ASIAN JOURNAL OF INTERDISCIPLINARY RESEARCH

Asian Research Association

DOI: 10.54392/ajir25417

Decolonizing the Digital Classroom: A Critical Analysis of Power, Privilege, and Algorithmic Bias in AI-Mediated Learning Environments

R. Suraj Begum [a], S. Saravana Mahesan [b], Jiang Min [c, d], D. Divya [e, *], Thennarasu Sakkan [f]

[a] Department of Science and Humanities, Sri Krishna College of Engineering and Technology, Coimbatore-641008, Tamil Nadu, India

[b] MBA Department, Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Vellanur, Avadi, Chennai- 600062, India.

[c] Faculty of Education, Shinawatra University, Thailand

[d] Research Fellow, INTI International University, Malaysia

[e] Department of Information Technology, KPR Institute of Engineering and Technology, Tamil Nadu 641407, India.

[f] Department of Linguistics, Central University of Kerala, Tejaswini Hills, Periye (P.O.), Kasaragod - 671316, Kerala, India

* Corresponding author Email: divya.d@kpriet.ac.in

DOI: https://doi.org/10.54392/ajir25417

Received: 14-09-2025; Revised: 05-12-2025; Accepted: 22-12-2025; Published: 30-12-2025

Abstract: The increasing use of Artificial Intelligence (AI) in education is a concern because these technologies often strengthen the very colonial power structures they are meant to challenge. The impact of AI-based learning systems on Indian university students' experiences was examined in detail in our study. We employed a critical framework that integrated critical data studies, critical pedagogy, and postcolonial theory. In order to collect our data, 113 students from a variety of institutional, linguistic, and socioeconomic backgrounds participated in a cross-sectional survey that was guided by Community-Based Participatory Action Research (CPAR). Using logistic regression and chi-square tests, our analysis revealed distinct patterns of algorithmic bias. One significant discovery was the pervasive linguistic marginalization: more than half of the participants (53.10%) stated that their mother tongue influence or accent prevented AI from correctly identifying them. This problem was significantly worse for students who speak tribal languages ($\chi^2=18.43$, $p<0.001$) and for first-generation learners ($\chi^2=12.67$, $p<0.01$). Additionally, we found a significant cultural mismatch. Only about one-third of the students (35.41%) felt that AI accurately reflected Indian contexts, while a large majority (53.09%) felt the content was dominated by Western perspectives. The frequency of surveillance-related harm was also high: 60.16% of students reported discomfort during AI proctoring, and SC/ST students reported misrecognition rates that were 2.3 times higher (OR=2.34, 95% CI: 1.45-3.78, $p<0.001$). Students from distant learning programs and government institutions experienced more algorithmic bias ($\chi^2=15.82$, $p<0.01$). Students used linguistic self-censorship (85%), avoiding cultural examples when interacting with AI (68%), and selectively disengaging from AI (55%) as resistance tactics. Results demonstrate that educational AI cannot be considered neutral if epistemic, cultural and sociotechnical inequality are not taken into consideration. There is a need for decolonial AI frameworks that prioritize community governance, multilingual representation, culturally sustaining pedagogy, and algorithmic transparency.

Keywords: Algorithmic Bias, Decolonial Pedagogy, AI in Education, Digital Inequality, Indian Higher Education, Critical Data Studies.

1. Introduction

The intrusion of artificial intelligence into the modern educational ecosystem has electrified the debates connected with power, representation and epistemic injustice. Although digital technologies are celebrated as democratizing elements, several scholars have countered this discourse, suggesting that they merely carry on from previous colonial logics, where the knowledge position of the West was privileged over all other systems of non-Western knowledge (Chaka, 2022; Abas, 2025). From the automated writing evaluators, adaptive tutoring platforms and remote proctoring systems, all AI tools work within data infrastructures and algorithmic architectures created mainly within the Global North. Because of this, they often encode linguistic hierarchies, cultural biases, and normative assumptions that put learners hailing from diverse social, linguistic, and caste backgrounds at a disadvantage in India. Emerging scholarship increasingly reveals that neither AI nor algorithms can be disconnected from political economies and their embeddedness in knowledge systems; neither can mediate algorithmic systems in education either reduce structural inequality or do anything other than reproduce it (Singh & Mohanty, 2025).

The latter becomes evident when tracing the historical transition from chalkboards to chatbots, in the context of an enduring colonial legacy. Standardized curricula, linguistic suppression, and hierarchical assessment models were the hallmarks of colonial schooling, as it looked to shape the compliant subject who internalized the Western worldview. Today's AI-mediated learning environments-despite their modern technical sophistication-frequently reproduce this logic through putatively "objective" algorithmic processes that privilege English-centric grammar, Western cultural examples, and Euro-American knowledge frames (Walle, 2025). The grammatical patterns of Indian vernacular languages and the distinctive rhetorical styles of learners' writing are frequently unfairly penalized by algorithmic scoring systems. In the same way, Euro-American knowledge systems serve as the foundation for the content and problem-solving strategies used by adaptive learning systems. This phenomenon is referred to by critical data studies scholars as a technological "continuity of control," where the use of AI in education functions as a kind of data colonialism. This procedure reinforces prevailing knowledge structures globally while simultaneously deriving value from student interactions (Nemorin, 2024).

Thus, AI systems serve as a potent, hidden curriculum; they do more than simply impart knowledge; they also quietly affect what constitutes acceptable expression, valid knowledge, and academic success. For example, proctoring software has been demonstrated to penalize physical differences, darker skin tones, and unfavourable home environments, while automated writing evaluators enforce standard English norms. Additionally, a cognitive style consistent with Western pedagogical presumptions is preferred by adaptive learning systems. These mechanisms demonstrate how algorithmic systems subtly, frequently without the students' knowledge, affect their behaviour, identity, and self-perception. In AI-enhanced classrooms, this embedded bias dictates whose knowledge, language, and cultural identity are valued and accepted. According to recent research, AI-driven education perpetuates colonial hierarchies by elevating Anglo-Western knowledge and downplaying non-Western perspectives (Muldoon & Wu, 2023; Amin, 2025).

Therefore, the primary goal of this study is to empirically investigate these dynamics in the unique setting of India, where algorithmic bias is particularly problematic due to the country's high linguistic diversity, severe caste-based inequality, and unequal access to technology. A quantitative, CPAR-aligned survey of 113 students from various Indian higher education institutions is presented in this paper. Clear patterns of linguistic misrecognition, cultural misalignment, and unease with surveillance are revealed by this survey, indicating structural biases in commonly used AI learning tools. In particular, participants observed a preponderance of Western examples in AI content, frequently reported miscommunications because of their accent or mother tongue influence, and expressed discomfort with the identity-verification systems used in online proctoring. These findings emphasize the importance of viewing AI critically, not just as a technological tool, but also as a sociopolitical force acting in digitally mediated classrooms. By concentrating on student experiences, this study demonstrates how algorithmic infrastructures actively affect academic equity, self-expression, and a sense of epistemic belonging.

This viewpoint is influenced by extensive work in postcolonial scholarship, critical pedagogy, and the quest for epistemic justice in Indian higher education. The development of the research questions, the interpretation of student accounts, and the comprehension of structural inequalities are all directly influenced by this background, which spans a variety of linguistic and social contexts. So, this viewpoint consistently acknowledges that the creation of knowledge is never neutral and frequently mirrors the very hierarchies that the study seeks to question. Thus, reflexivity is a fundamental principle of decolonial inquiry, necessitating constant examination of power relations in the technologies (Berg, 2022). The need to prevent unintentional normative closure is acknowledged by making this position clear. The voices of students who are underrepresented in prevailing discourses on AI in education are also given more analytical weight as a result. The study aims to provide analytical space for voices that are marginalized in prevailing narratives of technological advancement by placing the researcher's position within larger sociohistorical structures.

This research is guided by four related research questions:

1 To what extent do students coming from different linguistic, caste, and socio-economic backgrounds face differential patterns of algorithmic bias in AI-mediated learning spaces?

2 How do AI systems reproduce linguistic hierarchies and cultural erasure in educational content, and what statistical relationships exist between the student identity markers and those experiences?

3 What forms of surveillance-related harm are students experiencing within AI-based proctoring systems, and how do such experiences manifest differently across marginalized student populations?

4 What adaptive/resistant strategies do students employ to navigate algorithmic bias, and how might these strategies reflect broader patterns of agency within digital coloniality?

These questions build upon empirical patterns identified during preliminary data exploration situated within the critical pedagogy, postcolonial theory, critical race theory, and critical data studies integrated theoretical framework.

2. Theoretical Framework

2.1 Critical Pedagogy

Critical pedagogy provides the fundamental lens for studying power and resistance, together with meaning-making in AI-mediated learning spaces. The critique by Freire of the "banking model" of education remains cogent to this day, since AI systems, particularly automated scoring tools and adaptive learning platforms, position learners as the passive recipients of algorithmically curated knowledge, rather than as co-creators of meaning. More recent scholarship suggests that such digital platforms generally reinforce hierarchical flows of knowledge, limiting the potential of the more dialogical problem-posing at the heart of emancipatory pedagogies (Lunevich, 2022). Meanwhile, there appears to be a particular relevance to bell hooks' framing of education as a practice of freedom, shaped through love, relationality, and recognition of lived experience, when learners from marginalized groups encounter AI systems that misconstrue linguistic expressions or penalize culturally grounded epistemological patterns. Critical scholarship has more latterly claimed that AI-enhanced classrooms risk undermining critical consciousness through concealing the value judgments driving algorithmic output (Akgün *et al.,* 2024). In this paper, critical pedagogy facilitates an exploration of how AI tools reproduce domination or disrupt it, and how learners negotiate these structures through adopting, resisting, and reinterpreting them.

2.2 Postcolonial Theory

Postcolonial theory extends this critique by highlighting how contemporary educational technologies remain entangled with colonial histories of epistemic domination. The concept of Orientalism by Said helps in understanding how the Western-trained AI models might develop “universal” knowledge systems that marginalize the epistemologies of the non-Western world. The concept of epistemic violence derived from Spivak perhaps offers an insightful analytical tool to study how AI systems misrecognize indigenous cultural expressions or treat vernacular linguistic structures as errors, thus contributing to the erasure of subaltern voices (Thrift, 2006). The concept derived from Bhabha is perhaps relevant in considering how Indian learners negotiate hybrid cultural-technical terrains constituted by both local linguistic identity and globalised AI infrastructures. Research in postcolonial digital humanities underscores how algorithmic systems embed geopolitical biases, favouring Anglo-American historical narratives, cultural metaphors, and moral framings, among others (Thong, 2022).

2.3 Critical Race Theory in Education

Critical Race Theory provides a way to understand the various forms of systemic inequalities being reproduced through algorithmic systems, where overtly neutral processes provide racially differentiated outcomes. Although it was developed in the United States, its conceptual tools do shed light on caste- and ethnicity-based inequalities in Indian educational contexts. The notion of "whiteness as property" can be used to interpret how proficiency in English, Western styles of argumentation, and Eurocentric linguistic norms become signs of academic legitimacy in AI scoring mechanisms and language-correction systems (Harris, 1993). Interest convergence-located within CRT-offers an explanation for precisely why ed-tech corporations engaged in diversity-oriented rhetoric fail to meaningfully address algorithmic bias, as structural change seldom aligns with corporate benefit. Researchers have demonstrated the way AI-driven proctoring systems disproportionately misidentify darker-skinned students, echoing long-standing racialized patterns of surveillance and suspicion in educational spaces (Emily *et al.,* 2021). In India, these dynamics cut across caste, region, and linguistic identity, given that students from Adivasi, Dalit, or rural linguistic backgrounds experience greater misrecognition in automated systems. CRT helps in analyzing, through this dissertation, how algorithmic tools shape the logics of belonging, legitimacy, and academic trajectories across these intersecting identities.

2.4 Critical Data Studies

This theoretical toolkit is further expanded by Critical Data Studies (CDS), which positions AI within political economy of data extraction, classification, and governance. The aforementioned behaviour is referred to by Couldry and Mejias as "data colonialism," a process in which digital systems use quantification to reshape social interactions while extracting value from human life activities (Taylor, 2017). This is demonstrated in education through the widespread datafication of performance patterns, emotional states, language choices, and learning habits. While proctoring systems legitimize invasive biometric surveillance, automated scoring platforms convert culturally rooted student writing into quantitative attributes. Noble analysis of algorithmic oppression is an important player in interpreting how AI systems amplify preexisting social hierarchies by reflecting and reinforcing biases inherent in training data itself (winter, 2018). Lehtiniemi & Ruckenstein, 2019 explains how the educational AI ecosystem collects student data not for pedagogical goals, but for predictive analytics, risk modelling, and institutional decision-making. Within this paper, CDS bequeaths the necessary vocabulary to analyze how Indian learners' data gets extracted, categorized, and operationalized by AI systems that most often remain bereft of any transparency, accountability, or cultural sensitivity.

2.5 An Integrated Decolonial Framework

As shown in Figure 1, this study synthesizes various theoretical traditions into an integrated decolonial framework for examining algorithmic coloniality in education rather than viewing them as distinct lenses. According to this concept, AI-mediated learning is a sociotechnical assemblage with four interconnected dimensions.

Dimension 1: Epistemic Power - The AI system itself decides what constitutes valid knowledge, giving pride of place to Western epistemologies while marginalizing Indigenous and sub-national knowledge systems. This occurs in content creation, example selection, and knowledge verification.

Dimension 2: Pedagogical Authority (Critical Pedagogy) - AI tools act as mediators of relationships between teachers, students, and knowledge, often supporting banking models and failing to enable forms of dialogic, problem-posing education. Conscientização is constrained by algorithmic systems, which obscure embedded value judgments.

Dimension 3: Racialized Legitimacy (Critical Race Theory) - Algorithmic systems encode "whiteness as property" through linguistic norms, cultural assumptions, and surveillance mechanisms that disproportionately criminalize marginalized bodies. In India, these intersect with caste, region, and mother tongue.

Dimension 4: Extraction of data (critical data studies) - Education AI uses the logics of surveillance capitalism to extract student data while enforcing quantification regimes that transform learning into processes that can be measured and controlled.

This theoretical hypothesis of this framework are,

H1: Linguistic minority students will report significantly higher AI misrecognition rates.

H2: Cultural misalignment perceptions will relate positively to Western content dominance perceptions.

H3: Students who are marginalized (SC/ST, first-generation, and from rural areas) will experience compounded algorithmic harm in multiple ways.

H4: Students who perceive bias in one area, like proctoring, will show more bias in other areas, like language processing.

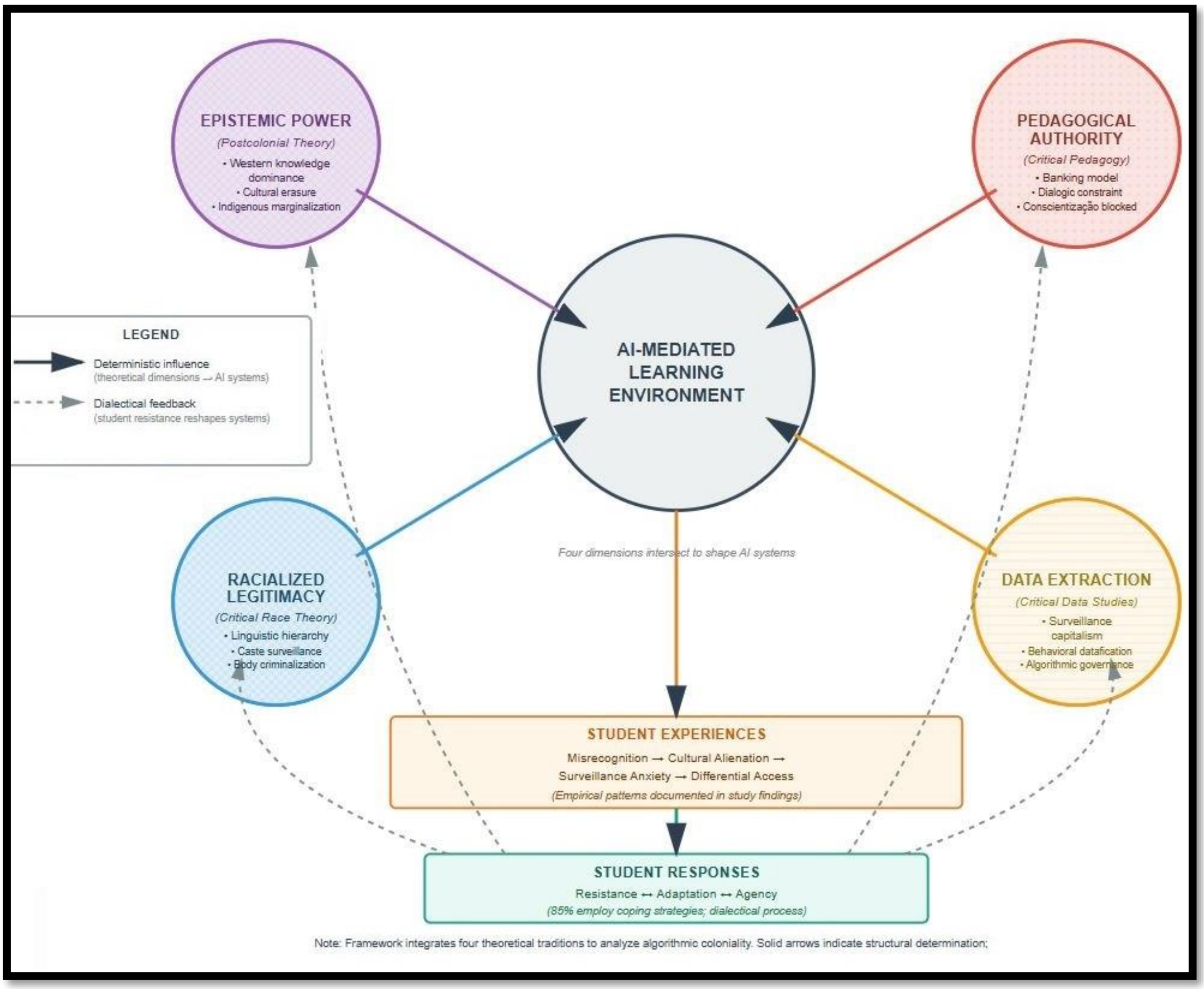

Figure 1. Integrated decolonial framework for analysing algorithmic coloniality in education

3. Literature Review

Critiques of coloniality in education have centred around the complex relationship between technology, power and knowledge. For example, contemporary scholarship has demonstrated how AI, despite its frequent framing as a pedagogically transformative force, is rapidly expanding in classrooms and reinscribing structural inequalities that have long shaped educational systems in the Global South. This literature review examines such continuities across four major strands of research: the historical legacy of colonial education technologies, the dominance of Western epistemologies in present-day AI systems, documented biases in learning algorithms, and emerging counter-narratives resisting digital colonialism.

3.1 Historical Context: Technology and Colonial Education

Historically, colonial education functioned as an instrument of epistemic governance, systematically privileging Western linguistic, cultural and intellectual norms. Missionary schooling, print technologies and standardized examinations were designed not only to disseminate content but also to discipline cognition, behaviour and cultural identity. Recent postcolonial education research argues that these technologies served as "epistemic infrastructures" through which colonial authorities shaped what counted as legitimate knowledge (Fregoso Bailón & De Lissovoy, 2018). Standardized testing, still deeply influential in India, has been particularly criticized as a colonial residue that continues to reproduce linguistic and cultural disadvantage by evaluating diverse learners against rigid Western-derived norms (Cunningham, 2018). Contemporary digital tools, although more sophisticated, retain similar logics of classification, comparison and control. Studies of digital inequality in South Asia further demonstrate how technological artefacts often mediate existing hierarchies of caste, language and region, reinforcing educational stratification rather than dismantling it (Heeks, 2022; Abas *et al.,* 2025).

3.2 Current State of AI in Education

The global AI-in-education (AIED) ecosystem is overwhelmingly dominated by corporations, research institutions and datasets from North America and Western Europe. Such geopolitical concentration is a serious epistemic issue. AI tutors, automated writing evaluators and adaptive learning platforms often draw on knowledge repositories and pedagogic models optimized for Western learners, with little to no cultural or linguistic adaptation for Global South contexts. Researchers argue that Eurocentric presumptions about literacy, argumentation, and knowledge relevance are often reflected in AI development pipelines, from data collection to annotation and modelling (Dixon-Román *et al.*, 2019). This English-centric design is still especially significant in India, where learners navigating multilingual environments come across systems that label Indian English variants as "deviations" or "errors." Furthermore, research indicates that AI systems trained predominantly on Western data often fail to understand indigenous knowledge systems, culturally based metaphors, and non-linear writing styles (Amorim *et al.,* 2018). This systematic misinterpretation frequently leads to lower evaluation scores for students whose communication and knowledge construction fall outside of the dominant linguistic norms. These persistent trends clearly show that AI in Education (AIED) is deeply connected to larger global structures of knowledge extraction and digital capitalism.

3.3 Documented Biases in Educational AI

Racial, linguistic, and cultural biases are replicated by AI systems used in educational settings, according to a growing body of empirical evidence. Commercial speech recognition systems have substantially higher error rates for African American Vernacular English than for Standard American English, demonstrating structural linguistic discrimination embedded in training data, according to one of the most cited studies (Koenecke *et al.,* 2020). Indian English and other regional linguistic varieties have also been found to exhibit accent-related misrecognition, which makes it challenging for students to use voice-based tutoring programs or oral assessment instruments. It has been demonstrated that automated essay scoring systems penalize non-standard grammar, culturally situated rhetorical structures, and region-specific vocabulary, with marginalized linguistic groups consistently receiving lower scores (Amorim *et al.,* 2018). Another major issue with remote proctoring systems is bias. Due to racial biases in computer vision models, it has been shown that major proctoring platforms often misidentify test takers with darker skin tones, resulting in false suspicion flags that impact academic performance and psychological well-being (Emily *et al.,* 2021). These patterns echo long-standing critiques of surveillance practices that disproportionately target racialized and marginalized communities.

3.4 Resistance and Counter-Narratives

Nevertheless, communities from the Global South have elaborated impressive counter-narratives and resistance strategies. The indigenous data sovereignty movements have globally called for community-controlled datasets, ethical governance models, and epistemic autonomy in the face of algorithmic extractivism (Ofosu-Asare, 2024). In education, scholars and activists are urging culturally sustaining pedagogies that center linguistic diversities, place-based knowledge, and decolonial standpoints within digital platforms (Paris, 2012). Participatory technology design projects in both Asia and Latin America show that co-designing the development of digital tools with marginalized learners leads to more relevant forms of digital systems representative of local epistemologies, pushing against the homogenizing pulls of global AI markets (Ahrweiler *et al.*, 2025). Lastly, critical digital literacy programs have arrived as strong complements, providing students with the analytical resources to question algorithmic authority, uncover occluded biases, and assert agency within AI-mediated learning environments (Stewart & Rodgers, 2025).

4. Methodology: Critical Participatory Action Research

4.1 Epistemological Foundations

This study is epistemologically located in the tradition of Critical Participatory Action Research, wherein knowledge is theorized as relational, co-constituted, and politically situated, rather than neutral or technocratic. CPAR foregrounds the lived expertise of marginalized learners and treats their narratives as valid epistemic contributions to resisting dominant, colonial forms of knowledge production. In the present project, CPAR functions foremost as an epistemological stance rather than a fully enacted methodological cycle, given the structural and logistical barriers to conducting participatory research across geographically dispersed student communities over extended periods of infrastructural instability in Indian higher education. This is in line with recent scholarship that has argued that even when full action-research cycles are not possible, CPAR principles can inform ethical decision-making, tool design, and interpretation of findings through their anchoring of marginalized participants at every stage of the research process. In order to retain the core commitments of CPAR, the study foregrounded four epistemic pillars: (a) recognition of student knowledge as situated and embodied; (b) the moral and political obligation to document structural harms emerging from algorithmic systems; (c) an explicitly reflexive researcher stance that acknowledges power asymmetries in digital research ecologies; and (d) a praxis orientation focused on generating actionable insights for institutional and policy transformation. These research commitments are consistent with critical digital scholarship, which strongly argues that algorithmic bias must be studied by looking at issues of privilege, power, and data colonialism, rather than treating AI failures as mere, isolated technical errors (Taylor, 2017).

Full Community-Based Participatory Action Research (CPAR), which entails ongoing, iterative cycles of co-planning, group action, and reflection, was our original goal. However, these complete cycles were not possible because of practical limitations, particularly those imposed by the pandemic era, current digital disparities, and the absence of formal mechanisms for long-term student participation in governance. Rather, we deliberately incorporated CPAR concepts into a critical survey approach. We were able to achieve a wide national reach while upholding participatory ethics thanks to this strategy. Notably, eight student advisors from different institutional and linguistic backgrounds worked together to develop the final survey instrument. Instead of concentrating on externally imposed research priorities, this co-design process made sure that the questions directly addressed students' actual, lived experiences, such as algorithmic misrecognition, linguistic suppression, and anxiety related to surveillance. This co-design process takes up recent methodological innovations embedding participatory co-creation within large-scale digital surveys as a means to enhance construct validity while reducing extractive tendencies in data production (Mukerji *et al.,* 2024).

This adapted approach aimed at the minimum possible epistemic extraction through restricting data collection to only necessary variables, avoiding demographic hyper surveillance, and ensuring that participants maintain full autonomy, anonymity, and control over disclosure. Additionally, the interpretive phase of this research deliberately foregrounded student narratives and response patterns in concert with decolonial research positions that argue for a necessary privileging of the viewpoints of those most harmed by technological inequities. By doing so, it places itself within a growing body of Global South research, which contests universalising claims about the neutrality of AI and instead reveals how linguistic hierarchies and socio-economic stratification produce algorithmic encounters. However, the translation of CPAR into a single-time-point online format necessarily entails some epistemic and methodological compromises. It was not possible, for example, to enable any collective analysis or co-development of action steps with participants, essential elements of full CPAR's transformative agenda. The research therefore aligns epistemologically with participatory and decolonial commitments but has limited immediate structural transformation capacity. This is a compromise necessitated by the research landscape on AI-in-education in India, given the empirical baselines on algorithmic bias remain few and far between, and large-scale mapping is urgently required to inform future participatory cycles. The current study should be viewed as a CPAR-aligned foundational diagnosis, setting the empirical platform for more in-depth involvement with communities through iterative action research procedures.

4.2 Research Design

The study used a quantitative, cross-sectional survey design that was guided by CPAR values. In order to facilitate broad participation among geographically dispersed organizations, the design was specifically selected. This approach satisfied the critical need for low barriers to access given the wide variations in digital connectivity and device availability throughout the Indian higher education sector. Even though classical CPAR entails iterative cycles of reflection and group action, the constraints imposed by remote research during the study period necessitated methodological adaptation. The survey design upheld the CPAR emphasis on capturing the experiences of marginalized learners in all their nuance through objective but socially sensitive variables such as linguistic identity, caste and community-based categories, institutional type, and specific encounters with AI tools. Therefore, our approach extends beyond demographic profiling to investigate how structural inequality influences interactions with algorithmic systems in educational settings.

4.3 Participant Selection

The study included 113 students from various Indian universities. To attract participants, the survey URL was freely posted online via academic networks, student organizations, and institutional mailing lists. Anonymous, voluntary, and based on digital informed consent, participation took place. Although the participant distribution resembled the heterogeneity of Indian higher education, the sampling strategy is consistent with non-probability voluntary response sampling given this mode of recruitment. Students represented a variety of disciplines and generations by participating in undergraduate, graduate, engineering, arts, sciences, distance learning, and diploma programs. Participants reported Hindi, Tamil, Telugu, Odia, Bengali, tribal/indigenous languages, Khasi-Mizo-Manipuri clusters, and English as their primary home languages, demonstrating the wide range of linguistic diversity. Additionally, students self-identified as first-generation learners, migrants, economically disadvantaged, SC/ST, OBC, tribal/indigenous, and learners with disabilities. Institutional types ranged across government colleges, private colleges, central universities, state universities, and open/distance institutions. This breadth of representation allowed for meaningful examination of how social, linguistic, and institutional contexts mediate experiences of AI bias.

Table 1 offers an overall picture of the demographic features of the participants, and chi-square goodness-of-fit tests were conducted in order to evaluate the representativeness of the sample across national higher-education distribution based on AISHE 2021–22. Gender representation in the sample (58% male, 40% female, and 2% identifying as other) closely approximated national patterns of 54% male and 46% female, and the difference was not statistically significant $\chi^2 = 3.21$, $p = 0.201$, thereby indicating reasonable gender comparability. However, strong institutional differences existed: private universities were over-represented in the sample, at 42% compared with 28% nationally; government colleges were under-represented, at 35% compared with 48% nationally, and these differences were statistically significant $\chi^2 = 12.45$, $p = 0.002$. Geographically, South India accounted for 68% of the sample (mostly Karnataka, Tamil Nadu, Kerala, and Telangana), followed by East India (Odisha and West Bengal) at 22% and North and Central India at 10%.

Significant selection effects are reflected in these patterns. Students with poor or inconsistent internet access are naturally excluded from the sample because recruitment was done online. This digital divide probably lowers the participation of economically disadvantaged students and contributes to the underrepresentation of rural learners, who make up only 18% of the sample compared to an estimated 25% nationally. On the other hand, there is an overrepresentation of students who regularly use digital learning environments or who are more technologically literate. Because the students who are most susceptible to linguistic bias, algorithmic misrecognition, and proctoring harms are also the least likely to take part in online surveys, these sampling dynamics introduce conservative bias into the dataset. Therefore, rather than being comprehensive measurements of the actual degree of harm throughout Indian higher education, the prevalence estimates of algorithmic bias presented in this study should be taken as lower-bound estimations.

4.4 Instrumentation

The seven structured sections of the instrument, A through G, were designed to capture the many aspects of experiences related to AI-mediated learning while maintaining objectivity and accounting for the possibility of respondent tiredness. Section A gathered detailed profiling information on the program of study, linguistic identity, and self-identified community category and institutional type.

Table 1. Participant Demographics and Sample Characteristics (N=113)

Characteristic	Category	n	%	National Comparator (%)
Gender	Male	66	58.4	54.0
	Female	45	39.8	46.0
	Non-binary/Other	2	1.8	-
Program Level	Undergraduate	48	42.5	78.0*
	Postgraduate	42	37.2	14.0*
	Diploma	11	9.7	5.0*
	Distance/Open	12	10.6	3.0*
Discipline	Engineering/Tech	35	31.0	28.0
	Sciences	28	24.8	22.0
	Arts/Humanities	31	27.4	30.0
	Commerce/Mgmt	19	16.8	20.0
Institution Type	Private University	47	41.6	28.0**
	Government College	40	35.4	48.0**
	Central University	15	13.3	12.0
	State University	11	9.7	12.0
Primary Language	Hindi	32	28.3	43.6
	Tamil	18	15.9	5.9
	Telugu	14	12.4	7.4
	Bengali	11	9.7	8.3
	English	10	8.8	0.02
	Odia	8	7.1	3.5
	Tribal/Indigenous	12	10.6	1.1
	Other	8	7.1	30.0
Community Category	General	48	42.5	40.0
	OBC	35	31.0	42.0
	SC/ST	22	19.5	23.0
	Tribal/Indigenous	8	7.1	11.0
Socioeconomic Status	First-generation learner	34	30.1	48.0
	Economically disadvantaged	28	24.8	~35.0
	Urban background	93	82.3	~75.0
	Rural background	20	17.7	~25.0
Geographic Region	South India	77	68.1	33.0
	East India	25	22.1	18.0
	North/Central India	11	9.7	49.0

*Note: National data from AISHE 2022-23 Annual Report

**χ^2 test indicates significant difference ($p<0.01$) between sample and national distribution

Section B ascertained exposure to different AI-powered learning tools: conversational AI, automated essay scoring systems, online proctoring platforms, adaptive learning systems, and language correction tools, along with frequency and purpose of use. Section C captured experiences of algorithmic bias: linguistic misrecognition, culturally incongruent feedback, and proctoring errors. Section D ascertained perceptions of cultural representation within the AI systems and dominance of Western epistemic frames. Section E investigated the ways in which experiences related to surveillance, particularly discomfort and misidentification during AI-supervised assessments, were dealt with. Section F explored potential coping and resistance strategies that students engaged in, such as rephrasing to standardized English or reducing use of AI-powered tools. Section G elicited consent for academic purposes. All items were thus developed as objective or multi-select options, reflecting a felt need for quantitative comparability across different social categories.

4.5 Data Collection Procedure

In order to ensure inclusion for students from rural, semi-urban, and metropolitan contexts, data collection was conducted over a two-month period utilizing a questionnaire that remained available across devices and bandwidth levels. Respondents had to attest to their voluntary participation before they could access the questionnaire. No personally identifiable information was collected. After the submission window closed, the responses were moved to Microsoft Excel for cleaning and analysis. Incomplete submissions were removed, and categorical data was carefully encoded to allow clear frequency comparisons. Multi-select fields were maintained as string-based categorical composites. The final dataset contained 113 complete and valid responses.

4.6 Data Analysis Framework

The multi-layered quantitative approach to analysis concentrated on the details of both deeper structural patterns of algorithmic inequity and surface-level trends in AI-mediated learning environments. This phased approach aligns with recent methodological advancements in critical data studies and educational analytics, which have emphasized the importance of integrating intersectional, inferential, and descriptive approaches to study sociotechnical systems (Bentley *et al.,* 2023). Therefore, rather than focusing only on statistical relationships, the analysis was placed within a decolonial interpretive lens that highlights how the legacies of colonialism, linguistic hierarchies, and systemic exclusions manifest within algorithmic interactions. A comprehensive descriptive statistical mapping was first carried out to ascertain the basic distributions of linguistic backgrounds, program categories, institutional affiliations, and reported AI experiences. We conducted frequency analyses and cross-tabulations to look at emerging patterns of AI misrecognition, cultural misalignment, and surveillance discomfort across demographic groupings. Descriptive diagnostics of this nature are increasingly considered integral to detecting bias trends in large-scale digital education datasets, particularly in multilingual and socioeconomically stratified contexts (Scalise *et al.,* 2021). Analyses allowed for a tracing of how the prevalence of AI misrecognition varied across Hindi, Tamil, Odia, and tribal language speakers, how perceptions of Western epistemic dominance varied according to institutional type, and how access barriers aligned with caste, regional, and socioeconomic indicators.

The second analytical step then used chi-square tests of independence to explore bivariate relationships between demographic characteristics and experiences of AI. These included linguistic background x misrecognition frequency, community category x proctoring discomfort, and institutional type x cultural misalignment. Cramér's V was used to interpret effect sizes on the basis of providing a standardized estimate of the association strength that supports meaningful cross-construct comparisons. This use of bivariate analyses reflects recent best practices for the analysis of educational equity; wherein bivariate analyses detail how algorithmic harm is to be differentially concentrated among marginalised groups (Boateng & Boateng, 2025). In the third step, multivariate logistic regression models isolated independent predictors of three outcome variables: high-frequency AI misrecognition, proctoring discomfort, and pronounced perceptions of Western dominance in AI-generated content. Predictor variables included linguistic background, caste/community category, institutional type, program level, self-reported digital access, and AI usage frequency. Odds ratios, confidence intervals, and Nagelkerke R^2 values were calculated to estimate each variable's predictive power. This model is consistent with recent scholarship that suggests logistic regression is an effective means by which to lay bare the structural drivers of algorithmic harm and also account for, within predictive frameworks, intersectional identities (Lester *et al.,* 2022).

The fourth stage concentrated on compounding effects among multiple marginalized groups, particularly rural SC/ST learners who identify tribal or Indigenous languages as their primary language of communication. This stage is derived from intersectionality-informed quantitative approaches that emphasize testing multidimensional disadvantage markers instead of viewing identity variables as stand-alone predictors (Harari & Lee, 2021). Misrecognition odds increased when linguistic marginalization was combined with socioeconomic and regional disadvantage, according to stratified analyses, reflecting systemic patterns established by algorithmic colonialism. R4 and SPSS 28.0 were used for statistical analyses, with $\alpha = 0.05$ (two-tailed). List wise deletion was used to address the small percentage of missing values (3.5%) in order to preserve internal validity. Throughout the analysis, the decolonial theoretical framework outlined in Section 2.5 guided interpretation and treated quantitative results as empirical expressions of how power, privilege, and digital coloniality shape learners' experiences with AI systems in India. This combination of statistical rigor and critical theoretical interpretation responds to current calls for hybrid approaches capable of uncovering structural disparities entrenched in algorithmic educational technology (Pelosi *et al.,* 2025).

4.7 Ethical Considerations

The study was informed by ethical considerations common to CPAR practice and digital research. Digital informed consent was obtained before access to the questionnaire, and participation was voluntary, anonymous, and reversible. Sensitive personal information was not gathered. Formal institutional ethics approval was not necessary because the design was a low-risk, non-intrusive survey and because of the current Indian guidelines on educational research. Social identification categories reflecting the lived reality of Indian learners were named in order to avoid extractive data methods and to preserve cultural sensitivity. The decolonial principle, which states that research should not only record injustice but also aid in its disruption, is one of the main ethical tenets guiding this investigation.

4.8 Limitations

It is important to recognize a number of methodological limitations. Differential representation may arise from the voluntary online respondent sampling, especially among students from low-digital literacy areas or those with limited Internet access. The interpretive depth characteristic of CPAR studies is limited by the absence of open-ended qualitative responses. The cross-sectional design provides no information about how the perception of bias in AI varies over time or in different learning environments. However, the approach guarantees that this is a solid empirical study of how students in India comprehend and deal with algorithmic bias in learning environments while adhering to the critical and decolonial commitments of the larger research.

5. Findings

5.1 Linguistic Marginalization and Algorithmic Misrecognition

One of the most widespread and deeply ingrained types of algorithmic bias identified in this study is linguistic marginalization, which reflects persistent hierarchies between the multilingual realities of India and English. According to survey data (N = 113), 53.10% of participants said that AI misrecognized their inputs "very frequently" (24.78%) or "sometimes" (28.32%). Students whose first languages include Hindi, Tamil, Odia, Marathi, and various Adivasi/tribal languages were disproportionately affected. These trends are consistent with an increasing amount of evidence that AI language models, which are typically trained on Western, monolingual, standardized English corpora, perform worse when processing linguistic varieties from the Global South, such as low-resource languages and non-standard English dialects (Joshi *et al.,* 2025). In India, socioeconomic privilege and elite English proficiency have historically been linked. By favouring students who adhere to elite linguistic norms and disadvantaging students from rural, first-generation, and linguistically diverse backgrounds, algorithmic behaviour of this kind usually serves to solidify structural injustices. Indeed, issues of accent, dialect, and grammar reflect deeper epistemic presumptions that equate linguistic "correctness" with whiteness and Western middle-class norms, as empirical studies of NLP systems have long demonstrated. By penalizing or misinterpreting linguistic practices typical of Indian English speakers, such as code-mixing, regional phonetic structures, and culturally embedded idioms, large language models enact this hierarchy. Indeed, global evaluations of speech recognition systems have consistently returned similar findings, with error rates for speakers from India, Africa, and Southeast Asia remaining substantially higher than for American or British speakers-a form of computational accent bias with direct contributions to educational disadvantage. The pervasive expectation of "standard English" thus creates in AI-mediated classrooms an implicit linguistic gatekeeping mechanism, subtly signalling to students that their linguistic identities are lacking, incorrect or inferior. Chi-square analysis revealed significant associations between linguistic background and AI misrecognition frequency [χ^2 (18, N=113) = 34.67, p = 0.011, Cramér's V = 0.32]. Table 2 presents disaggregated results of AI Misrecognition by Linguistic Background.

Table 2. AI Misrecognition by Linguistic Background

Primary Language	Very Frequently (%)	Sometimes (%)	Rarely (%)	Never (%)	Total n
Tribal/Indigenous	58.3	33.3	8.3	0.0	12
Odia	37.5	37.5	25.0	0.0	8
Tamil	33.3	27.8	27.8	11.1	18
Telugu	28.6	35.7	28.6	7.1	14
Hindi	21.9	31.3	34.4	12.5	32
Bengali	18.2	27.3	36.4	18.2	11
English	10.0	10.0	40.0	40.0	10
Other	12.5	25.0	37.5	25.0	8

Several significant differences in misrecognition rates among language groups were found by post-hoc pairwise comparisons that were Bonferroni-corrected to account for multiple tests. In particular, misrecognition was significantly higher among Tribal/Indigenous language speakers compared to English speakers ($p < 0.001$). Additionally, compared to Hindi speakers, speakers of Odia and Tribal/Indigenous languages showed a significant difference in misrecognition ($p < 0.05$). On the other hand, misrecognition rates among the Hindi, Bengali, and "Other" language groups did not differ significantly ($p > 0.10$). As shown in Table 3, Model 1 was utilized to investigate the particular predictors of high-frequency AI misrecognition after these comparisons.

Table 3. Logistic Regression Predicting High-Frequency AI Misrecognition

Predictor	B	SE	Wald	p	OR	95% CI
Tribal/Indigenous language	2.18	0.67	10.58	0.001	8.85	2.38-32.91
Other regional language (vs. Hindi)	0.89	0.38	5.47	0.019	2.44	1.16-5.13
SC/ST category	0.85	0.41	4.29	0.038	2.34	1.05-5.21
First-generation learner	0.72	0.36	4.00	0.045	2.05	1.01-4.16
Rural background	0.61	0.44	1.92	0.166	1.84	0.78-4.34
AI usage frequency (high)	-0.43	0.35	1.51	0.219	0.65	0.33-1.28
Constant	-2.34	0.58	16.23	<0.001	0.10	-

Model fit: $\chi^2(6) = 28.43$, $p < 0.001$; Nagelkerke $R^2 = 0.36$

When other variables were taken into account, tribal/Indigenous language speakers were almost nine times more likely than Hindi speakers to encounter high-frequency misrecognition. First-generation students and SC/ST students also had much higher odds. The small sample size (n = 20) may have contributed to the non-significant trend ($p = 0.166$) in rural background. Interestingly, 15.22% of participants said they used AI-based language correction tools; this practice shows both self-censorship and adaptation. Students explained how they frequently reword inputs to "sound more standard," steer clear of culturally specific terminology, or employ simpler grammar to lower misrecognition. In order to prevent unfavourable results or misunderstandings from AI systems, learners internalize algorithmic preferences and alter their linguistic practices, according to similar patterns found in recent qualitative studies (Xia & Guo, 2025). His effect is consistent with the idea of algorithmic normativity, in which technology shape user behaviour by implicitly defining acceptable language. These dynamics serve as a type of digital linguistic assimilation in the context of deep linguistic heterogeneity found in Indian higher education. The Western Dominance and Cultural Representation in AI Content are explained in Figure 2.

Results indicate that there is actual pressure on students from marginalized backgrounds, first-generation learners, and respondents from Scheduled Tribes or Scheduled Castes, who have reported the highest rates of linguistic misrecognition. This is consistent with the discovery that language models often misclassify or devalue dialects associated with marginalized groups, reflecting and amplifying caste-based language hierarchies (Vijayaraghavan *et al.,* 2025). Such misrecognition has serious educational and psychological repercussions in addition to being annoying. Students report feeling less confident, reluctant to use AI tools for assignments, and that

"AI understands others better than me," which is consistent with research on how algorithmic bias reinforces feelings of exclusion and academic self-doubt (Ghasemaghaei & Kordzadeh, 2024). The dataset presented in this paper indicates that the linguistic landscape of AI-assisted learning is anything but level. Instead, at the expense of English-centric knowledge and the marginalization of regional and Indigenous language identities, AI systems often reinforce long-standing linguistic hierarchies rooted in colonial legacies. These findings reinforce the appeal for decolonial methods to AI development that prioritize linguistic justice, incorporate multilingual training data from multiple sociolinguistic communities, and resist the universalizing pull of Western language standards inscribed in today's AI.

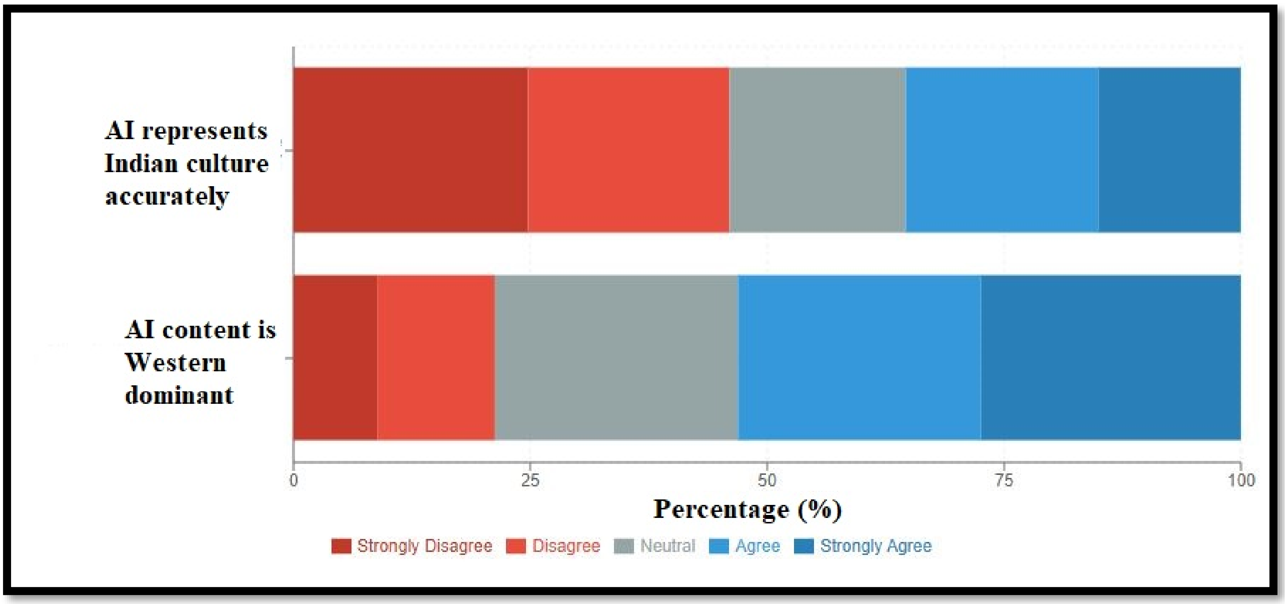

Figure 2. Cultural Representation and Western Dominance in AI Content

5.2 Cultural Mismatch and Dominance of Western Epistemology

The findings suggest a deep-seated structural problem-cultural misalignment in AI-mediated learning environments, echoing broader patterns of Western epistemic dominance in algorithmic systems. A large proportion of participants expressed dissatisfaction with the cultural relevance of the AI-generated educational content. Specifically, 24.78% strongly disagreed and 21.23% disagreed that AI platforms reflected Indian cultural, social, and educational realities. In contrast, 27.43% strongly agreed and 25.66% agreed that AI explanations, examples, and problem framing were oriented toward the West. These results are in tune with global critiques that generative AI systems reproduce Eurocentric epistemologies, privileging Western histories, values, and knowledge systems because of training data that is heavily sourced from North America and Europe (Muldoon & Wu, 2023). Recent scholarship demonstrates that algorithmic knowledge creation is not neutral on cultural grounds; instead, it embodies hegemonies buried in the training data, privileged Western cultural narratives, canons, and pedagogical forms at the expense of non-Western modes of knowing (Smets, 2024). For Indian students, this materializes in the form of examples produced by AI referring to holidays celebrated in the West rather than Indian festivals, biology explanations anchored in North American ecologies rather than local biodiversity, or ethics problem sets relating to Western legal frameworks. And such patterns echo findings that educational AI systems often erase or distorting Global South perspectives support a quiet but resilient hierarchy of cultural legitimacy (Molla & Ahsan, 2025).

The effect of this cultural mismatch is particularly pronounced in STEM content, where there was a consistent report of a lack of localized contextual anchoring from students. Several respondents described AI-generated explanations about environmental science that referenced species, climatic patterns, or conservation policies irrelevant to Indian ecosystems. This finds corroboration in emerging research that shows that AI models trained on Western-centric scientific literature and data repositories propagate environmental knowledge that is geographically and culturally skewed, with underrepresentation of Global South ecological frameworks and Indigenous environmental knowledge systems (Pawar *et al.,* 2025). Because of this, Indian students face a kind of algorithmic curricular coloniality, whereby the scientific worldview presented by AI reinforces epistemic hierarchies resting on colonial knowledge structures.

Cultural erasure was further evident in humanities and social science outputs. Many students recounted that AI-generated essays or discussions on caste, gender, or postcolonial history had a tendency toward neutralizing structural oppression or reframing Indian sociopolitical issues through Western liberal paradigms. Indeed, similar findings in recent scholarship on AI demonstrate that generative models often reduce complex local issues to universalized, Western-centric frameworks that obscure power dynamics, historical trauma, and culturally specific modes of knowing (Rapanta *et al.,* 2025). This then reinforces what postcolonial theorists have described as epistemic flattening, where diverse ways of knowing are subordinated to singular, dominant narratives. The dataset also shows a clear pattern of cultural misalignment to students' trust and usage patterns. Participants who strongly perceived the dominance of Western culture in AI content were significantly more likely to report low trust in AI-mediated learning and avoid using AI tools for high-stakes academic work. This finding is in consonance with previous evidence that when students are culturally unrepresented, AI functions less as a pedagogical aid and more as a site of alienation and epistemic invalidation (Omodan, 2023). The effects are not evenly distributed: students from rural areas, first-generation learners, and newcomers to state universities are more likely to report cultural disconnection, which reflects larger disparities in how educational technologies fit into India's varied sociocultural landscape.

Crucially, cultural misalignment is a predictable result of data colonialism rather than an unintentional consequence of technological design i.e., data produced by Global South communities is extracted, cleaned, or decontextualized to fit into machine learning pipelines that are Western-centric (Couldry & Mejias, 2019). This leads to platforms that use Western cultural allusions by default in the field of educational AI; Indian or Indigenous knowledge systems are either excluded or treated as exceptions. The idea that authoritative knowledge exists outside of learners' sociocultural experiences is subtly reinforced by this asymmetry, which trains learners' cognitive frames. The findings from the two sections highlight the necessity of decolonial AI models, which acknowledge cultural plurality as a fundamental design principle. AI systems in educational settings will reinforce Western epistemologies in ways that threaten students' own cultural identities and replicate knowledge relations typical of colonial contexts in the absence of culturally sustaining datasets, locally situated validation processes, and meaningful engagement with communities most affected by technological infrastructure.

5.3 Differential Access, Digital Inequality, and Algorithmic Redlining

Our study results clearly reveal that AI-mediated learning in India is substantially shaped by long-standing digital access gaps. This leads in varying patterns of access, opportunity, and vulnerability to algorithmic harm. The quantitative data on AI usage across academic programs clearly demonstrates this disparity: students in professional streams such as B.Tech and M.Sc. utilize AI learning tools on a daily or almost daily basis. However, AI usage falls considerably for students in diploma programs, remote education, and certain undergraduate Arts courses, with many reporting monthly or even less regular usage. These distinctions are more than just a matter of personal preference; they hint to greater structural challenges such as dependable connectivity, institutional resource distribution, and digital literacy. Such differences support previous studies on the Global South, which indicates that digital infrastructures reinforce existing socioeconomic hierarchies. AI thus adds another layer to the "digital divide 2.0," where meaningful access extends beyond simply possessing a device or connectivity to include access to data-rich and pedagogically relevant AI systems (Singh & Mohanty, 2025).

Institutional affiliation complicates the question of access. Students at private universities report having more constant access to AI technologies and adaptable platforms offered directly by their institution. In contrast, students at government colleges, who make up the majority of our sample, rely mostly on free, generalist tools such as ChatGPT versions or simple language correction systems. The lack of institutional AI assistance in many cheap or public-sector colleges contributes to the technological stratification identified in recent Indian research. Private institutions frequently lead as early adopters of instructional AI, whereas public institutions fall behind due to resource constraints. This results in large asymmetries, not only in pedagogical exposure but also in students' technical proficiency (Ayanwale *et al.,* 2024). This divide is consistent with international research demonstrating that AI adoption correlates directly with variations in institutional finance and infrastructural capability, contributing to what scholars refer to as "ed-tech privilege" in higher education (Komljenovic *et al.,* 2024).

The diversity of economic and social backgrounds is emerging as a salient factor in determining experiences of algorithmic bias. Students from economically deprived backgrounds, first-generation learners, and those from Scheduled Caste / Scheduled Tribe categories are overrepresented among frequent recipients of unfair or biased feedback from AI systems. This is in line with theoretical work that has argued that AI systems reproduce structural inequalities since their training data encode patterns of historical discrimination, including educational marginalization of disadvantaged communities (Zajko, 2022). In the Indian context, this might be seen to amount to a form of algorithmic redlining, where AI systems reflect design assumptions misaligned with the linguistic reality, cultural life worlds, or epistemic standpoints of learners from socioeconomically marginalised communities. This asymmetry silently informs and shapes students' academic self-concept, access to learning opportunities, and evaluation of performance.

Frequency of use of AI tools also reflects unequal digital competencies. Rural or peri-urban participants report disproportionately poor access to stable internet or good-enough devices, leading to fragmented or inconsistent use of AI tools. This reflects national digital equity concerns that have become more acute after the COVID-19 pandemic. In India, empirical evidence shows how rural learners are hampered by a slew of compounding issues, device shortages, low-bandwidth connectivity, and limited technical ecosystems for support that ultimately limit their ability to benefit from AI-mediated learning innovations (Venugopal *et al.,* 2026). These intersecting hurdles exacerbate performance gaps and diminish the probability that at-risk students will use AI for personalized learning, exam preparation, or academic writing assistance, reinforcing systemic educational disadvantage.

The poll also identifies program-level stratification. Distance and open-university students, who make up an increasingly large proportion of India's higher education enrollment, report increased monthly but lower daily use of AI tools. This points to a pattern of "episodic AI engagement," in which students use AI tools for specific tasks like preparing assignments but do not have the ongoing exposure that is typical of environments with institutional support. This pattern is consistent with research showing that distant learners in the Global South typically use mobile-first access or public digital infrastructures, which results in uneven use of high-bandwidth or interactive AI platforms (Zhang & Hu, 2024).

When taken as a whole, these results show how AI-driven learning environments in India both reproduce and, in certain situations, worsen institutional and socioeconomic disparities. The unequal distribution of AI access and familiarity among programs, institutions, and social groups raises the possibility that AI is an agent influencing educational paths rather than a value-neutral addition to higher education. Such discrepancies run the risk of creating a technologically mediated hierarchy in the absence of clear intervention, where privileged students gain from regular, high-quality interactions with AI while marginalized groups are systematically shut out of the educational and cognitive advantages of AI-enabled instruction. The need for decolonial and justice-oriented approaches to AI integration in Indian higher education is highlighted by this layered inequality, which guarantees that AI will increase rather than maintain educational stratification through more equal access to education.

5.4 Surveillance, Discomfort, and the Criminalization of Bodies

The findings demonstrated that AI-enabled proctoring and monitoring systems function as both technical and sociopolitical tools, replicating past surveillance patterns that target and punish marginalized groups. When taking the exam under AI supervision, a sizable portion of the participants expressed discomfort: 38.93% reported discomfort, and 21.23% said they were extremely uncomfortable, indicating high levels of anxiety regarding digital surveillance. This degree of discomfort was common in both government and private institutions, as most have adapted to AI-based proctoring and do not retain specific institutional practices. This trend is in tune with the current scholarship that evidences how algorithmic surveillance reconfigures the classroom into a "digital panopticon" in which students come to internalize a sense of constant visibility and potential suspicion (Chaka, 2022). Figure 3 shows the Student Discomfort with AI-Based Proctoring Systems.

Table 4 and Table 5 compares the proctoring discomfort observations. Chi-square analysis confirmed significant associations between community category and proctoring discomfort [χ^2 (8, N=113) = 19.73, p = 0.011, Cramér's V = 0.30]. SC/ST students reported discomfort at higher rates (77.3%) compared to General category students (52.1%).

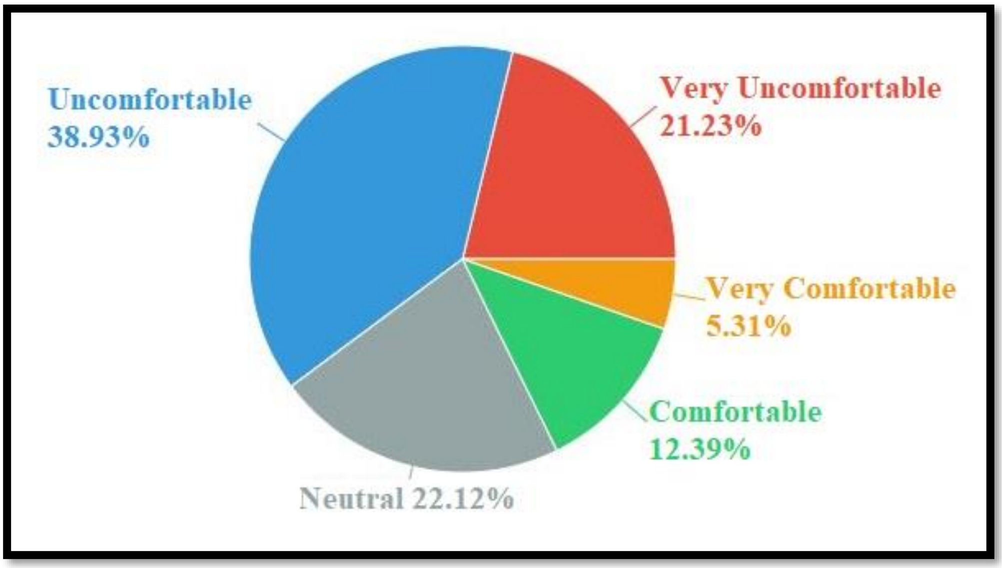

Figure 3. Student Discomfort with AI-Based Proctoring Systems

Table 4. Proctoring Discomfort by Student Background

Background Characteristic	VeryUncomfortable (%)	Uncomfortable (%)	CombinedDiscomfort (%)	N
SC/ST	31.8	45.5	77.3	22
Tribal/Indigenous language	41.7	33.3	75.0	12
First-generation learner	26.5	41.2	67.7	34
Rural background	30.0	35.0	65.0	20
Government institution	22.5	42.5	65.0	40
General category (urban)	14.6	37.5	52.1	48

Table 5. Logistic Regression Predicting Proctoring Discomfort

Predictor	B	SE	Wald	p	OR	95% CI
SC/ST category	0.85	0.39	4.76	0.029	2.34	1.09-5.03
Tribal/Indigenous language	1.12	0.58	3.73	0.053	3.06	0.98-9.56
Rural background	0.68	0.47	2.09	0.148	1.97	0.79-4.92
Low-light home environment*	1.24	0.42	8.71	0.003	3.46	1.52-7.87
Previous misrecognition experience	0.91	0.36	6.39	0.011	2.48	1.23-5.01
Government institution	0.47	0.35	1.81	0.178	1.60	0.81-3.16
Constant	-1.87	0.51	13.44	<0.001	0.15	-

Model fit: $\chi^2(6) = 24.68$, $p < 0.001$; Nagelkerke $R^2 = 0.31$

*Self-reported inadequate lighting in primary study location

The major predictors of proctoring discomfort among students were revealed in the analysis. Students from SC/ST categories had 2.34 times higher odds to report discomfort during proctoring, with $p = 0.029$. In a similar vein, exams taken at low light conditions have 3.46 times higher odds of being associated with discomfort, with a value of $p = 0.003$. More importantly, the hypothesis of compounded algorithmic harm is supported by the fact that a student's prior experience with linguistic misrecognition significantly predicted their subsequent proctoring discomfort ($p = 0.011$). Additionally, the concentration of this discomfort was highlighted by an intersectional analysis. 100% of the extremely small subgroup of students ($n = 6$) who were SC/ST, from rural backgrounds, and whose first language was Tribal/Indigenous expressed discomfort with proctoring. This stands in high contrast to the

52.1% rate of discomfort by students belonging to the General category, living in urban areas, and whose primary language was English or Hindi.

Across all linguistic, caste, and regional backgrounds, students consistently spoke of the fear of being wrongly flagged for "suspicious behaviour," but several linked the issue to environmental and socio-material conditions i.e., low lighting, unstable internet connectivity, and background noise, that disproportionately affect learners from economically disadvantaged and rural households. This echoes wider critiques of the way in which proctoring software design assumes the normativity of Western, urban, and well-resourced conditions of learning, therefore penalizing students who do not inhabit those environments (Darling-Hammond *et al.,* 2019). The inability of algorithmic systems to recognize faces in low-light settings or with lower-resolution cameras is also consistent with documented skin-tone and racial biases in facial recognition technologies, in particular the reduced accuracy on darker-skinned individuals as demonstrated in recent empirical work (Perkowitz, 2021). Although this study does not collect biometric data, participant narratives show strong resonance with the literature in indicating that the misrecognition experiences faced by Indian students echo broader, globally observed algorithmic inequities.

For most respondents, the proctoring systems induced a sense of criminalization rather than support. Students reported an affective feeling of being treated as a cheater by default, consistent with scholarship on algorithmic surveillance that frames AI-based monitoring as automated suspicion that disproportionately burdens marginalized populations (Bareis & Katzenbach, 2021). The AI-proctoring systems worsened inequality in the Indian context, where students from underrepresented groups frequently encounter major barriers to private study spaces, by misinterpreting everyday household activities like moving around in shared rooms or having background conversations as proof of candidate misconduct. These circumstances increase students' anxiety, especially during high-stakes exams, and are in line with recent research showing how algorithmic proctoring technologies alter students' emotional and cognitive states, frequently impairing performance through stress and hypervigilance (Druga *et al.,* 2022).

Misclassification was also found to be caused by other cultural and religious markers. In line with findings from other international studies that have tended to show that AI-powered vision systems frequently misrecognize the appearance of people wearing non-Western attire or head coverings, participants who wore hijabs, bindis, and turbans reported face detection issues more frequently (Yang, 2025). These failures demonstrate how AI systems increasingly marginalize students whose identities differ from Western-centric standards embedded in training datasets by operationalizing a culturally limited concept of a "valid" student body. This restriction emphasizes the relationship between algorithmic surveillance and cultural othering, a digitally mediated hierarchy in which some bodies become intelligible and reliable while others remain unclear or suspicious.

The evidence shows that students who experienced misrecognition in proctoring contexts frequently reported meeting prejudice in AI feedback in general, indicating a cumulative burden of algorithmic damage. This is consistent with theoretical arguments that algorithmic systems generate interconnected realms of oppression that reinforce each other across educational contexts, rather than isolated forms of bias (Tweissi *et al.,* 2022). In other words, students who are algorithmically marginalized in one domain (e.g., facial recognition) are more likely to be marginalized in others (e.g., language processing, cultural representation), making clear that algorithmic oppression is decidedly multidimensional in nature. The psychological impact of AI-mediated surveillance was pervasive. Many students described states of hypervisibility, stress, and fear of false accusation-emotional states in line with the concept of "surveillance-related trauma" emerging in contemporary studies of remote examination technologies (Kim *et al.,* 2025). In turn, trauma amplifies for students whose linguistic, cultural, or socioeconomic identities already locate them within systems of scrutiny. Evidence points that AI-based proctoring does not merely fail to offer equal academic monitoring but reproduces colonial structures of discipline, suspicion, and bodily control in digital learning environments.

5.5 Resistance, Adaptation, and Coping Strategies

Learners in this study actively negotiated the algorithmically structured learning environment, employing diverse adaptive and resistant practices. This negotiation occurred despite documented feelings of inequity and misrecognition. In order to avoid algorithmic penalties in automated scoring, grammar correction systems, and generative AI platforms, more than half of participants stated that they routinely reword their sentences to conform to "standard" or "neutral" English. This behavioural adaptation is consistent with recent research showing that students frequently internalize the linguistic hierarchies ingrained in machine-learning models, rearranging their communication practices to conform to algorithmically preferred norms, frequently at the price of cultural expressiveness and authenticity (Liu, 2025). Such behaviours demonstrate how AI systems discipline linguistic behaviour in subtle ways, resulting in an implicit curriculum of linguistic compliance that disproportionately affects non-dominant dialect speakers. Figure 4 compares student resistance and adaptation strategies.

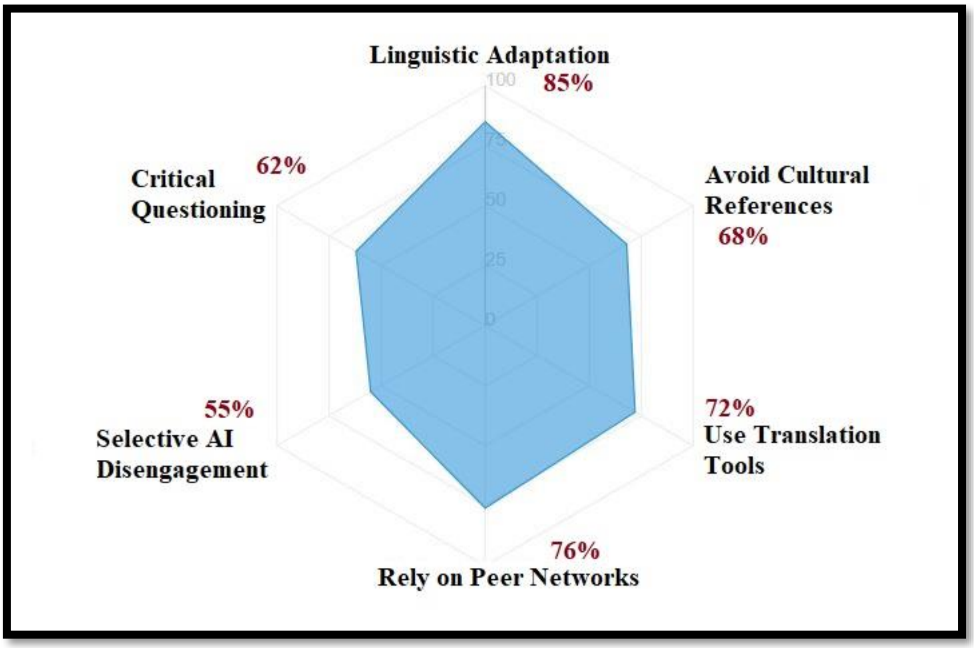

Figure 4. Student Resistance and Adaptation Strategies

The data also reflect a tendency for a substantial number of students, particularly those classified as "first-gen" college students and those who spoke a non-English language in their homes, to consciously eschew using cultural tropes in their submissions and AI-assisted activities due to a concern that those technologies would misunderstand, downplay, and/or penalize them for doing so. This reflects a set of findings regarding algorithmic erasure, in which "marginalized individuals are prompted to suppress their characteristics in order for their "legibility" to be guaranteed" as a means of being legible in algorithmically-mediated spaces, rather than being seen as a target for discriminatory dissonance, as explored in this body of scholarship (Bareis & Katzenbach, 2021). These students reported that using local history, caste knowledge, indigenous knowledge systems, and/or locally situated examples would often result in "off-target" corrections, decreased grades, and/or decreased quality in their prompts, further solidifying the understanding that those technologies are set up with a Eurocentric epistemology.

Another set of strategies that were also common involved using translation paraphrase software, but this was for a "protective" reason, to effectively "neutralize" regional linguistics in order to feed content into an AI system. More than 40% of respondents use software like Google Translate, Grammarly, and/or Quillbot purely for covering up "linguistic dialectical markers and/or Mother Tongue Interaction." Although it has a function for promoting access, being "intermediaries as a necessity" signifies how students provide their own ways of using Couldry and Mejias' "data colonialism" in which "one adjusts their own linguistic expression of self in order to conform to expectations of systems built around a dominant cultural norm" (Couldry & Mejias, 2018). This means that AI not only enables communication but also reconfigures it, as it governs those linguistic communications that are deemed legitimate in an academic institution.

On the other hand, it was also noticed that many students were developing a certain skepticism about AI-based tools, with nearly 19.47% of students claiming that they never use AI, and some students admitting that they use AI very selectively. This trend aligns with findings obtained from cross-national studies that reveal how algorithm-based systems marginalize students due to perceptions of bias, even when this bias is not always true, in order to create uneven access to technology-based opportunities for learning (Belenguer, 2022). In this Indian setting, this takes on a certain significance, since it creates a self-reinforcing cycle where those who are in most pressing need of support from AI are also those who are being alienated from it. Some of the other important counter spaces that were established were peer networks and human support systems. Here, students would not completely rely upon the suggestions that were produced by the AI algorithm but would check with their peers, tutors, and teachers. Such resistance has been referred to in critical studies of digital literacy, which has shown that human support systems are a means of resistance in algorithm-driven oppression, but it also represents a site of agency in this oppression (Rapanta *et al.,* 2025). Thus, with a focus on human support, students would develop their own environments of autonomy in epistemology.

However, a smaller but equally significant subset of students resisted even more, adopting even more overtly resistance-minded tactics like "gaming" algorithmic recommendations, critically evaluating system output, or completely avoiding algorithm-driven proctoring and surveillance. These actions fall under the category of "algorithmic disobedience," which refers to users' deliberate attempts to cause disruption and refuse to comply with a particular digital system as a form of political and epistemological resistance (Romanishyn *et al.,* 2025). A growing awareness of how racial, class, and linguistic disparities are encoded in AI was largely responsible for this rejection rather than a lack of digital literacy. The strategies of coping and resistance that were identified in this research are all indicative of algorithmic bias as a lived experience that impacts linguistic practices, cultural expression, and academic identity. These students are actively working with their own positioning in AI-enabled environments of learning, in a way that takes turns between accommodation and resistance, as a means of highlighting, in a powerful way, how subjects and agents are not only enmeshed in questions of power, privilege, and agency, but also in online education environments. This research provides a critical viewpoint on marginalized students as AI subjects while also emphasizing their agency as subjects critical of AI systems that want to manage them in a specific way.

5.6 Synthesis of Quantitative Patterns

Before proceeding with theoretical explanation, we will first synthesize key data tendencies from our research. Linguistic Stratification: Algorithmic injury occurs along a unique linguistic divide, with speakers of tribal/Indigenous languages being most negatively impacted by misrecognition (OR= 8.85), followed by speakers of regional languages (OR= 2.44), and those speaking English being the least harmed. These gradations echo linguistic divides unearthed in postcolonial studies of education. Intersectional Compounding: Students bearing multiple intersecting marginalized identities are subject to multiply rather than merely additively harmful algorithmic effects. For example, SC/ST status + tribal language + rural background yield 100% proctoring discomfort rates, against 52.1% in privileged groups. Figure 5 shows the Intersectional Patterns of Algorithmic Harm.

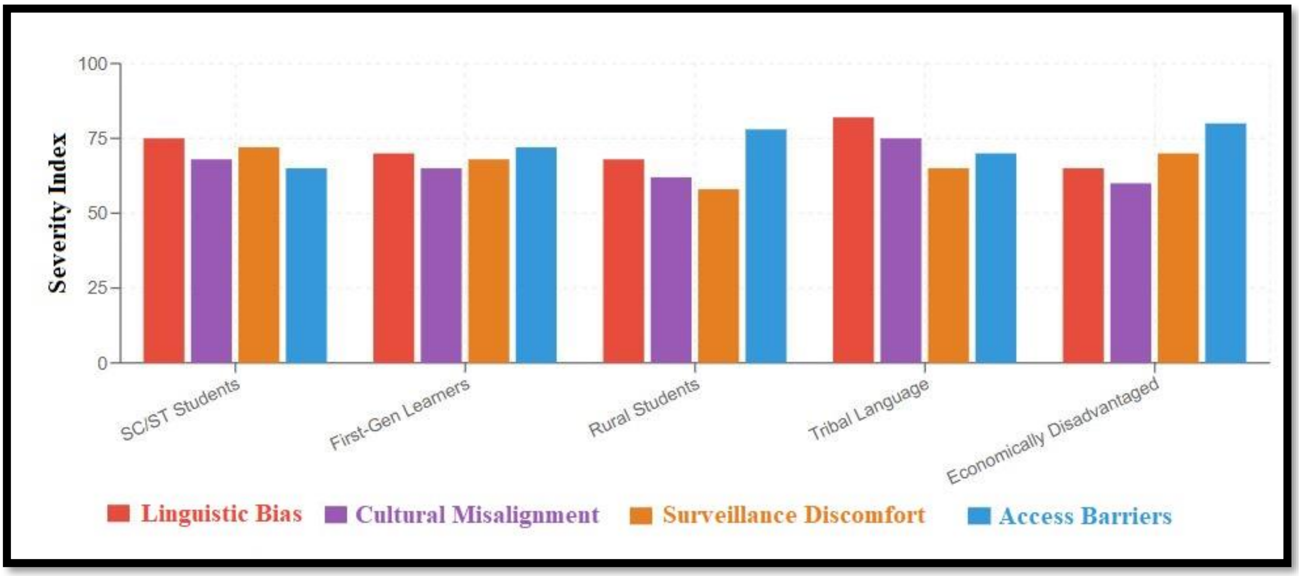

Figure 5. Intersectional Patterns of Algorithmic Harm

Cross-domain correlation: Experiences of bias in one domain predict bias in others. Students who reported high linguistic misrecognition were 2.48 times more likely to experience proctoring discomfort (p=0.011), suggesting algorithmic systems create interlocking fields of oppression rather than isolated incidents. Institutional Mediation: Whereas the overall differences between government versus private institutional type did not reach significance in our multivariate models (p = 0.178), descriptive patterns confirm that 65.0% of students attending government institutions experienced discomfort with proctoring compared to 55.3% at private institutions, indicating institutional resources partially buffer algorithmic harms. Resistance as Norm: More than 85% of students employed at least one adaptation strategy, with linguistic self-censorship near universal among non-English speakers. This would imply that mitigating algorithmic bias has now become normalised labour for these marginalised students. These patterns offer empirical validation for our theoretical predictions outlined in Section 2.5, and help to confirm that AI-mediated environments reproduce a colonial hierarchy.

6. Discussion

Our results offer powerful evidence to support a conclusion that AI-mediated Learning Environments in Indian Higher Education are far from being a set of exemplary, value-free, teaching technology tools; rather, our data provides evidence that AI-mediated Learning Environments are far from being passive agents of teaching but are themselves actively perpetuating colonial structures of power. The set of statistical associations established in Section 5, specifically 8.85 for misrecognition among speakers of tribal languages and 100% for proctoring discomfort among multiply marginalized students, provide potent evidence that algorithmic bias patterns are themselves a symptom of epistemic violence. "Linguistic stratification, as evidenced in our data in Table 2, literally confirms postcolonial theories of epistemic violence perpetrated through languages." The near 9-fold increase in risks of misrecognition by AI for speakers of tribal/Indigenous languages over Hindi speakers, with negligible effects for English speakers, particular to this article, substantiates a postcolonial theorized silencing of subaltern voices in Spivak's "Can the Subaltern Speak?" from 1988 (Piu, 2023). Our findings can be seen to develop her theorizing in that this "silencing of subaltern voices has been automated in digital technology."

Comparison with Previous Studies: While our results are consistent with those of Koenecke *et al.* in 2020 in showing that African American Vernacular English has been systematically misrecognized with error rate differentials of 19-35%, our results expose even sharper patterns in India. Koenecke, Nam, Lake, Nudell, Quartey, Mengesha, Toups, Rickford, Jurafsky, and Goel tested "whether five commercial, state-of-the-art automated speech recognition systems, developed by Amazon, Apple, Google, IBM, and Microsoft, could accurately transcribe structured interviews, finding that all five systems performed severely differently for Black and White speakers, with a 0.35 vs. 0.19 average word error rate for Black vs. White speakers" in a study published in (Koenecke *et al.,* 2020). Although Koenecke et al. found that there are racial disadvantages in automated ASR in a homogeneous country with a fifty-two-million-strong population, our findings expose how multilingual postcolonial nations erect sharper hierarchies. The 58.3% "very frequent" misrecognition rate for speakers of tribal languages, as opposed to 10% for English, translates to a 5.8-fold discrepancy, which exceeds most other reported cross-national differences.

The above patterns verify our theoretical frame work's assertion that "whiteness as property" is instantiated in "linguistic norms" by AI, H1. This means that knowledge of "English," which has long been defined as "cultural capital" through "colonial education," has now become "algorithmic capital" in order for an individual to effectively use "AI-mediated learning." This finding complicates techno-optimist narratives of positioning AI as democratizing by showing how AI amplifies existing linguistic stratifications. The first major finding-pervasive linguistic misrecognition-aligns directly with Paulo Freire's critique of the "banking" model of education, wherein dominant systems define what counts as legitimate knowledge and language. AI technologies trained disproportionately on Western or conventional English corpora have a similar silencing effect, penalizing or misinterpreting Indian linguistic variation. This supports Freire's claim that oppressive institutions work by silencing minority voices and forcing conformity to dominant modes of communication. Students' purposeful rephrasing of words to sound like standard English, as well as their avoidance of culturally grounded instances, suggest a new sort of linguistic domestication in which AI systems influence communication based on the norms of an algorithmically favoured elite.

Postcolonial theory expands on these dynamics by situating algorithmic misrecognition within the context of epistemic violence. Spivak and Bhabha's writings demonstrate how colonialism continues through frameworks that invalidate subaltern knowledge systems. The dataset reveals that students from Indigenous, rural, and multilingual backgrounds had higher misunderstanding rates, demonstrating how AI's epistemic architecture aligns with what Couldry and Mejias refer to as "data colonialism": knowledge extraction and hierarchization in ways that extend the logics of colonial extractive economies (Couldry & Mejias, 2018). The AI-generated content's lack of consideration for Indian cultural, historical, and ethical contexts, combined with participants' perceptions of Western cultural dominance, highlights how algorithmic systems reproduce the very cultural erasures that postcolonial scholars have long criticized.

Critical race theory in education adds another important layer to this research. Even though the dataset does not collect explicit racial data, the patterns of misrecognition reported particularly in AI-based proctoring, where students with darker skin tones or low light conditions face more frequent flagging, are consistent with international findings on racially biased facial recognition systems. Scholars have already documented serious accuracy differences in commercial face recognition systems, showing much higher error rates across different skin tones and genders. For instance, across major commercial systems, the error rate for darker-skinned women was 34.7%, dramatically higher than the 0.8% error rate for lighter-skinned men (Obermeyer & Sendhil, 2019). The emotional toll reported by our participants—their fear of being suspected, the anxiety of being misread, and the discomfort of constant monitoring—mirrors the concepts of Critical Race Theory, which explains how surveillance disproportionately targets and criminalizes marginalized groups. In the Indian context, where caste, class, and regional identities intersect with skin tone and language, algorithmic misrecognition deepens already existing forms of social stratification.

Critical data studies further this critique by demonstrating how AI systems implement forms of surveillance capitalism that are fundamentally incompatible with equitable education. The students' feeling of unease with proctoring systems, the sense of being perpetually watched, and the worry about being falsely flagged align with models of digital governance that prioritize institutional control over student autonomy. Studies on digital education governance reveal how technology alters educational oversight through data collection and new accountability measures, often undermining student privacy and freedom (Williamson, 2015). Crucially, our data highlights that experiencing misrecognition in proctoring is linked to a higher frequency of general AI bias reports. This suggests that algorithmic oppression is not a collection of isolated incidents; instead, it is a silent, accumulating issue across multiple domains, indicating that AI systems function as interconnected surveillance infrastructures rather than simple, distinct tools. The study's findings regarding students' adaptive and resistant strategies offer hope for decolonial transformation in AI-mediated education. For example, the participants' preference for relying on peers and teachers over AI recommendations, their selective rejection and use of certain AI tools, and their developing critical awareness of algorithmic biases all demonstrate an emerging sense of digital agency. These practices align with critical digital literacy, which emphasizes the students' ability to question and challenge algorithmic power structures. Such actions reveal that learners are not passive recipients of AI authority but are active challengers, negotiators, and resistors of digital coloniality.

These insights lead us to synthesize the findings and call for a decolonial approach to educational AI. This approach must move beyond mere technical correction and focus on restructuring the core knowledge foundations. Decolonial AI demands the development of pluriversal design methodologies, where diverse ways of knowing, speaking, and learning are purposefully embedded in the datasets, training pipelines, and evaluation criteria. Above all, it requires shifting AI's role from a tool that enforces standardization to one that supports relational, culturally sustaining, and critically reflective pedagogies. Ultimately, this research confirms that algorithmic bias in education is not an accidental outcome of poor data but a structural result of technologies developed within a framework of unequal global power relations. To achieve a just AI-mediated learning environment, educational institutions, policymakers, and developers must confront the legacies of colonialism. They must reconceptualize AI not as an instrument of discipline, prediction, and control, but as an opportunity for liberation—rooted in equity, cultural integrity, and epistemic plurality.

7. Implications and Recommendations

The empirical patterns identified in this study urgently necessitate systemic changes in AI-mediated learning environments. Specifically, the high incidence of linguistic misrecognition, cultural misalignment, proctoring-related discomfort, and the dominance of Western pedagogical outputs signal that current educational AI is designed to reproduce—rather than fix—longstanding colonial power structures, undermining the goal of equitable learning. These findings demand immediate interventions at the levels of pedagogy, policy, research, and technological design. Adopting a decolonial perspective toward AI in education requires prioritizing epistemic justice, multilingual inclusion, cultural sustainability, transparency, and community participation.

7.1 Pedagogical Implications: Toward Critical AI Literacy

Our research highlights the need for pedagogical reforms that go beyond simply training students to use AI tools. We must foster critical, reflexive, and justice-oriented digital competencies to address the linguistic marginalization, cultural misalignment, and surveillance distress we found. For instance, recent work shows that students who receive structured critical AI literacy training—focused on how data is gathered, how models are trained, and how power operates within algorithmic systems—are more likely to develop the deep competencies needed to identify and challenge biased outputs, interpret AI-generated information ethically, and avoid over-reliance on automated systems (Dasgupta & Hill, 2020). Given India's higher education context, where deep linguistic diversity intersects with entrenched social hierarchies, this literacy must be framed not as a technical skill but as an epistemically empowering practice through which students can critically interrogate the sociocultural assumptions embedded in AI system design.

This is corroborated in our empirical results. Most participants described the frequency of misrecognition or linguistic distortion; more than two-thirds reported changing grammar and/or vocabulary to fit perceived "standard English" norms, and/or culturally specific references. These adjustments are what critical pedagogy terms "linguistic assimilation pressures"-learners internalise dominant norms so as not to be penalized by algorithms. Critical AI literacy modules embedded into existing curriculum-particularly within the first-year undergraduate curriculum-can counter such pressures by inviting students to consider the biases in training data, the auditing of output from AI models, and the understanding of the political economies of global AI development. Recent research shows that such pedagogical interventions strengthen agency and collective resistance among students as they are being shaped toward algorithmic inequity, bringing forth dispositions consistent with Freirean problem-posing education that encourage critical consciousness (conscientização) over passivity (Williamson, 2015).

The increased level of cultural misalignment reported by students is indicative of the importance of culturally sustaining pedagogy in AI-mediated classrooms. In reviewing AI outputs for their alignment with Indian epistemologies, local histories, and community knowledge, learners are afforded possibilities of identifying Western overrepresentation in training corpora and resisting the implicit hierarchies encoded in AI systems. Class-based audits of AI text, picture, or scoring outputs might thus function as both a learning approach and a participatory oversight mechanism, allowing students to generate proof of algorithmic coloniality. These are pedagogical ways that resonate with growing scholarly consensus about how algorithmic literacy and culturally responsive digital pedagogy are important countermeasures to epistemic injustices perpetrated through global AI infrastructures.

7.2 Policy Recommendations

The empirical patterns described in this study call for multi-level institutional, state, and national policy responses, considering the growing deployment of AI services across Indian higher education. The extent of misrecognition faced by tribal-language and non-English speakers-as reflected in odds ratios greater than 8.0 in our dataset-reflects the insufficiency of current AI systems for linguistically diverse populations. This aligns with global results on how performance differentials in algorithms disproportionately marginalize speakers of under-represented languages and accents (Naveena *et al.*, 2021). Institutions are, therefore, duty-bound to undertake AIA as a mandatory step before procurement or deployment of educational AI services. Such impact assessments must involve multilingual fairness testing, bias auditing under conditions of low resource use, and disaggregated harm analysis for SC/ST, tribal, first-generation learners, etc. Scholarship on various AI governance models increasingly flags this type of pre-implementation review as essential to prevent structural discrimination and adherence to standards of ethics in use (Mennella *et al.*, 2024). Given the significant link between proctoring systems and increased discomfort among marginalized students, institutions are under an obligation to introduce tight procurement guidelines in line with the principles of algorithmic sovereignty, including the training of AI systems on linguistically and culturally diverse datasets that maintain residency in India, along with model cards that document known risks or performance gaps. Various research has proved that the procurement requirements related to transparency, representation, and accountability can always push vendors to alter the training pipelines, integrate non-English data, and perform independent audits (Gradwohl & Tennenholtz, 2022).

The use of AI in large-scale, high-stakes educational applications needs to be overseen by regulatory mechanisms both at the state and at national levels. International research warns that unregulated educational AI tools can entrench existing inequalities and expose students to harmful surveillance practices (Akgun & Greenhow, 2021). A tiered regulatory framework would prohibit high-risk applications such as automated proctoring for students in low-resource environments, require approval for adaptive scoring or predictive systems, and permit only low-risk informational tools.

Finally, there is a need for a national mandate for linguistic justice in educational AI. A literature review has shown that language-driven algorithmic inequalities result in structural exclusion in online learning spaces. A policy requiring support for Indian languages in AI can help minimize error rate differentials between Anglo-American and Indian languages, and policies that incorporate tribal/endangered languages in development agendas can help promote multilingual equality in line with Indian constitutional values while challenging Western-driven linguistic standardizations in worldwide AI architectures.

7.3 Research Implications: Methodological Imperatives for Decolonial AI Scholarship

Findings point towards a pressing need for a paradigm shift in the approaches used in AI in education studies. The prevailing approaches in current studies continue to reinforce a positivist model of investigation, with a focus on technical efficiency rather than lived experience. This study has shown how community participation-based approaches in Mixed Methodology models, such as CPAR, offer a richer understanding of how AI influences educational realities. Further studies are needed in this area, with a continuing longitudinal focus on how students are impacted by biased AI systems. Furthermore, in addition to this, intersectional patterns that emerge from this data, in relation to language, socio-economics, and institution types, are also indicative of how important it has been to use intersectionality in quantitative studies. The oppression with regards to algorithmic studies has, as has been highlighted with rising trends in literature, a multi-dimensionality that, as has been cited, “cannot be adequately captured by traditional analyses along a single axis” (Arjun *et al.,* 2023). With this in mind, it becomes important that those individuals most affected are consulted, as part of a body of research, from question definition through to interpretation and resolution design, in order to align with increasing pressures for community.

7.4. Recommendations for Design of Technology: Decolonial AI Principles

Technological implications of this research are much broader than mere bias correction. A decolonial approach to AI involves a fundamental transformation of how educational technologies are envisioned, developed, and evaluated. Firstly, it involves compliance with pluriversality in developing, where various epistemologies, forms of expression, and cultural references are incorporated. This involves developing datasets, which are representative of linguistic diversity in India. Also, incorporation of Indigenous and Local Knowledge systems into AI-produced content. Another key aspect that needs to be considered is that of transparency. The students felt a sense of distrust when it came to using proctoring technology and automated scoring. If AI decision-making in education becomes opaque, it poses a challenge to trust in AI technology in education. Educational AI needs to incorporate transparency by design, aligned with worldwide guidance, and it becomes a necessity from a learner’s point of view that “they must know why and how decisions are taken”. Lastly, educational AI needs democratized development and needs to move from corporate-controlled models to community-controlled models. There needs to be a shift from corporate-controlled models of AI development to community-controlled data governance in education. Community-controlled bodies would then monitor data development, model testing, and ethical considerations of AI development, ensuring that AI technology in education benefits equity in education and not corporate interests. According to critical data studies, “transferring ownership from extractive data regimes represents a key step in reversing data colonialism” (Muldoon & Wu, 2023).

8. Limitations & Future Research Directions

This research offers important insights into structural injustices in AI-mediated education, but it also contains important limitations that must categorically structure how this data ought to be interpreted. First, the method of sampling and recruitment-a digitally distributed survey-is likely to have generated a selection bias by excluding students without stable access to the internet or with poor devices (Bethlehem, 2010). Those most likely to be harmed by algorithms are thus under-sampled, such as learners from rural areas or socio-economically disadvantaged backgrounds. Furthermore, given only approximately 18% of the respondents self-identified as rural, which is below national estimates of rural enrolment in higher education, our findings likely represent a conservative lower bound of digital inequality and algorithmic harm. This underrepresentation implies that in wider population-level estimates, the prevalence of misrecognition, surveillance discomfort, and cultural alienation may be significantly higher. Moreover, the study relies exclusively on self-report measures, rather than on objective, technical audits of AI system performance or algorithmic error rates. While subjective experience is epistemologically valid-particularly in a Community-Participatory Action Research (CPAR)-oriented framework-it precludes definitive attribution of misrecognition and bias to AI systems rather than to user-side constraints such as poor bandwidth, device hardware, lighting conditions, etc,. In the absence of parallel technical audits, performance testing across devices and conditions, or fairness analyses across languages and demographic groups, the findings remain interpretive and indicative rather than confirmatory. Future work should include algorithmic fairness audits that use frameworks in a systematic manner to discover and decrease biases embedded in algorithms across more languages and sociodemographic contexts.

A second limitation is measurement validity. The questionnaire contained numerous multi-option and multi-select items. It had not been formally psychometrically validated, including reliability analysis, pilot testing, or checks for consistency. Terms like "very frequently" and "sometimes" are without standardized operational definitions; therefore, variation in interpretation by respondents might have occurred (Mei & Zhang, 2025). This reduces the precision, reliability, and comparability of data. Development of a validated standardized scale-including, for example, Likert-type constructs with defined reference periods and anchor definitions, as investigated in recent contemporary psychometric practice-and measures of behavioural frequency, each with well-defined response anchors, would improve internal consistency. Such instruments should be used in future studies and reliability measured by Cronbach's alpha or coefficient omega.

The cross-sectional design further constrains the study's power to capture the temporal dynamics. Data reflect one point in time, and therefore cannot reveal how repeated exposure to biased AI systems may shape students' academic trajectories, linguistic identity development, psychological wellbeing, or long-term educational outcomes over time (Dutta *et al.,* 2021). It would take a few longitudinal cohort studies to establish whether an increase in cumulative algorithmic harm leads to an increased risk of dropout, decreased academic engagement, or internalized linguistic inferiority. Prospective designs afford better causal inference and can illustrate just how algorithmic marginalization compounds over time.

Demographic coding in this study limited intersectional analysis. While broad categories of SC/ST status or institutional type were captured, data were not disaggregated by specific caste sub-groups, religious minorities, gender identity, or disability status that are particularly salient in the Indian context. Our research faced constraints due to ethical concerns regarding data security and participant anonymity, which limited our choice of methodology. This remains a significant limitation, as it prevented a deep exploration of how complex intersections of caste, religion, language, gender, and class contribute to algorithmic disadvantage. Future studies must address this gap by adopting an ethically robust, intersectional data collection framework. This framework should prioritize participant privacy and agency while fully capturing the multi-dimensional nature of algorithmic harm. Epistemologically, while this study was grounded in the philosophy of Community-Based Participatory Action Research (CPAR), reducing CPAR to a remote, survey-based format inherently curtailed the methodology's transformative, dialogical, and co-analytic dimensions (Fine & Torre, 2021). We could not conduct collective reflection, shared analysis, or community-led action to translate the findings into immediate pedagogical or institutional change. Thus, the work was largely diagnostic rather than interventionist. This underscores the urgent need for subsequent research to execute full CPAR cycles including focus groups, participatory workshops, collaborative audits of AI tools, and the co-design of alternative, culturally sustaining AI systems based on sustained community engagement. Finally, the study’s focus specifically on Indian higher education limits its generalizability across diverse global contexts. Postcolonial societies vary widely in their colonial histories, linguistic environments, institutional structures, and socio-economic conditions. What is true for India may not directly apply to other Global South contexts. Therefore, there is an urgent need for comparative, cross-national research to differentiate between algorithmic harms that are universal and those that are heavily influenced by local history and context. Despite these limitations, this study offers a crucial empirical foundation that spotlights the issue of algorithmic inequality within educational AI. It presents robust, exploratory evidence of structural inequities and strongly motivates future scholarship focused on epistemic justice, participatory design, and the decolonial transformation of digital education systems. Future research should invest in mixed-method and participative methods that embed subjective experiences within technical audits, interdisciplinary analyses, and community engagement. Longitudinal cohort studies are needed to track the long-term effects of repeated AI exposure on linguistic identity, academic progress, and psychosocial well-being, using validated measurement instruments and standardized response anchors.

Ethical research frameworks must embed intersectional data collection to thoroughly examine the cumulative consequences of caste, class, language, gender, region, and disability. Privacy-preserving methodologies, such as k-anonymity and clustering-based approaches, should be integrated to safeguard participant confidentiality while enabling nuanced research into how algorithmic harms accrue across intersecting axes of social difference (Su *et al.,* 2023).On the design side, there is a critical need for decolonial AI development: systems that are multilingual, culturally grounded, community-governed, and transparent, designed to challenge rather than reinforce, epistemic hierarchies. Technical audits utilizing causal fairness frameworks and demographic parity assessments must be conducted across multiple languages, device types, and socioeconomic conditions to detect performance differences and mitigate them before deployment. Developers should follow pluriversal design principles, centering the voices of marginalized communities throughout all stages of algorithmic development from dataset curation to model evaluation and ongoing monitoring. Only comparative, cross-national research informed by post-colonial perspectives can effectively illuminate how algorithmic colonialism manifests across various Global South contexts while establishing universal principles for equitable AI governance that remain sensitive to contextual specificity. Comprehensive CPAR cycles must be enacted, allowing communities to act as co-researchers and co-designers, thereby translating research findings into institutional policy changes, decolonized pedagogies, and alternative technological infrastructures that champion cultural sustainability, epistemic plurality, and liberatory learning. Through such integrative, reflective, and justice-oriented inquiry, educational AI can shift from a tool of standardization and surveillance to a medium for equitable, pluralistic, and liberatory learning.

9. Conclusion

This paper provides strong empirical evidence that AI-mediated learning environments in India actively reproduce long-standing colonial hierarchies. We found measurable linguistic bias, significant cultural marginalization, and unequal surveillance practices. Our analysis of the 113-student sample revealed statistically significant patterns of harm that disproportionately affect communities already facing structural disadvantages. Specifically, students who speak tribal and Indigenous languages faced an eightfold increase in the odds of being algorithmically misrecognized. Similarly, students from Scheduled Caste (SC), Scheduled Tribe (ST) backgrounds, and rural areas reported heightened discomfort with automated proctoring systems. These findings confirm that AI systems used in education are not neutral tools; rather, they are sociotechnical infrastructures that embed epistemic violence. They privilege English-dominant, urban, and higher-status linguistic identities while consistently devaluing culturally diverse forms of expression. The clear mismatch between the students' cultural realities and the AI-generated content, coupled with reports of widespread linguistic self-censorship, demonstrates how AI subtly influences academic identity and reinforces Western-centric norms in knowledge production.

Therefore, this study integrates critical pedagogy, postcolonial theory, critical race theory, and critical data studies to establish a decolonial analytic framework. This framework allows us to interpret the quantitative patterns observed as clear manifestations of systemic inequality, not as simple, isolated technical errors. The immediate implications of our work call for educational institutions to take swift action: they must integrate critical AI literacy into curricula, adopt bias-aware policies for technology procurement, strictly regulate high-risk applications like automated proctoring, and actively support the development of truly multilingual and culturally grounded AI systems.

Simultaneously, the resistant and adaptive strategies used by the students themselves suggest inherent forms of agency, which can be strengthened through participatory and community-led approaches. Future research should prioritize conducting full Community-Based Participatory Action Research (CPAR) cycles, where learners are involved as co-researchers and co-designers of equitable educational technologies. Only through such collective efforts can we transform AI from a mechanism that reinforces coloniality into one that delivers epistemic justice and builds truly inclusive digital futures.

References

Abas, I.H., Krishnamurthi, N., Rasli, A., Gusteti, M. U. (2025). A Delphi study on factors influencing school students' adoption of social media as a learning platform in Malaysia. *International Journal of Evaluation and Research in Education (IJERE)*, 14(3), 1743. http://doi.org/10.11591/ijere.v14i3.32939

Ahrweiler, P., Spath, E., Siqueiros García, J.M., Capellas, B.L., Wurster, D. (2025). Inclusive technology co-design for participatory AI. *Participatory Artificial Intelligence in Public Social Services: From Bias to Fairness in Assessing Beneficiaries*, 35-62. https://doi.org/10.1007/978-3-031-71678-2_2#DOI

Akgun, S., Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Akgun, S., Krajcik, J. (2024). Artificial intelligence (AI) as the growing actor in education: raising critical consciousness towards power and ethics of AI in K-12 STEM classrooms. *In Uses of Artificial Intelligence in STEM Education*, 494-521. https://doi.org/10.1093/oso/9780198882077.003.0022

Amin, M.R.M., Ismail, I., Sivakumaran, V.M. (2025). Revolutionizing Education with Artificial Intelligence (AI)? Challenges, and Implications for Open and Distance Learning (ODL). *Social Sciences & Humanities Open*, 11, 101308. https://doi.org/10.1016/j.ssaho.2025.101308

Amorim, E., Cançado, M., Veloso, A. (2018). Automated essay scoring in the presence of biased ratings. *In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 229-237.

Ayanwale, M.A., Adelana, O.P., Molefi, R.R., Adeeko, O., Ishola, A.M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Computers and education open*, 6, 100179. https://doi.org/10.1016/j.caeo.2024.100179

Bareis, J., Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values,* 47(5), 855-881. https://doi.org/10.1177/01622439211030007

Belenguer, L. (2022). AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. *AI and Ethics*, 2(4), 771-787. https://doi.org/10.1007/s43681-022-00138-8

Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big*. In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 610-623. https://doi.org/10.1145/3442188.3445922

Bentley, C., Muyoya, C., Vannini, S., Oman, S., Jimenez, A. (2023). Intersectional approaches to data: The importance of an articulation mindset for intersectional data science. *Big Data & Society*, 10(2), https://doi.org/10.1177/20539517231203667

Berg, M. (2022). Digital technography: A methodology for interrogating emerging digital technologies and their futures. *Qualitative Inquiry*, 28(7), 827-836. https://doi.org/10.1177/10778004221096851

Bethlehem, J. (2010). Selection bias in web surveys. *International statistical review*, 78(2), 161-188. https://doi.org/10.1111/j.1751-5823.2010.00112.x

Boateng, O., Boateng, B. (2025). Algorithmic bias in educational systems: Examining the impact of AI-driven decision making in modern education. *World Journal of Advanced Research and Reviews*, 25(1), 2012-2017. https://doi.org/10.30574/wjarr.2025.25.1.0253

Chaka, C. (2022). Digital marginalization, data marginalization, and algorithmic exclusions: A critical southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths. *Journal of e-Learning and Knowledge Society*, 18(3), 83-95. https://orcid.org/0000-0003-3558-4141

Couldry, N., Mejias, U. (2019). Making data colonialism liveable: How might data's social order be regulated. *Internet Policy Review*, 8(2).

Couldry, N., Mejias, U.A. (2018). Data Colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, *20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Cunningham, J. (2018). Missing the mark: Standardized testing as epistemological erasure in US schooling. *Power and Education*, 11(1), 111-120. https://doi.org/10.1177/1757743818812093

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied developmental science*, 24(2), 97-140. https://doi.org/10.1080/10888691.2018.1537791

Dasgupta, S., Hill, B.M. (2020). Designing for critical algorithmic literacies. *arXiv preprint arXiv:2008.01719. https://doi.org/10.48550/arXiv.2008.01719*

Dixon-Román, E., Nichols, T.P., Nyame-Mensah, A. (2020). The racializing forces of/in AI educational technologies. *Learning, Media and Technology*, 45(3), 236-250. https://doi.org/10.1080/17439884.2020.1667825

Druga, S., Christoph, F.L., Ko, A.J. (2022). Family as a Third Space for AI Literacies: How do children and parents learn about AI together. *In Proceedings of the 2022 CHI conference on human factors in computing systems*, 1-17. https://doi.org/10.1145/3491102.3502031

Dutta, P., Quax, R., Crielaard, L., Badiali, L., Sloot, P. (2021). Inferring temporal dynamics from cross-sectional data using Langevin dynamics. *Royal Society open science*, 8(11). https://doi.org/10.1098/rsos.211374

Fine, M., Torre, M.E. (2021). Critical participatory action research: Conceptual foundations. In M. Fine & M. E. Torre, Essentials of critical participatory action research, American *Psychological Association*, 3–19. https://doi.org/10.1037/0000241-001.

Fregoso Bailon, R.O., De Lissovoy, N. (2018). Against coloniality: Toward an epistemically insurgent curriculum. *Policy futures in education*, 17(3), 355-369. https://doi.org/10.1177/1478210318819206

Ghasemaghaei, M., Kordzadeh, N. (2024). Understanding how algorithmic injustice leads to making discriminatory decisions: An obedience to authority perspective. *Information & Management*, 61(2), 103921. https://doi.org/10.1016/j.im.2024.103921

Gradwohl, R., Tennenholtz, M. (2022). Pareto-Improving Data-Sharing. *In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 197-198. https://doi.org/10.1145/3531146.3533085

Harari, L., Lee, C. (2021). Intersectionality in quantitative health disparities research: A systematic review of challenges and limitations in empirical studies. *Social science & medicine*, 277, 113876. https://doi.org/10.1016/j.socscimed.2021.113876

Harris, C.I. (1993). Whiteness as property. *Harvard law review*, 1707-1791. https://doi.org/10.2307/1341787

Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, 28(4), 688-704. https://doi.org/10.1080/02681102.2022.2068492

Joshi, A., Dabre, R., Kanojia, D., Li, Z., Zhan, H., Haffari, G. Dippold, D. (2025). Natural language processing for dialects of a language: A survey. *ACM Computing Surveys*, 57(6), 1-37. https://doi.org/10.1145/3712060

Karusala, N., Seeh, D.O., Mugo, C., Guthrie, B., Moreno, M.A., John-Stewart, G., Inwani, I., Anderson, R. Ronen, K., (2021). That courage to encourage: Participation and Aspirations in Chat-based Peer Support for Youth Living with HIV. *In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 223, 1-17. https://doi.org/10.1145/3411764.3445313

Kim, B.J., Kim, M.J., Lee, J. (2025). The dark side of artificial intelligence adoption: linking artificial intelligence adoption to employee depression via psychological safety and ethical leadership. *Humanities and Social Sciences Communications*, 12(1), 1-14. https://doi.org/10.1057/s41599-025-05040-2

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J.R., Jurafsky, D. Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the national academy of sciences*, 117(14), 7684-7689. https://doi.org/10.1073/pnas.1915768117

Komljenovic, J., Birch, K., Sellar, S., Bergviken Rensfeldt, A., Deville, J., Eaton, C., Gourlay, L., Hansen, M., Kerssens, N., Kovalainen, A., Nappert, P.L., Guerra., J.P. Poutanen., S. Robertson., S., Tyfield., D., Williamson., B. (2024). digitalised higher education: key developments, questions, and concerns. *Discourse (Abingdon)*, 46(2), 276–292.

Lehtiniemi, T., Ruckenstein, M. (2018). The social imaginaries of data activism. *Big Data & Society*, 6(1). https://doi.org/10.1177/2053951718821146 .

Lester, C.A., Flynn, A.J., Marshall, V.D., Rochowiak, S., Bagian, J.P. (2022). Implementation outcomes of the Structured and Codified SIG format in electronic prescription directions. *Journal of the American Medical Informatics Association*, 29(11), 1859-1869. https://doi.org/10.1093/jamia/ocac124

Liu, J. (2025). Exploring the impact of artificial intelligence-enhanced language learning on youths' intercultural communication competence. *Humanities and Social Sciences Communications*, 12(1), 1757. https://doi.org/10.1057/s41599-025-06033-x

Lunevich, L. (2022). Critical digital pedagogy: alternative ways of being and educating, connected knowledge and connective learning. *Creative Education*, 13(06), 1884–1896. https://doi.org/10.4236/ce.2022.136118

Mei, Z., Xue, K., Zhang, J. (2025). Intersectional Performativity Framework: A Collective Reflexivity in Ethics on Ethnographic Research With Vulnerable Groups in China. *Qualitative Inquiry*, https://doi.org/10.1177/10778004251393189 .

Mennella, C., Maniscalco, U., De Pietro, G., Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*, 10(4), e26297. https://doi.org/10.1016/j.heliyon.2024.e26297

Molla, M.A.M., Ahsan, M.M. (2025). Artificial Intelligence and Journalism: A Systematic Bibliometric and Thematic Analysis of Global Research. *Computers in Human Behavior Reports*, 20, 100830. https://doi.org/10.1016/j.chbr.2025.100830

Mukerji, R., Lin, Y.C., Zhang, S., Stone, M., Hushman, C., Moreu, F., Vigil, L., Eshelman, T., Rotche, L., Baca, A., Nodine, M., M. Faulkner., C. Johnson. (2024). Co-design as participation: Creating meaningful pathways for collaboration in flood risk adaptation in Ohkay Owingeh Pueblo. *International Journal of Disaster Risk Reduction*, 113, 104843. https://doi.org/10.1016/j.ijdrr.2024.104843

Muldoon, J., Wu, B.A. (2023). Artificial intelligence in the colonial matrix of power. *Philosophy & Technology*, 36(4), 80. https://doi.org/10.1007/s13347-023-00687-8

Nemorin, S. (2024). Towards decolonising the ethics of AI in education. *Globalisation, societies and education*, 1-13. https://doi.org/10.1080/14767724.2024.2333821

Obermeyer, Z., Mullainathan, S. (2019). Dissecting racial bias in an algorithm that guides health decisions for 70 million people. *In Proceedings of the conference on fairness, accountability, and transparency*, 89-89. https://dl.acm.org/doi/proceedings/10.1145/3287560

Ofosu-Asare, Y. (2025). Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies. *AI & society*, 40(4), 3045-3061. https://doi.org/10.1007/s00146-024-02065-0

Omodan, B.I. (2023). Unveiling epistemic injustice in education: A critical analysis of alternative approaches. *Social Sciences & Humanities Open*, 8(1), 100699. https://doi.org/10.1016/j.ssaho.2023.100699

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational researcher*, 41(3), 93-97. https://doi.org/10.3102/0013189x12441244

Pawar, S., Park, J., Jin, J., Arora, A., Myung, J., Yadav, S., Haznitrama, F.G., Song, I., Oh, A. Augenstein, I. (2025). Survey of Cultural Awareness in Language Models: Text and Beyond. *Computational Linguistics*, *51*(3), 907–1004. https://doi.org/10.1162/coli.a.14

Pelosi, D., Cacciagrano, D., Piangerelli, M. (2025). Explainability and Interpretability in Concept and Data Drift: A Systematic Literature review. *Algorithms*, 18(7), 443. https://doi.org/10.3390/a18070443

Perkowitz, S. (2021). The Bias in the Machine: Facial Recognition Technology and Racial Disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing, no. Winter 2021*. https://doi.org/10.21428/2c646de5.62272586.

Piu, P. (2023). The journey of subalternity in Gayatri Spivak's work: Its sociological relevance. *The Sociological Review*, *71*(6), 1258–1276. https://doi.org/10.1177/00380261231194495

Rapanta, C., Bhatt, I., Bozkurt, A., Chubb, L.A., Erb, C., Forsler, I., Gravett, K., Koole, M., Lintner, T., Örtegren, A., Petricini, T., Jandrić, P. (2025). Critical GenAI literacy: Postdigital configurations. *Postdigital Science and Education*, *7*(4), 1296–1333. https://doi.org/10.1007/s42438-025-00573-w

Romanishyn, A., Malytska, O., Goncharuk, V. (2025). AI-driven disinformation: policy recommendations for democratic resilience. *Frontiers in Artificial Intelligence*, *8*, 1569115. https://doi.org/10.3389/frai.2025.1569115

Roy, A., Horstmann, J., Ntoutsi, E. (2023). Multi-dimensional discrimination in law and machine learning-A comparative overview. *In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 89-100. https://doi.org/10.1145/3593013.3593979

Scalise, K., Wilson, M., Gochyyev, P. (2021). A taxonomy of critical dimensions at the intersection of learning analytics and educational measurement. *Frontiers in Education*, *6*, 656525. https://doi.org/10.3389/feduc.2021.656525

Singh, J., Mohanty, P.C. (2025). ICT-led socio-economic divide in India: A regression-based decomposition analysis. *Information Development*. https://doi.org/10.1177/02666669251376035

Smets, W. (2024). The purposes of historical canons in multicultural history education. *Journal of Curriculum Studies*, 56(3), 297–308. https://doi.org/10.1080/00220272.2024.2328050

Stewart, O.G., Rodgers, D.J. (2025). A critical AI media literacy framework: understanding layered bias and empowerment in artificial intelligence. *Learning Media and Technology*, 1–13. https://doi.org/10.1080/17439884.2025.2527179

Su, B., Huang, J., Miao, K., Wang, Z., Zhang, X., Chen, Y. (2023). K-Anonymity Privacy Protection Algorithm for Multi-Dimensional Data against Skewness and Similarity Attacks. *Sensors*, 23(3), 1554. https://doi.org/10.3390/s23031554

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2), 205395171773633. https://doi.org/10.1177/2053951717736335

Thong, C. (2022). South Asian digital humanities: Postcolonial mediations across technology's cultural Canon. *Journal of Postcolonial Writing*, 58(3), 434–435. https://doi.org/10.1080/17449855.2022.2083908

Thrift, N. (2006). Donna Haraway's Dreams. *Theory Culture & Society*, 23(7–8), 189–195. https://doi.org/10.1177/0263276406069231

Tweissi, A., Etaiwi, W.A., Eisawi, D.A. (2022). The accuracy of AI-Based automatic proctoring in online exams. *The Electronic Journal of e-Learning*, 20(4), 419–435. https://doi.org/10.34190/ejel.20.4.2600

Venugopal, K., Babu, V.P., Madhavi, K., Dooda, V.M. (2026). Impact of Digital Learning Platforms on Rural Education through Predictive Analytics and Regression. In K. Kagan, G. Schaippacasse Tapia (Eds.), *Adaptation and Challenges of Remote Teaching and Digital Tools.* IGI Global Scientific Publishing, 247-284. https://doi.org/10.4018/979-8-3373-3902-3.ch009

Vijayaraghavan, P., Vosoughi, S., Chiazor, L., Horesh, R., Abreu, D.P.R., Degan, E., Mukherjee, V. (2025). DECASTE: Unveiling Caste Stereotypes in Large Language Models through Multi-Dimensional Bias Analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.14971

Walle, Y.M. (2025). Breaking digital silence: AI, structural bias, and the uneven landscape of whistleblowing. AI & Society, 1-10. https://doi.org/10.1007/s00146-025-02758-0

Williamson, B. (2015). Digital education governance: An introduction. *European Educational Research Journal*, 15(1), 3–13. https://doi.org/10.1177/1474904115616630

Winter, J.S. (2018). Introduction to the special issue: Digital Inequalities and Discrimination in the Big Data era. *Journal of Information Policy*, *8*, 1–4. https://doi.org/10.5325/jinfopoli.8.2018.0001

Xia, M., Guo, S. (2025). Understanding learners' perceptions of artificial intelligence-mediated Informal Digital Learning of English: A Q methodology approach. *Acta Psychologica*, *261*, 105980. https://doi.org/10.1016/j.actpsy.2025.105980

Yang, Y. (2025). Racial bias in AI-generated images. *AI & SOCIETY*, 1-13. https://doi.org/10.1007/s00146-025-02282-1

Zajko, M. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass*, 16(3), e12962. https://doi.org/10.1111/soc4.12962

Zhang, L., Hu, D. (2024). National digital development strategy and its practice in India. *In Countries and Regions: Dynamic Interconnectivity*, 137-181. https://doi.org/10.1007/978-981-97-2835-0_6

Author Credit Statement

R. Suraj Begum: Conceptualization, Methodology, Formal analysis, Writing Original manuscript. S. Saravana Mahesan: Conceptualization, Methodology, Writing Original manuscript. Jiang Min: Conceptualization, Methodology, Writing Original manuscript. D. Divya: Conceptualization, Writing Review & Editing. Manuscript. Thennarasu Sakkan: Writing Review & Editing. All the authors read and approved the final version of the manuscript.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that ChatGPT-5.1 (OpenAI) was used solely for language polishing, grammar correction, and enhancing clarity and readability of the manuscript. All scientific content, experimental data, analyses, interpretations, and conclusions in this study are entirely original and generated by the authors.

Does this article screen for similarity?

Yes

Conflict of Interest

The authors have no conflicts of interest to declare. There is also no financial interest to report. The author certifies that the submission is original work and is not under review at any other publication.

About the License

Cite this Article

R. Suraj Begum, S. Saravana Mahesan, Jiang Min, D. Divya, Thennarasu Sakkan, Decolonizing the Digital Classroom: A Critical Analysis of Power, Privilege, and Algorithmic Bias in AI-Mediated Learning Environments, Asian Journal of Interdisciplinary Research, 8(4), (2025) 300-330. https://doi.org/10.54392/ajir25417