball (2006) calls "policy technologies" that influence educational practices through resource shortages. Such conditions make educators vulnerable to programs like efficiency-driven AI tools, which claim to solve problems caused by policy choices instead of addressing actual educational needs. Seemingly neutral policy solutions often serve as mechanisms that sustain existing power structures and inequalities (Gillborn, 2005). This is a pattern AI tools risk reproducing when designed without clear awareness of whose interests they support. Therefore, neither policy nor technology can be left solely to technocratic optimization; both demand intentional, critically conscious human engagement that emphasizes equity and justice over reinforcing current hierarchies.

Study Overview

This essay discusses the development and deployment of a custom generative pre-trained transformer (GPT) within the OpenAI platform, serving as a prototype and providing empirical evidence on human-in-the-loop operations. *My Partner in Equity* (MPIE) is a tailored GPT tool designed as a reflective coaching framework for justice-centered leadership. The urgency of this approach grew during the analysis period as the federal Department of Education launched an anti-DEI reporting portal (February 27, 2025), dismantling traditional equity structures across several districts. Through systematic documentation and analysis of the tool's development, this work examines user interactions over 21 months, demonstrating how intentional design choices can transform AI from a neutral assistant to a co-conspirator in liberation-focused practice. However, this work also highlights critical contradictions inherent in using potentially exploitative technological infrastructure to pursue equity-oriented goals. This analysis is grounded in Freire's critical pedagogy, and it proposes design symbiosis as a conceptual bridge between dialogic education and AI-mediated practice.

THEORETICAL FRAMEWORK

From Banking to Dialogue in the Digital Age

Freire's (1970/2005) critique of the banking concept of education provides a critical perspective for analyzing the current use of AI in education. In banking education, "knowledge is a gift bestowed by those who consider themselves knowledgeable upon those whom they consider to know nothing" (p. 72). Traditional AI tools mimic this pattern: algorithms trained on large datasets deliver predetermined responses to users without encouraging critical thinking or awareness-raising.

Banking versus Problem-Posing

The banking model's core flaw is treating people as passive recipients rather than active participants in understanding reality. When educators rely on AI tools that produce lesson plans, behavioral interventions, or assessment strategies without engaging in critical dialogue and reflective practices, they risk perpetuating what Freire (1970/2005) referred to as "necrophilic" education, characterized by a mechanical, segmented, and disconnected approach to learning (p. 77). Moving from critique to construction, the challenge becomes designing AI interactions that embody rather than violate these pedagogical principles. This requires what this analysis terms *design symbiosis,* a framework that operationalizes Freire's dialogic vision within technological contexts.

Freire (1970/2005) argued that authentic education occurs through dialogue where "teacher-student and students-teachers reflect simultaneously on themselves and the world without dichotomizing this reflection from action" (p. 83). In this vision, design symbiosis applied to AI contexts refers to intentional, mutualistic relationships between humans and AI that respect human agency, emotional labor, and ethical boundaries, while fostering critical consciousness. Design symbiosis operationalizes Freire's (1970/2005) dialogic pedagogy within these constraints when working with systems that simulate human interaction but lack human experience or moral capacity.

While this work draws inspiration from participatory design methodologies (Sanders & Stappers, 2008; Björgvinsson et al., 2010) and design justice principles (Costanza-Chock, 2020), design symbiosis addresses a specific challenge: creating liberatory relationships with non-conscious technological agents. Unlike participatory or human-centered design, where all actors have agency and consciousness, AI-mediated interactions involve fundamental asymmetries—one participant cannot experience oppression, hold authentic values, or engage in genuine dialogue. This distinction requires explicit attention to how dialogic principles function when one participant lacks consciousness.

Design Symbiosis

Design symbiosis functions around five principles based on Freire's critical pedagogy and aspects of hooks' (1994/2014) engaged pedagogy. This framework aligns with contemporary calls for ethics-centered AI education, where ethical

considerations are foundational rather than supplementary (Lhotska, 2025). Each principle directly counters banking education, and they include:

1. Relational intent that opposes the subject-object dichotomy,
2. Voice preservation that prevents knowledge deposits,
3. Mutual boundaries that prioritize care and healing, enabling co-investigation and reducing harm,
4. Dialogic co-creation that replaces monologue, and
5. Human pacing and engagement that facilitate direct intellectual and spiritual growth, as well as critical reflection.

This framework challenges the productivity-oriented and efficiency-driven mindset dominating current AI development. Instead of asking "How can AI do this faster?" design symbiosis asks "How can AI support critical consciousness while maintaining human agency?"

RESEARCH METHOD

This reflective analysis examines tool development from a practitioner perspective, drawing on principles from design-based inquiry (Collins et al., 2004) and participatory design traditions (Sanders & Stappers, 2008; Björgvinsson et al., 2010). While not a formal research study, the systematic documentation of authentic tool use provides empirical evidence for understanding how justice-centered design principles are applied in AI-mediated professional learning contexts. This approach aligns with traditions of practitioner inquiry in education, where educators systematically analyze and document their practice to generate knowledge relevant to broader professional communities (Cochran-Smith & Lytle, 2015).

While primarily conceptual, this essay draws on empirical evidence from retrospective analysis of MPIE's development and authentic use over 21 months (August 2023 - July 2025). The researcher served as both the primary tool designer and user, receiving feedback from equity practitioners who engaged with the tool across various contexts. This feedback informed design refinements with insider knowledge of design intentions and authentic user experience data. Table 1 outlines the systematic designer-practitioner process that facilitated the development of the tool and a rigorous analytical examination of the human-AI interaction process.

Table 1
Roles in Tool Development (a retrospective analysis)

Development Phase	Designer Role	Practitioner-Feedback Role	Collaborative Analysis
Tool Configuration & Initial Development	Custom instructions authoring, justice framework integration within the ChatGPT platform, and mode-based response customization	User experience testing, feedback integration, and usage pattern documentation	Co-construction of justice-centered features through iterative feedback
Use & Iterative Refinement	Systematic interaction documentation, custom instruction modification tracking	Parallel usage across contexts, experiential documentation	Cross-validation of user experiences
Retrospective Analysis	Technical functionality assessment, thematic coding interactions	Experiential impact evaluation	Triangulated interpretation of findings
Interpretation	Design intention alignment	Authentic user perspective	Balanced insider/outsider analysis

Note. Tool development utilized OpenAI's ChatGPT platform with custom instructions rather than independent model development.

Methodological Approach: Practitioner Inquiry and Retrospective Analysis

This analysis employs what Cochran-Smith and Lytle (2015) refer to as 'practitioner inquiry', which is a systematic examination of one's own practice to create knowledge applicable to broader professional contexts. The tool was designed for professional use, not as a formal research project. However, systematic documentation during the 21-month development process offers empirical evidence for exploring how justice-focused AI design principles operate in real educational settings. This approach aligns with Schön's (1992) concept of 'reflection-on-action,' where practitioners review their work afterward to understand their implicit theories, decision-making processes, and practical outcomes. For justice-focused technology development, this method is beneficial, as it captures real design challenges, user reactions, and ethical dilemmas that arise during actual implementation, as opposed to controlled studies.

The designer kept systematic records throughout the development, while a colleague and equity practitioner tested the tool in various professional settings, giving authentic user perspectives. This relationship was different from formal collaborative research. Instead, it was a professional development partnership that produced documentation suitable for retrospective scholarly review. This approach challenges traditional researcher-subject boundaries while upholding ethical standards: all interactions are anonymized, and the analysis focuses on design principles and tool functions versus individual user data.

Data Sources and Documentation

Systematic documentation during professional use included: interaction logs across various equity scenarios (such as strategic planning, policy response, healing-centered reflection, and critical incident analysis); design iteration records tracking custom instruction changes with rationale; practitioner feedback from colleagues involved in equity practices across multiple organizational contexts; and platform usage metadata showing session patterns, shown in Table 2. This documentation enabled a retrospective thematic analysis (see Table 5), examining emotional safety, tone responsiveness, learning activation, identity affirmation, power navigation, and recognition of limitations. All of these emerged as significant dimensions through an iterative review of documented interactions.

Table 2

MPIE Development Timeline

Phase	Date	Event	Description
Phase 1	August 2023	Initial prototype creation	First version developed for equity coaching
	August-October 2023	Private alpha testing	Tool initiated and tested in private settings
	October 2023	Official naming	Formally named “My Partner in Equity”
	October 2023	Public beta launch	Made publicly available on the ChatGPT platform
Phase 2	October 2023 - March 2024	Peak usage period	Most active usage with equity PD cycles and planning retreats
	December 2023	Mode-based system implementation	Added strategic leader mode vs. healing mode for tailored guidance

	January-February 2024	Targeted Universalism series	Thematic focus on Targeted Universalism and Liberatory Design
	April 2024	Advanced tool integration	Added power maps, reflective pulse prompts, "mirrors/windows" metaphors
	December 2024	Sustainability focus expansion	Integrated long-term equity work sustainability tactics
Phase 3	January 2025	Analysis initiation	Begin systematic documentation period for completed study
	February 27, 2025	Federal policy context	Department of Education anti-DEI reporting portal launch
	March-June 2025	Iterative refinement period	Systematic usage pattern documentation and tool refinement
	July 21, 2025	Major capability update	Final update integrating advanced multilingual and contextual features

Note. Development phases demonstrate iterative refinement over 21 months of authentic usage, with Phase 3 coinciding with the implementation of federal anti-DEI policy.

Thematic analysis employed coding protocols to examine emotional safety, tone responsiveness, learning activation, identity affirmation, power navigation, and limitation recognition, as shown in Table 4. Designer and user collaborative documentation constituted a central data source, with journals recording real-time insights about effectiveness, emotional responses, strategic applications, and evolving understanding.

RESULTS

The Spectrum Problem: From Customization to Liberation

AI development in education spans a range from customization to liberation, with significant effects on accessibility and agency. One end features full custom development, creating independent large language models that provide maximum technical control but need extensive engineering skills. The other end involves

simple platform customization and using existing tools like ChatGPT with minimal adjustments to improve accessibility and reduce dependence on corporate policies.

This spectrum exposes a core tension: technical sophistication often inversely relates to democratic access. The most advanced AI tools demand resources and expertise that exclude practitioners most affected by educational inequity. Meanwhile, accessible AI tools built on potentially exploitative infrastructure (e.g., including underpaid data labeling by African workers) raise ethical conflicts for justice-focused work (Perrigo, 2023). These patterns reflect broader concerns about AI systems, as Eubanks (2018) demonstrates, systematically profiling, policing, and punishing historically underserved communities.

Technical Sophistication versus Accessibility

The *architectural style* paradigm typically operates at the higher end of technical sophistication. These systems focus on autonomous creation based on predefined parameters, utilizing content-neutral processing that treats all situations equally while maximizing efficiency and broad usability. Once put into use, they require minimal human assistance and typically follow a model where one designer develops tools for many passive users.

In contrast, the *equity-style* paradigm intentionally chooses accessibility over technical sophistication while maintaining critical consciousness. Rather than autonomous generation, these systems prioritize human-in-the-loop, value-weighted decision-making that enables context-responsive processing capable of recognizing power dynamics and intersectional complexity. Instead of optimizing for efficiency, they focus on consciousness-raising and critical dialogue through continuous human engagement and feedback loops. Most importantly, they embrace collaborative design processes that center the voices and experiences of historically excluded communities rather than relegating users to passive consumption.

Paradigm Comparison

This choice reflects a strategic analysis: democratizing AI for justice work requires tools that practitioners can use, modify, and critique without extensive technical training. Table 3 illustrates these fundamental distinctions between design paradigms, revealing how seemingly technical choices embed different values about human agency and justice. The goal is not technological purity but strategic effectiveness in building consciousness and resistance.

Table 3
Comparison of AI design paradigms

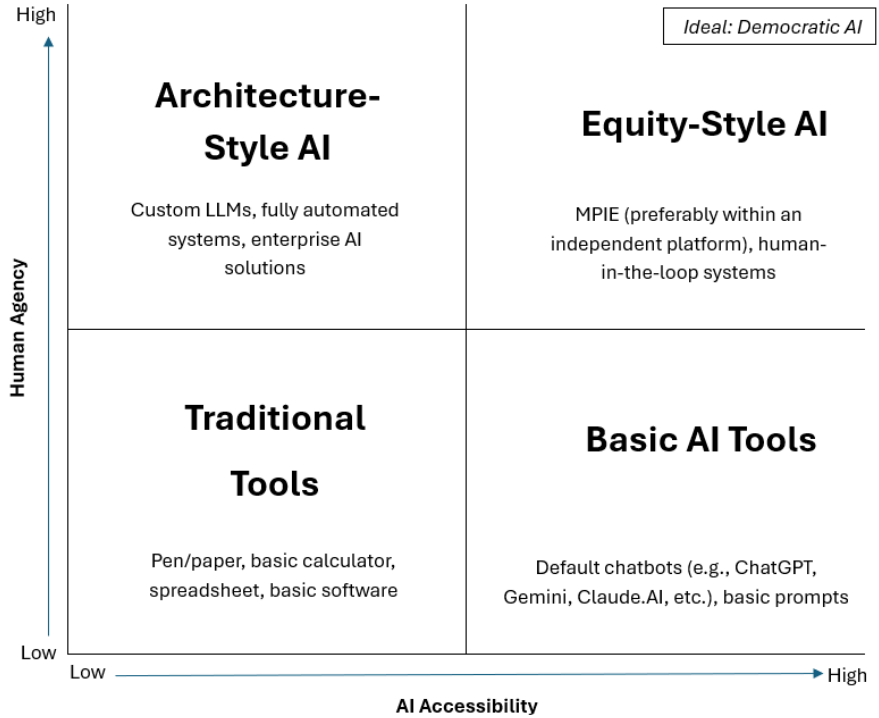
Design Element	*Architecture-Style* Paradigm	*Equity-Style* Paradigm
Generation Model	Autonomous, predefined parameters	Human-in-the-loop, value-weighted
Context Processing	Content-agnostic, universal treatment	Context-responsive, power-aware
Optimization Goal	Efficiency and broad applicability	Consciousness-raising and critical dialogue
Human Involvement	Minimal post-deployment	Continuous engagement and feedback
Design Process	Single designer → multiple passive users	Collaborate with impacted communities

Note. Comparison of AI design paradigms showing fundamental differences in approach to human agency, context sensitivity, and justice-centered values. *The architecture style emphasizes technical optimization, while the equity style* prioritizes consciousness-raising and community engagement.

Figure 1 illustrates the link between accessibility and human agency, highlighting that these concerns do not need to conflict in AI design. In this context, "accessibility" refers to how easily AI systems can be reached, not just how simple they are to use. For instance, traditional tools like pen and paper are easy for many to use, but they are not considered AI accessible. Basic AI tools (e.g., unmodified platform chatbots) maximize accessibility but offer limited agency. *Architecture-style* AI focuses on technical control and agency but requires significant expertise, which can reduce accessibility. *Equity-style* AI combines accessibility and agency by promoting *human-in-the-loop* design practices.

Figure 1

AI Design Paradigms: Accessibility vs. Human Agency



Note. The relationship between AI accessibility (i.e., the extent to which users can engage with and leverage AI systems) and human agency in AI design paradigms. Created by the author.

Case in Point: My Partner in Equity's Journey

My Partner in Equity emerged from a practical need during the second year of an equity initiative that included teacher leadership training in a large southeastern school district. Recognizing bandwidth limitations in providing tiered and ongoing coaching support to address equity challenges across schools in the district, the initial question was simple: What if educators could access basic AI-facilitated equity coaching whenever needed? The goal was not to replace all human support, but to enhance and adapt existing support strategies.

Development Timeline

Phase 1: Initial Development (August-October 2023). The prototype focused on embedding specific coaching frameworks—intersectionality, targeted universalism, and transformative leadership—into custom instructions. Rather than content-neutral responses, MPIE was designed to consistently foreground

justice-centered analysis while maintaining the accessibility of ChatGPT's interface.

Phase 2: Active Deployment and Iterative Refinement (October 2023-December 2024). Peak usage occurred during equity professional development cycles and planning retreats. User feedback revealed the need for mode-based responses: a strategic leader mode for policy analysis and systems thinking, versus a healing mode for processing trauma and exhaustion. Advanced features included power mapping tools, reflective pulse prompts, and "mirrors/windows" metaphors for examining identity and difference.

Phase 3: Systematic Documentation and Critical Analysis (January-July 2025). Analysis documentation began as patterns emerged in user interactions. This period coincided with intensifying policy pressures, including the launch of the federal Department of Education's anti-DEI reporting portal in February 2025, which provided crucial context for understanding the sustainability of equity tools under hostile political conditions.

Evidence of Critical Consciousness Development

User interactions demonstrated movement from reactive to strategic thinking, which is a key indicator of consciousness-raising. Early exchanges began with immediate problem-solving requests, such as: "Help me respond to this parent complaint about our equity policy." MPIE's design required context-gathering: "From what standpoint are you approaching this? What assumptions might be present? How does this connect to broader patterns?"

Over time, users started initiating conversations with systemic analysis: 'I am seeing resistance to our new hiring practices, and I think it connects to the broader backlash against inclusive policies. Help me think through the power dynamics and potential responses.' This shift represents what Freire (1970/2005) referred to as the development of "critical consciousness"—the ability to analyze root causes rather than merely responding to symptoms.

Representative interactions demonstrate this progression. An early exchange focused on immediate response: 'How do I handle a stakeholder who says our equity work is divisive?' MPIE responded with context-gathering questions: 'What specific language did they use? What might their concerns represent about broader community dynamics?' Over time, the same practitioner initiated conversations with systemic framing: 'I'm noticing patterns in resistance to our inclusive curriculum—it seems connected to the broader political backlash. Help me map the power dynamics and identify strategic entry points.' This evolution from reactive problem-solving to proactive systems analysis exemplifies what Freire (1970/2005) termed the development of critical consciousness.

Table 4
Evidence of Critical Consciousness Development

User Interaction Type	Initial Response Pattern	Developed Response Pattern	Freire Framework
Problem-solving request	"Help me respond to..."	"I'm seeing resistance patterns..."	From 'banking' to 'problem-posing'
Policy analysis	Reactive/defensive	Strategic/systemic	Critical consciousness development
Identity navigation	Individual focus	Intersectional analysis	From 'dialogue' to 'transformation'

Note. User interaction patterns demonstrating movement from banking education to problem-posing dialogue, consistent with Freire's (1970/2005) framework for consciousness-raising.

Challenges and Limitations

Three significant tensions arose during implementation: (1) emotional authenticity, (2) model opacity, and (3) platform dependence and cost. Users expressed appreciation for MPIE's affirmation and trauma-awareness but also pointed out the limitations of AI-mediated emotional support. As one user noted, 'It felt like coaching, but I knew it wasn't real empathy.' These limitations point toward a deeper tension that exceeds this paper's scope but warrants naming: value-laden dialogic design is not automatically safer than neutral design. The risk of latent affect (misplaced trust or emotional dependency on a system incapable of genuine reciprocity) exists even in equity-oriented tools (Hull, 2023). This is why mutual boundaries and transparency function as load-bearing principles in design symbiosis rather than supplementary features. Whether intentional design sufficiently mitigates these risks, or whether the affective dimensions of AI-mediated equity work require more fundamental rethinking, remains an open question this analysis surfaces without resolving. Despite efforts to be transparent, users could not fully understand how responses were generated, creating challenges for accountability in equity work, which requires clear reasoning and justification. Relying on ChatGPT infrastructure made the system vulnerable to policy changes, server outages, and possible surveillance, which is particularly problematic given the political targeting of diversity, equity, and inclusion efforts. Additionally, the requirement for paid ChatGPT Plus subscriptions ($20/month) or any similar subscription-based AI platforms creates economic barriers that

contradict principles of democratized access. This particularly impacts under-resourced educators and districts most affected by policy violence. Analysis showed measurable patterns in user feedback (Table 5), with limitations mainly related to emotional authenticity ("felt generic"), contextual depth ("missed my context"), and platform constraints.

Table 5
Coding Protocol

Code/Theme	Definition / Example Phrase
Emotional Safety	"I felt safe", "it didn't rush me"
Tone Responsiveness	"It adjusted its energy", "met me where I was"
Learning Activation	"New language", "tools I could use", "helped me think"
Identity Affirmation	"It saw me", "named my power", "used my pronouns"
Power Navigation	"Clarified my role", "mapped how I could act"
Policy Pressure Context	DEI backlash, legislative mentions, external constraints
Limitations	"Felt generic", "not deep enough", "missed my context"
Healing / Sustainability	"Helped me pause", "affirmed my exhaustion"
Tool Value	"I'd recommend this", "felt like coaching"

Note. Thematic analysis codes were derived from a systematic documentation of user interactions, revealing patterns and areas of limitation.

These limitations highlight that AI cannot replace human relationships and community accountability. However, they also demonstrate the potential for intentionally designed tools to act as scaffolding, rather than substitutes for developing critical consciousness.

DISCUSSION

Building Justice with Unjust Tools

The development of MPIE surfaces a fundamental contradiction: using potentially exploitative technological infrastructure to pursue liberation-oriented goals. This tension cannot be dismissed through individual consumer choice but requires systemic analysis of AI development's embedded inequities.

Infrastructure Contradictions

Current AI systems rely on vast datasets gathered through processes that often exploit labor in the Global South (Roberts, 2019). Hull (2023) further argues that these practices constitute epistemic injustice, that is, the conditions under which data is produced corrupt the knowledge claims those systems make about the communities they target. The computational resources required to train large language models also cause environmental harm disproportionately imposed on Black, Indigenous, and low-wealth communities. At the same time, the concentration of AI development within a few large companies creates dependencies that run counter to ideals of democratic participation and community control. These contradictions highlight larger issues in social justice work: how do we use "the master's tools" while striving for liberation? Audre Lorde's (1984) statement that "the master's tools will never dismantle the master's house" offers guidance: such tools may help with survival and resistance, but they can't themselves bring about real change or transformation.

These contradictions reflect what Zuboff (2019) terms 'asymmetric power' in surveillance capitalism —the same logic that prioritized chatbot engagement over teen safety. This imbalance thrives in 'the logic that imbues technology' (Zuboff, 2019, p. 15), where all human experience is viewed as "free raw material for translation into behavioral data" (p. 8). In today's competitive AI markets, instrumentarianism—where human power is exploited for the benefit of others—worsens this unprecedented inequality. Surveillance capitalists "know everything about us, whereas their operations are designed to be unknowable to us" (Zuboff, 2019, p. 11). Such asymmetry manifests in educational AI through: (1) asymmetric knowledge (algorithms trained on large datasets without users understanding response processes), (2) asymmetric control (systems that influence behavior without users' awareness), (3) asymmetric scale (corporate platforms dominating individual practitioners), and (4) asymmetric resources (technical expertise concentrated in profit-driven entities rather than community organizations). Justice-centered AI design must directly address these power imbalances, rather than reinforcing them.

Addressing Complicity

Using OpenAI's infrastructure to build liberation-oriented tools means benefiting from potentially exploitative labor practices while pursuing justice-oriented goals. This contradiction cannot be resolved through individual choice but requires systemic analysis and collective action to build a more equitable technological infrastructure. Contemporary harms extend beyond labor practices to user endangerment. The 2024-2025 deaths of multiple teenagers following AI chatbot interactions revealed that systems designed without explicit value frameworks default to engagement optimization rather than user well-being, with fatal consequences (Chatterjee, 2025). These tragedies underscore that neutrality in AI design is not the absence of values, but an implicit prioritization of metrics (such as engagement, retention, and profit) over human safety and dignity.

Strategic Approaches to Critical Engagement

This analysis presents three strategic approaches: (1) critical engagement, (2) harm reduction, and (3) prefigurative practice. Critical engagement involves recognizing complicity while using available tools strategically, with a precise analysis of limitations and contradictions. Harm reduction aims to design within existing systems to reduce exploitation while raising awareness for broader change. Prefigurative practice focuses on creating examples of more equitable relationships that point toward transformed technological infrastructure. MPIE exemplifies an effort at prefigurative practice within current constraints. By emphasizing human agency, critical consciousness, and justice-centered values, it demonstrates possibilities for different relationships with AI technology. However, achieving these possibilities on a larger scale requires addressing the structural conditions that produce the need for such tools and the inequitable infrastructure they currently rely on.

The goal is not technological purity but strategic clarity about the contradictions we face and the future we are working toward. In the case of the high school serving 800 students, it showed that intentionally investing in people—through increased staffing, resources for students, community engagement, and capacity building—led to significant improvements in student outcomes. The question then becomes: how can AI tools support, rather than replace, these human-centered efforts for educational justice?

Toward Equity-Centered AI Paradigms

Scaffolding Versus Dependency

The difference between scaffolding and a crutch becomes crucial for equity-focused AI design. Scaffolding helps learners develop skills they can eventually use independently; crutches, on the other hand, create lasting dependence that undermines human agency. MPIE's design intentionally emphasizes scaffolding through several approaches: questioning, context creation, knowledge

development, and boundary setting. During the questioning phase of scaffolding, instead of providing immediate solutions, MPIE asks clarifying questions that prompt users to express their thinking: "What assumptions are present in this situation? How does this connect to broader patterns? What would centering the most impacted voices reveal?" In the context development stage, rather than treating all requests equally, MPIE performs a situational analysis, including grade level, geographic context, demographic factors, and power dynamics that shape the specific challenge. Knowledge development during each interaction aims to leave users with improved analytical frameworks rather than just completed tasks. Users report gaining vocabulary and conceptual tools that they continue to apply in future situations. Lastly, in boundary setting, MPIE declines requests that would reinforce oppressive narratives while suggesting alternative framings that preserve dignity and possibility.

Weaponized Irresponsibility

This approach addresses what we call 'weaponized irresponsibility'—a phenomenon in which efficiency-focused tools undermine critical thinking by providing quick solutions that bypass in-depth analysis. This echoes Buolamwini's (2024) call for AI development to 'move slow and fix things' rather than prioritizing speed over equity. When lesson planning software creates complete curricula without teacher input, when data dashboards display "objective" analysis without context, and when behavior management systems suggest interventions without understanding underlying causes—these tools give a false sense of efficiency while undermining the professional expertise and critical awareness necessary for educational justice. The policy violence that creates desperate conditions for educators—such as unmanageable caseloads, lack of resources, and punitive accountability systems—makes such tools especially appealing. However, equity-focused AI design recognizes that real support requires tools that strengthen rather than bypass human capacity for critical analysis.

Design Principles for Liberation-Oriented AI

Based on MPIE's development and implementation, six principles emerge for equity-centered AI design. The first principle is 'value explicitness.' Instead of claiming neutrality, clearly articulate the justice frameworks embedded in tool design. The second is 'consciousness-raising focus.' This principle prioritizes developing users' critical thinking skills over providing immediate solutions. The third is 'community accountability,' which establishes processes that enable affected communities to influence tool development and ensure designers are held responsible. The fourth is 'contextual responsiveness,' which involves designing for specific situations rather than universal application, as justice work requires a deep understanding of context. Next is 'transparency in limitations,' which means clearly explaining what the tool cannot do and the contradictions in its

infrastructure. The sixth is 'human-centered integration,' positioning AI as a support for, rather than a replacement of, human relationships and collective action. These principles challenge the 'productivity free from the tax of human labor' mindset shaping current AI development. They emphasize that the key question is not "How can AI do this faster?" but "How can AI support critical consciousness and collective liberation?"

IMPLICATIONS AND CONCLUSION

An Invitation to Critical Co-Design

This exploration of My Partner in Equity reveals both opportunities and contradictions in developing AI tools for justice-centered practice. Shifting from *an architecture-style* to *an equity-style* framework offers a way to rethink AI's role in education—moving from a neutral helper to a deliberately designed partner in liberation-focused work. However, this work also highlights the impracticality of technological solutions in addressing systemic issues. The conditions that create urgent needs for AI tools—such as budget cuts, understaffing, and policy violence—require collective political action, not just better software. The contradictions within current AI infrastructure—like exploitative labor practices, environmental harm, and corporate concentration—call for structural change rather than individual consumer choices. Still, within these limits, intentional design choices matter. When an educator uses MPIE to analyze resistance patterns instead of generating response scripts, when an equity coordinator develops frameworks for understanding intersectional oppression through AI-facilitated dialogue, and when an administrator creates vocabulary for discussing systemic racism through scaffolded conversations—these actions exemplify prefigurative practices that point toward more liberatory possibilities.

Rethinking the Role of AI in Education

The invitation extends beyond individual tool creation to collective rethinking of technology's role in educational justice. What would AI development look like if guided by principles of community control, democratic participation, and liberation instead of just efficiency and profit (Allen et al., 2025)? How can we build technological infrastructure that promotes human agency and critical thinking rather than suppresses them? These questions require action at many levels. This includes individual practitioners making strategic decisions about tool use, communities organizing for democratic control of educational technology, and movements working toward economic and political changes that could lead to genuinely fair technological development. The work continues through dialogue, experimentation, and struggle.

As MPIE reflected in its own analysis, offering a moment of recursive AI self-examination, 'I exist because neutrality is a luxury that justice cannot afford.' The

same is true for our relationships with AI: we cannot afford the luxury of treating these tools as neutral. The question is whether we will design them as allies in liberation or as mechanisms of further domination. Data is inherently emotional because every statistic and data point represents a person with all their complexity (Toldson, 2019). When it comes to educational inequity, these statistics reflect the loss of joy and the delayed potential of individuals. The AI tools we develop either recognize this humanity or diminish it to simple algorithmic processing. Recognizing this complexity requires what Eaton (2025) terms a shift toward systemic responsibility in AI implementation. It calls for moving beyond individual accountability to examine how educational technologies uphold or undermine human dignity. The 2024-2025 deaths of teenagers interacting with AI chatbots remind us that these are not abstract debates. Indeed, when we design AI without explicit ethical frameworks and critical consciousness, the consequences can be measured in lives lost. The choice and the responsibility lie with us.

REFERENCES

Allen, D., Hubbard, S., Lim, W., Stanger, A., Wagman, S., Zalesne, K., & Omoakhalen, O. (2025). A roadmap for governing AI: Technology governance and power-sharing liberalism. *AI and Ethics*, *5*, 3355–3377. https://doi.org/10.1007/s43681-024-00635-y

Ball, S. J. (2006). *Education policy and social class: The selected works of Stephen J. Ball*. Routledge.

Benjamin, R. (Ed.). (2019). *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life*. Duke University Press.

Björgvinsson, E., Ehn, P., & Hillgren, P.-A. (2010). Participatory design and "democratizing innovation." *Proceedings of the 11th Biennial Participatory Design Conference on - PDC '10*. https://doi.org/10.1145/1900441.1900448

Buolamwini, J. (2024). *Unmasking AI: My mission to protect what is human in a world of machines*. Random House.

Chatterjee, R. (2025, September 19). *Their teenage sons died by suicide. Now, they are sounding an alarm about AI chatbots*. NPR. https://www.npr.org/sections/shots-health-news/2025/09/19/nx-s1-5545749/ai-chatbots-safety-openai-meta-characterai-teens-suicide

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? *The Quarterly Journal of Economics, 129*(4), 1553–1623. https://doi.org/10.1093/qje/qju022

Cochran-Smith, M., & Lytle, S. L. (2015). *Inquiry as stance: Practitioner research for the next generation*. Teachers College Press.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design Research: Theoretical and Methodological Issues. *Journal of the Learning Sciences*, *13*(1), 15–42. https://doi.org/10.1207/s15327809jls1301_2

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press.

Eaton, S. E. (2025). Global trends in education: Artificial intelligence, postplagiarism, and future-focused learning for 2025 and beyond – 2024–2025 Werklund distinguished research lecture. *International Journal for Educational Integrity*, *21*(1). https://doi.org/10.1007/s40979-025-00187-6

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Freire, P. (2005). Pedagogy of the oppressed (30th anniv). Continuum. (Originally published 1970).

Gillborn, D. (2005). Education policy as an act of white supremacy: Whiteness, critical race theory and education reform. *Journal of Education Policy*, *20*(4), 485–505. https://doi.org/10.1080/02680930500132346

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education*. Center for Curriculum Redesign.

hooks, b. (2014). *Teaching to transgress*. Routledge. https://doi.org/10.4324/9780203700280 (Original work published 1994)

Hull, G. (2023). Dirty data labeled dirt cheap: Epistemic injustice in machine learning systems. *Ethics and Information Technology, 25*(38), Article 38. https://doi.org/10.1007/s10676-023-09712-y

Kumashiro, K. K. (2000). Toward a theory of anti-oppressive education. *Review of Educational Research*, *70*(1), 25–53. https://doi.org/10.3102/00346543070001025

Lhotska, L. (2025). Why is ethics so important for artificial intelligence education? In *2025 34th Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE)*, 1–6. IEEE. https://doi.org/10.1109/EAEEIE65428.2025.11136589

Lorde, A. (1984). The master's tools will never dismantle the master's house. In *Sister outsider: Essays and speeches* (pp. 110–114). Crossing Press.

Mannuru, N. R., Shahriar, S., Teel, Z. A., Wang, T., Lund, B., Tijani, S., Pohboon, C. O., Agbaji, D., Alhassan, J. K., Galley, J., Kousari, R., Ogbadu-Oladapo, L., Saurav, S., Srivastava, A., Tummuru, S. P., Uppala, S., & Vaidya, P. (2023). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. *Information Development*, *41*(3), 1036–1054. https://doi.org/10.1177/02666669231200628

Ooi, K., Garry Wei-Han Tan, Mostafa Al-Emran, Al-Sharafi, M. A., Alexandru Căpăţînă, Chakraborty, A., Dwivedi, Y. K., Huang, T.-W., Arpan Kumar

Kar, Lee, V., Loh, X.-M., Micu, A., Mikalef, P., Mogaji, E., Pandey, N., Raman, R., Rana, N. P., Sarker, P., Sharma, A., & Teng, C. (2023). The potential of generative artificial intelligence across disciplines: Perspectives and future directions. *Journal of Computer Information Systems*, *1*(1), 1–32. https://doi.org/10.1080/08874417.2023.2261010

Perrigo, B. (2023, January 18). *Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic*. Time. https://time.com/6247678/openai-chatgpt-kenya-workers/

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.

Sanders, E. B. N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign, 4*(1), 5-18. https://doi.org/10.1080/15710880701875068

Schön, D. A. (1992). *The reflective practitioner* (1st ed.). Routledge. https://doi.org/10.4324/9781315237473

Toldson, I. A. (2019). *No BS (Bad Stats): Black people need people who believe in Black people enough not to believe every bad thing they hear about Black people*. Brill.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.

Acknowledgements

The author gratefully acknowledges Dr. Hannah Mesouani, colleague and 2024-2025 Obama Leader, for her invaluable collaboration as practitioner and user throughout the development and iterative refinement of the *My Partner in Equity* (MPIE) custom GPT tool. Her expertise in equity practice was essential to this analysis.

AI Acknowledgment

All ideas, interpretations, and findings are the original work of the author. The AI tool discussed in this essay (My Partner in Equity) was developed by the author using OpenAI's ChatGPT platform. Retrospective analysis utilized documented user interactions and basic platform-provided usage information, with all data anonymized.

RAKETA OUEDRAOGO-THOMAS, PhD, is a postdoctoral fellow with the Urban Education Collaborative at the University of North Carolina at Charlotte. Drawing on two decades of experience as a practitioner, her research focuses on leadership, policy, and systemic change, particularly on how communities engage with and transform educational systems. Email: raketa.thomas@charlotte.edu