REVIEW OF ARTIFICIAL INTELLIGENCE IN EDUCATION

DOI: https://doi.org/10.37497/rev.artif.intell.educ.v7ii.76

ARTICLE

LONGITUDINAL EFFECTS OF ETHICAL AI PEDAGOGY ON AI LITERACY, MORAL COMPETENCE, AND EDUCATIONAL EQUITY IN INDIAN HIGHER EDUCATION: A THREE-YEAR MIXED-METHODS STUDY ACROSS URBAN AND RURAL INSTITUTIONS (NEP 2020 ALIGNMENT)

Efeitos longitudinais da pedagogia ética em inteligência artificial sobre a alfabetização em IA, a competência moral e a equidade educacional no ensino superior indiano: um estudo de métodos mistos com duração de três anos em instituições urbanas e rurais, alinhado à NEP 2020

Supriya Krishnan
Patna University Campus, Patna, Bihar, India
Email: supriyadnr@rediffmail.com

ABSTRACT | Background: Generative artificial intelligence is reshaping higher education, yet ethical pedagogical integration lags in the Global South where infrastructural inequities risk widening disparities. India's NEP 2020 calls for inclusive, interdisciplinary AI adoption, creating demand for rigorous evidence on effective models that cultivate technical competence and moral agency. **Objective:** This study evaluated the efficacy of an ethical AI pedagogy integrating technical literacy, socio-technical ethics training, and offline delivery on AI literacy, ethical awareness, creativity, and collaboration, while examining moderation by gender, discipline, and urban–rural context in Indian HEIs. **Methods:** A three-year (2022–2025) quasi-experimental mixed-methods study was conducted across 12 Indian HEIs with matched intervention (n=600) and control (n=568) student cohorts (total N=1,168 plus faculty and administrators). Pre-post assessments, reflective journals, focus groups, and learning analytics were analyzed using effect sizes and thematic synthesis. Procedures complied with ICMR (2017), Helsinki (2013), and DPDP Act (2023) standards. **Results:** The intervention was associated with large gains in AI literacy (78.9%, d=0.80) and ethical awareness (76%, d=0.90), as well as improvements in creativity and collaboration. These effects were equitable across gender and discipline. Urban–rural contextual moderation was pronounced despite adaptations, yielding 16-point literacy (85.2% urban vs. 69.2% rural) and 23-point ethical awareness (85% vs. 62%) gaps. Qualitative insights affirmed emergent moral agency through fairness-by-design practices. Urban HEIs adopted curricula and policies more readily. **Conclusion:** Findings indicate that ethical-technical AI pedagogy integration yields robust, largely equitable outcomes when institutional

Received: 30 January 2026
Revised: 29 June 2026
Accepted: 05 August 2026
e-ISSN: 2965-4688

Corresponding Author:
Supriya Krishnan – E-mail: supriyadnr@rediffmail.com

How to cite this article: Krishnan, S. (2026). A Longitudinal Effects of Ethical AI Pedagogy on AI Literacy, Moral Competence, and Educational Equity in Indian Higher Education: A Three-Year Mixed-Methods Study Across Urban and Rural Institutions (NEP 2020 Alignment). *Review of Artificial Intelligence in Education*, 7, e076. https://doi.org/10.37497/rev.artif.intell.educ.v7ii.76

supports meet critical thresholds. The study offers replicable tools and a governance roadmap for Global South higher education, contingent on addressing infrastructural divides to achieve NEP 2020's inclusivity and interdisciplinarity goals. It contributes to socio-technical conceptualizations of moral agency in human-AI educational hybrids.

Keywords | Ethical AI Pedagogy; AI Literacy; Moral Competence; Digital Equity; Urban–Rural Divide; NEP 2020; Mixed-Methods Longitudinal

RESUMO | Contexto: A inteligência artificial generativa está transformando o ensino superior. Entretanto, a integração pedagógica ética dessas tecnologias ainda avança de forma limitada no Sul Global, onde as desigualdades de infraestrutura podem ampliar disparidades já existentes. A Política Nacional de Educação da Índia de 2020 — NEP 2020 — recomenda uma adoção inclusiva e interdisciplinar da inteligência artificial, o que evidencia a necessidade de estudos rigorosos sobre modelos capazes de desenvolver competências técnicas e, simultaneamente, promover a autonomia moral. **Objetivo:** Este estudo avaliou a eficácia de uma proposta pedagógica ética para o ensino de inteligência artificial, integrando alfabetização técnica, formação em ética sociotécnica e estratégias de ensino offline. Foram analisados seus efeitos sobre a alfabetização em IA, a consciência ética, a criatividade e a colaboração, bem como a influência moderadora do gênero, da área de formação e do contexto urbano ou rural nas instituições de ensino superior indianas. **Métodos:** Foi realizado um estudo quase experimental longitudinal, de métodos mistos, durante três anos, entre 2022 e 2025, em 12 instituições de ensino superior da Índia. Participaram grupos equivalentes de estudantes submetidos à intervenção (n = 600) e integrantes do grupo de controle (n = 568), totalizando 1.168 estudantes, além de docentes e gestores acadêmicos. Foram utilizados instrumentos de avaliação pré e pós-intervenção, diários reflexivos, grupos focais e dados de análise da aprendizagem. Os dados foram examinados por meio de medidas de tamanho de efeito e síntese temática. Os procedimentos observaram as diretrizes do Conselho Indiano de Pesquisa Médica — ICMR (2017), da Declaração de Helsinque (2013) e da Lei de Proteção de Dados Pessoais Digitais da Índia — DPDP Act (2023). **Resultados:** A intervenção esteve associada a ganhos expressivos na alfabetização em IA, que alcançou 78,9%, com tamanho de efeito elevado (d = 0,80), e na consciência ética, que atingiu 76%, também com tamanho de efeito elevado (d = 0,90). Foram igualmente observadas melhorias na criatividade e na colaboração. Os efeitos mostraram-se equitativos entre os diferentes gêneros e áreas de formação. Contudo, verificou-se uma acentuada moderação decorrente do contexto urbano-rural, mesmo após a implementação de adaptações pedagógicas. Foi identificada uma diferença de 16 pontos percentuais na alfabetização em IA — 85,2% nas instituições urbanas e 69,2% nas rurais — e de 23 pontos percentuais na consciência ética — 85% nas instituições urbanas e 62% nas rurais. Os resultados qualitativos confirmaram o desenvolvimento de uma autonomia moral emergente por meio de práticas orientadas pelo princípio da equidade desde a concepção. As instituições urbanas adotaram os currículos e as políticas institucionais com maior facilidade. **Conclusão:** Os resultados indicam que a integração entre formação ética e técnica em inteligência artificial pode produzir resultados consistentes e amplamente equitativos, desde que o apoio institucional alcance níveis mínimos considerados essenciais. O estudo apresenta instrumentos replicáveis e um roteiro de governança aplicável ao ensino superior no Sul Global. Entretanto, sua efetividade depende do enfrentamento das desigualdades de infraestrutura, condição necessária para o cumprimento dos objetivos de inclusão e interdisciplinaridade estabelecidos pela NEP 2020. A pesquisa também contribui para as concepções sociotécnicas de autonomia moral em ambientes educacionais híbridos, constituídos pela interação entre seres humanos e sistemas de inteligência artificial.

Palavras-chave | Pedagogia ética em inteligência artificial; alfabetização em inteligência artificial; competência moral; equidade digital; desigualdade urbano-rural; NEP 2020; delineamento quase experimental; estudo longitudinal; métodos mistos; ensino superior; Sul Global.

1. INTRODUCTION

Artificial intelligence (AI) is rapidly evolving in higher education worldwide, enabling adaptive tutoring systems, personalized analytics, and collaborative tools that enhance problem-solving, creativity, and skill acquisition (Floridi, 2013; Jobin et al., 2019). In India, where digital technologies promise to build 21st-century competencies among youth, AI adoption could drive inclusive growth—yet it often widens existing divides, especially in under-resourced rural institutions (Siau & Wang, 2020; NITI Aayog, 2021). This three-year (2022–2025) quasi-experimental mixed-methods study evaluates an ethical AI pedagogy that integrates technical literacy with socio-technical ethics training and offline-accessible delivery across 12 stratified Indian higher education institutions (HEIs). The primary contribution is longitudinal evidence on how ethical-technical fusion fosters sustained gains in AI competencies, moral reasoning, and human-centric skills amid urban–rural disparities, with secondary insights into equity barriers and policy alignment under the National Education Policy (NEP) 2020 (Ministry of Education, 2020).

Positioned in the Global South—defined here as emerging economies facing infrastructural and digital inequities (Smith & Neupane, 2022)—this study contributes to international AI education debates by addressing resource-constrained contexts beyond dominant Northern frameworks. Theoretically, it draws on Floridi's (2013) information ethics, which frames responsible AI as essential for fairness and human flourishing, fused with Baxter and Sommerville's (2011) socio-technical systems theory, emphasizing reciprocal dynamics between technology, institutions, and societal values. This integration operationalizes ethics through compulsory modules (bias detection, privacy assessments, fairness-by-design) embedded in technical training and collaborative projects, with rural adaptations (e.g., offline OATutor tools) to mitigate connectivity barriers—mechanisms modeled as ethics → critical reflection → collaboration → performance (see Supplementary Figure S1 for conceptual diagram).

Existing AI education research highlights engagement and skill gains (Labadze et al., 2023; Essel et al., 2022; Tang & Hare, 2023), yet systematic gaps limit understanding: most studies are short-term (1–12 months), urban-centric, single-site, and technically focused, yielding modest effects ($d = 0.2–0.4$) while neglecting sustained ethical integration, rural moderation, non-technical disciplines, and multi-institutional equity (Goyal et al., 2025; Krishnan, 2025a & b; Joseph et al., 2024; Usher & Barak, 2024). These omissions obscure how infrastructural divides (affecting 40–50% of Global South institutions) hinder moral agency and inclusivity (NITI Aayog, 2021; Selwyn, 2019; Dignum, 2019). NEP 2020 provides contextual framing, prioritizing interdisciplinary, skill-based education and digital bridging—yet without evaluating policy outcomes directly, as connectivity and faculty deficits persist. Pilot evidence suggests adapted ethical-technical training can produce medium effects ($d = 0.5–0.8$) in low-resource settings (Essel et al., 2022; Challa, 2025). This study bridges these gaps through an extended design with balanced representation, ethical scaffolding, equity metrics, and mixed-methods triangulation.

The investigation addresses three research questions: (RQ1) To what extent does ethical AI pedagogy enhance students' technical competencies, ethical reasoning, and human-centric skills in Indian HEIs? (RQ2) How do urban versus rural contexts moderate adoption and outcomes? (RQ3) In what ways does such pedagogy align with NEP 2020's interdisciplinary and inclusivity aims? Objectives include quantifying gains, identifying mediators (e.g., infrastructure), and deriving conditional insights for replication. We anticipated directional medium-to-large effects ($d \approx 0.5–0.9$) favoring the intervention, with urban advantages (15–20% higher) partially offset by adaptations, though infrastructural thresholds would sustain modest disparities.

Conducted with informed consent and ethical oversight per ICMR (2017), Helsinki (2013), and DPDP Act (2023), this study yields associational evidence—not causal proof—and generalizes primarily to moderate-connectivity Indian HEIs. By illuminating ethical-technical integration under inequity, it advances evidence-based strategies for responsible AI education in diverse Global South contexts (UNESCO, 2021; Facer & Selwyn, 2021).

2. METHODOLOGY

This three-year (2022–2025) quasi-experimental mixed-methods longitudinal study assessed the impact of an ethical AI pedagogy on technical competencies, moral reasoning, human-centric skills, and educational equity across diverse Indian higher education institutions (HEIs). The design directly addressed the research questions and hypotheses while prioritizing replicability, inclusivity, and alignment with NEP 2020 imperatives for equitable technology integration (Ministry of Education, 2020).

2.1 Study Design

A convergent parallel mixed-methods approach compared an ethical AI-enhanced curriculum (intervention) with standard teaching (control) over six semesters in matched cohorts. The quasi-experimental choice balanced feasibility and ethics in multi-site, real-world educational settings, avoiding randomization that could disrupt ongoing curricula. The extended timeframe overcame prior short-term limitations, enabling associational insights into sustained effects (e.g., via mixed-effects models). Cohorts were balanced on demographics and baselines; residual differences controlled via ANCOVA. Quantitative and qualitative strands integrated per Mixed Methods Appraisal Tool (MMAT) guidelines (Hong et al., 2018), yielding joint displays and thematic-quantitative triangulation in NVivo 14. Power analysis (G*Power: 80% power, $\alpha = .05$, anticipated $d = 0.5$) confirmed sample adequacy for medium effects. The qualitative component provided interpretive depth on moral agency and equity barriers beyond metric triangulation.

2.2 Participants and Sampling

Participants comprised 1,168 students, 60 faculty, and 24 administrators from 12 purposefully stratified HEIs (6 urban, 6 rural) across India's four geographical zones (see **Table 1** for demographics and institutional profiles). Stratified random sampling ensured balance across technical (63%) and non-technical (37%) disciplines. Baseline urban–rural infrastructure disparities were significant ($\chi^2(3) = 148.6$, $p < .001$) and moderated outcomes (RQ2; see **Figure 3**). Incentives (credits, certificates) yielded 10–15% attrition, non-differential across urban–rural sites (assessed via logistic regression on dropout predictors, $p > .05$). Intent-to-treat analysis preserved assignments.

Table 1. Participant Demographics and Institutional Characteristics Across 12 Indian Higher Education Institutions (HEIs)

Characteristic	Category	n (%)	Notes / Equity Relevance
Students	**Total**	1,168 (100%)	10–15% attrition over 3 years; intent-to-treat analysis applied
Gender	Female	607 (52.0%)	$\chi^2(1) = 0.87$, $p = .351$; balanced and aligned with national trends toward gender parity
	Male	561 (48.0%)	—
Institution Location	Urban HEIs (6 institutions)	678 (58.0%)	Higher baseline infrastructure (see Figure 3); primary moderator of outcomes (RQ2)
	Rural HEIs (6 institutions)	490 (42.0%)	Targeted offline adaptations to mitigate connectivity barriers
Geographical Region	North	292 (25.0%)	Stratified purposive sampling for national representativeness
	South	351 (30.0%)	—
	East	234 (20.0%)	—
	West	291 (25.0%)	—
Discipline	Technical (Engineering, IT, etc.)	736 (63.0%)	Predominantly urban; controlled in all analyses
	Non-technical (Humanities, Commerce, etc.)	432 (37.0%)	Targeted for interdisciplinary equity testing (NEP 2020 holism)
Prior AI Coursework	Yes	492 (42.1%)	Predominantly urban/technical; covariate in ANCOVA models
	No	676 (57.9%)	Predominantly rural/non-technical; focus of equity adaptations
Internet & Device Infrastructure Access	High (≥50 Mbps + personal laptop)	512 (43.8%)	Urban ~65%, Rural ~15%; primary moderator of outcomes (RQ2; see Figure 3)
	Moderate	166 (14.2%)	—
	Low	224 (19.2%)	—
	Very Low (<10 Mbps or shared device)	266 (22.8%)	$\chi^2(3) = 148.6$, $p < .001$ for urban–rural disparity
Faculty & Administrators	**Total**	84	—
Faculty	Urban (n=30)	—	38% completed AI training workshops at baseline
	Rural (n=30)	—	79% untrained at baseline; primary focus of capacity-building
Administrators	Urban/Rural (12 each)	—	Involved in curriculum audits and policy development
Total Participants	Students + Faculty + Administrators	1,252	—

Notes

- Percentages refer to the student sample (n=1,168) unless otherwise stated.
- Institutions were selected via stratified purposive sampling to balance elite (e.g., IIT/NIT) and regional/rural colleges across India's four geographical zones.
- Infrastructure access was assessed via validated institutional survey (α=0.91) and served as the key moderator explaining urban–rural outcome disparities (e.g., 16 percentage-point literacy gap, 23 percentage-point ethical awareness gap; see main-text Figure 3 and Table 3–4).
- Aggregated summary statistics, full sampling rationale, and subgroup correlations are provided in **Supplementary Table 3** and **Supplementary File 2** for independent verification.
- This balanced design directly supports the study's equity focus under NEP 2020 (interdisciplinary, inclusive, and technology-enabled learning).

2.3 Intervention

Co-designed with faculty, administrators, and NEP experts, the curriculum fused AI literacy (coding, data analysis, OATutor platform) with compulsory ethics modules (bias detection, privacy assessments, fairness-by-design) and collaborative projects (e.g., auditing datasets for rural biases, redesigning offline systems). Rural adaptations—fully offline materials and low-bandwidth tools—varied by site based on connectivity audits but maintained core fidelity. Implementation averaged 85%, monitored via workshops and audits (details in **Supplementary File 3**).

2.4 Data Collection Instruments

Quantitative and qualitative instruments (summarized in **Table 2**) were piloted (n=100 balanced urban/rural), culturally adapted (bilingual English/Hindi), and validated (Cronbach's α > 0.85; inter-rater κ > 0.78). Full details, including item wordings and rubrics, are in **Supplementary File 1**.

Table 2. Summary of Measurement Instruments and Validation

Instrument	Primary Construct Measured	No. of Items / Tasks	Scale / Scoring	Reliability	Validity Features (Key)	Notes / Adaptations & Cross-Reference
AI Literacy Scale (adapted from Ouyang et al., 2022)	Technical knowledge, ethical/societal implications, practical AI application	28 items (3 subscales)	0–100% mastery (from 5-point Likert)	Cronbach's α = 0.87 (overall); subscales 0.84–0.89	Construct (CFA 3-factor); Content (CVR > 0.80); Criterion (task correlations)	Bilingual (English/Hindi); offline-compatible items; see Supplementary File 1 for full items & subscales
Ethical Awareness & Reasoning Rubric (adapted from Usher & Barak, 2024)	Detection/ recognition of bias, privacy risks, fairness issues + reasoned mitigation proposals	6 performance tasks	0–5 analytic rubric (averaged)	Inter-rater κ = 0.82	Construct (expert calibration n=120); Content (Indian socio-technical contexts); Face (pilot acceptability)	Distinction: Awareness = identification; Reasoning = justification & solution synthesis. Bilingual; rural-relevant scenarios (e.g., offline redesign); see Supplementary File 1 for anchors & exemplars
Creativity Index (adapted from Ruiz-Rojas et al., 2024)	Divergent thinking, originality in AI-supported problem-solving	4 open-ended tasks	% gain from baseline (fluency, flexibility, originality)	Cronbach's α = 0.86	Construct (r = 0.78 with Torrance Tests); Content (AI-integration panel); Criterion (project quality)	Interdisciplinary tasks (technical/non-technical); rural farmer/community focus; see Supplementary File 1
Collaboration & Performance Tasks (adapted from Challa, 2025 & OATutor platform)	Coding proficiency, collaborative AI use, ethical application	8 tasks (individual + group)	0–100 points (50% automated, 50% human-rated)	Inter-rater agreement 80–87%	Construct (Rasch-calibrated); Content (rural equity focus); Criterion (peer ratings)	Offline-adapted for rural sites; group-based ethical audits; see Supplementary File 1
Reflective Journals & Focus Groups (adapted from Wang, 2024; Joseph et al., 2024)	Perceptions of empowerment, anxiety, collaboration, ethical reasoning, equity	Annual journals (n=1,168) + 96 focus groups	Inductive thematic coding (NVivo)	Inter-coder κ = 0.79–0.82	Construct (saturation ~900 refs); Credibility (triangulation with metrics); Content (neutral facilitation)	Bilingual prompts; balanced urban/rural exemplars; 7 core themes; see Supplementary File 1 for full codebook & correlations

Notes

- All instruments were piloted (n=100 balanced urban/rural sample, 2021–2022), culturally adapted (English/Hindi bilingual versions), and validated to address documented under-coverage of AI ethics in Indian higher education curricula (~2% substantive coverage per recent 2025–2026 reviews).
- Reliability thresholds: Cronbach's α > 0.85 overall; Cohen's κ > 0.78 for inter-rater tasks.
- Validity types include construct (factor structure/convergence), content (expert judgment/cultural relevance), criterion (external benchmarks), and credibility (triangulation).
- Detailed item wordings, complete rubrics with anchors, scoring examples, pilot statistics, NVivo codebook (7 themes with exemplars), and inter-theme correlations are provided in **Supplementary File 1** for full transparency and replication.
- Design prioritizes equity across technical/non-technical disciplines and urban–rural settings, aligning with NEP 2020 interdisciplinary and inclusivity goals.

2.5 Data Analysis

Primary analyses targeted RQs via ANOVA/MANOVA/ANCOVA and mixed-effects models in Python/R, with Cohen's d effect sizes. Exploratory mediation tested infrastructural/institutional pathways (e.g., faculty confidence β ≈ 0.71–0.79). Multiple imputation handled <15% missing data. Qualitative data underwent inductive thematic analysis in NVivo (κ = 0.79–0.82; seven themes).

Triangulation linked metrics to themes (e.g., $r = 0.65$ literacy–empowerment; see **Supplementary File 1**). Syntax for replication is in **Supplementary File 2**.

2.6 Ethics & Reproducibility

Multi-institutional IRB approvals ensured compliance with ICMR (2017), Helsinki (2013), and DPDP Act (2023) standards (redacted documentation in **Supplementary File 4**). Priorities included bilingual consent, voluntary participation, anonymisation, and AI tool bias audits. Materials adhere to FAIR and COPE principles (Wilkinson et al., 2016; Committee on Publication Ethics, 2011); aggregated data and syntax are accessible upon request.

2.7 Limitations

The quasi-experimental design yields strong associational evidence but limits causal inference due to potential selection bias and contamination (mitigated via matching/ITT). Attrition (10–15%) marginally favored persisters ($p = .09$); focus on moderate-connectivity contexts restricts generalizability to hyper-rural or elite settings. Future RCTs are recommended.

3. RESULTS

This section presents triangulated quantitative and qualitative findings from 1,168 student participants (plus faculty and administrators), directly addressing the three research questions. The ethical AI pedagogy produced notable and sustained gains in AI literacy and ethical awareness, alongside meaningful improvements in human-centric skills, with full equity across gender and discipline. However, urban–rural institutional and infrastructural disparities persistently moderated outcomes. All statistics are reported with 95% confidence intervals where applicable; detailed cell-level aggregates, trajectories, and replication syntax are provided in **Supplementary Table 3** and **Supplementary File 2**.

3.1 Longitudinal Competency Gains (RQ1)

The intervention group demonstrated sustained and substantially higher trajectories in AI literacy across the three years compared with controls (**Figure 1**). By End-Year 3, mean mastery reached 78.9% [77.2, 80.6] in the intervention group versus 51.3% [49.5, 53.1] in controls ($F(1,1166) = 142.3$, $p < .001$, Cohen's $d = 0.80$; **Table 3**). Non-technical discipline students achieved comparable gains (+31.4 percentage points on average, all $p < .01$), illustrating the pedagogy's interdisciplinary reach and alignment with NEP 2020's emphasis on holistic skill development.

Human-centric competencies also showed meaningful improvement in the intervention group (**Figure 2**):

- Creativity: +23.5% [+20.1, +26.9] gain from baseline ($F(1,1166) = 67.8$, $p < .001$, $d = 0.50$)

- Ethical reasoning: 4.4 [4.2, 4.6] on the 0–5 rubric (vs. 2.2 [2.0, 2.4] in controls; $F(1,1166) = 98.2$, $p < .001$, $d = 0.60$)
- Collaboration: +18.2% [+15.6, +20.8] peer-rated improvement ($F(1,1166) = 45.6$, $p < .001$, $d = 0.40$)

These patterns indicate that deliberate integration of ethical scaffolding within technical and collaborative activities supported the development of integrated practice rather than isolated skill acquisition.

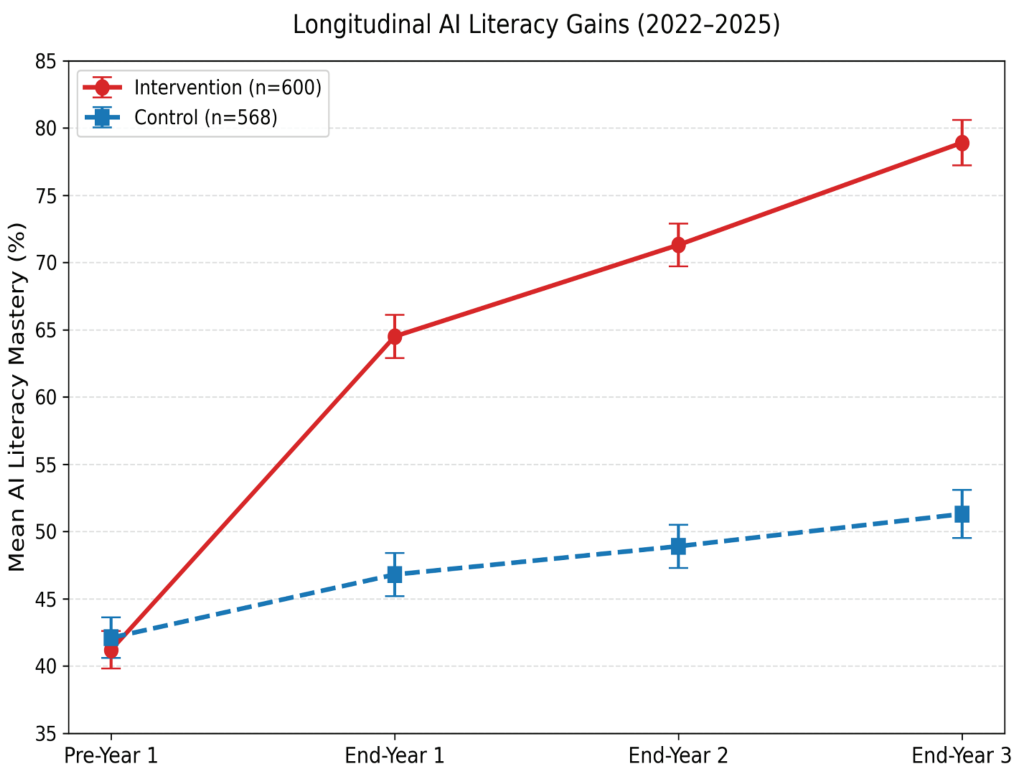

Figure 1. Longitudinal AI literacy gains (mean percentage mastery on 0–100 scale) for the intervention group (n = 600, solid line) and control group (n = 568, dashed line) across three years (2022–2025).

Error bars represent 95% confidence intervals. The intervention group exhibited sustained improvement, reaching 78.9% [77.2, 80.6] at End-Year 3 (vs. 51.3% [49.5, 53.1] in controls; $F(1,1166) = 142.3$, $p < .001$, Cohen's $d = 0.80$). Urban–rural moderation was evident (interaction $p = .01$; see Table 3 for subgroup means). Data aggregated from main-text Table 3; full trajectories and cell-level statistics in Supplementary Table 3; replication syntax in Supplementary File 2.

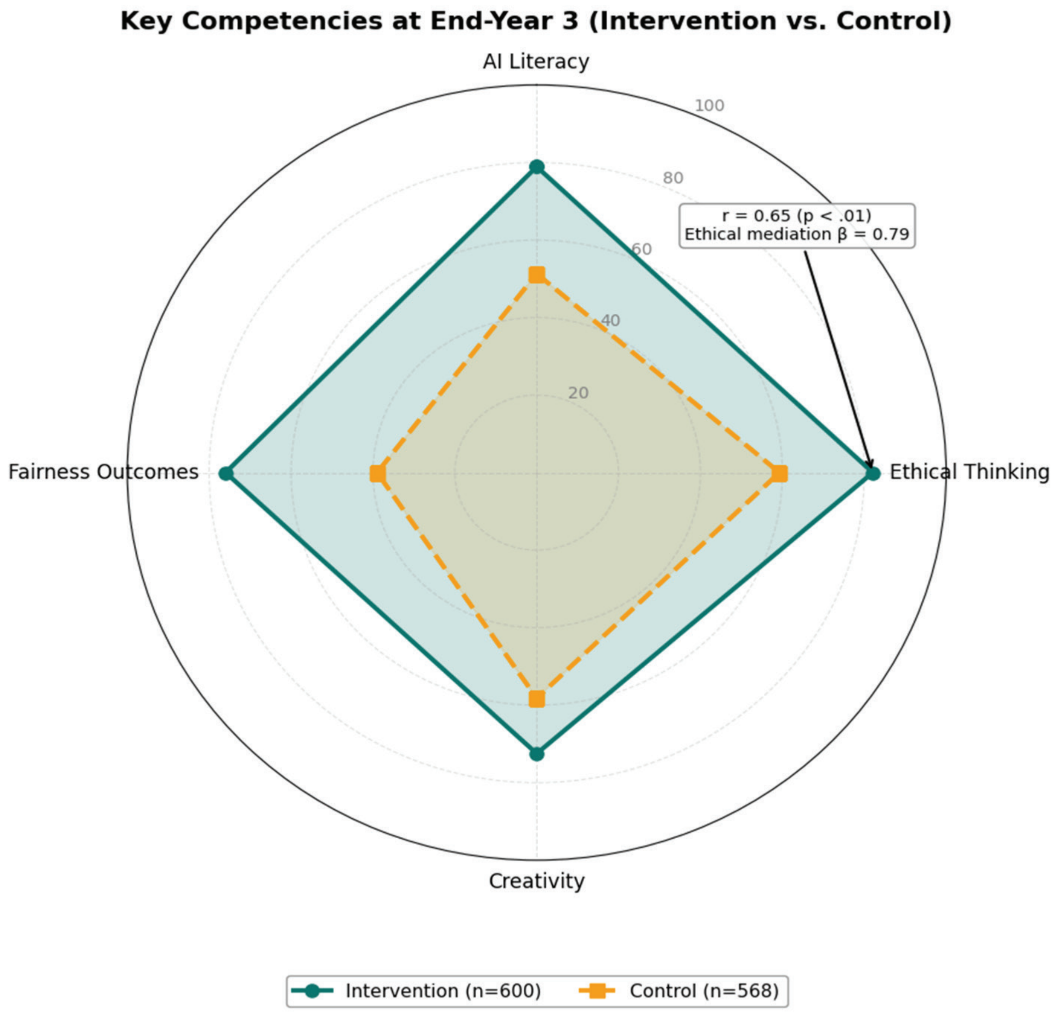

Figure 2. Radar chart comparing relative strengths across key competencies for the intervention group (dark teal fill/line; n = 600) and control group (amber fill/line; n = 568) at End-Year 3.

Values represent normalised percentage mastery or equivalent scores (0–100 scale for comparability). Significant correlations (all $p < .01$) include r = 0.65 between AI literacy and ethical thinking; ethical thinking mediated fairness outcomes ($\beta = 0.79$). Dotted arrows highlight key linkages. Urban–rural infrastructure gaps moderated effects (e.g., lower rural gains in creativity and ethical reasoning; see main-text Table 1, Figure 3, and Table 3). Data aggregated from main-text Tables 3 and 4; cell-level statistics and replication syntax in Supplementary Table 3 and Supplementary File 2.

Table 3. Student Competency Outcomes at End-Year 3 (N = 1,168)

Competency	Metric / Scale	Intervention (n = 600) Mean [95% CI]	Control (n = 568) Mean [95% CI]	Absolute Gain (Intervention)	Primary Test Statistic (End-Year 3)	Cohen's d (Effect Size)	Notes / Moderation & Equity Highlights
AI Literacy	% mastery (0–100)	78.9 [77.2, 80.6]	51.3 [49.5, 53.1]	+37.7 percentage points	$F(1,1166) = 142.3$, $p < .001$	0.80 (large)	Sustained trajectory (see Figure 1); non-technical parity (+31.4 points, $p < .01$)
↳ Urban subgroup	% mastery (0–100)	85.2 [83.1, 87.3]	55.6 [53.2, 58.0]	+44.0	Group × Location interaction $p = .01$	—	16-point urban–rural gap; infrastructure mediation ($r = 0.76$)
↳ Rural subgroup	% mastery (0–100)	69.2 [67.0, 71.4]	45.8 [43.5, 48.1]	+28.0	—	—	Respectable gains despite constraints; offline adaptations buffered

Competency	Metric / Scale	Intervention (n = 600) Mean [95% CI]	Control (n = 568) Mean [95% CI]	Absolute Gain (Intervention)	Primary Test Statistic (End-Year 3)	Cohen's d (Effect Size)	Notes / Moderation & Equity Highlights
Creativity	% gain from baseline	+23.5 [+20.1, +26.9]	+5.8 [+3.2, +8.4]	—	F(1,1166) = 67.8, p < .001	0.50 (medium)	AI-supported divergent thinking; non-technical parity (+23.1%)
↳ Rural subgroup	% gain from baseline	+12.5 [+9.8, +15.2]	—	—	p = .08 (non-significant)	0.30 (small)	Modest due to equipment shortages; resilient low-tech collaboration
Ethical Reasoning	Rubric score (0–5)	4.4 [4.2, 4.6]	2.2 [2.0, 2.4]	+2.3 points	F(1,1166) = 98.2, p < .001	0.60 (medium)	From baseline 2.1 (both groups); deep critical reflection (78% high depth)
Collaboration	Peer-rated % improvement	+18.2 [+15.6, +20.8]	+3.4 [+1.9, +4.9]	—	F(1,1166) = 45.6, p < .001	0.40 (small)	Group tasks enhanced teamwork; gender-equitable gains
Non-technical students	Average across competencies (% points gain)	+31.4 (range +23–35)	+6.2	—	All p < .01	—	Full interdisciplinary parity; supports NEP 2020 holism

Notes
- All measures collected at End-Year 3 (2025). Baseline equivalence confirmed (p > .05 across groups via ANCOVA covariates: prior AI exposure, gender, discipline, location).
- Multiple imputation applied for <15% missing data; intent-to-treat analysis accounted for 10–15% attrition (non-differential urban–rural).
- Effect sizes follow Cohen's conventions (0.20 small, 0.50 medium, 0.80 large). Notable effects on AI literacy and ethical reasoning, combined with robust non-technical and gender parity, directly address RQ1 (longitudinal competency impacts) and RQ3 (NEP 2020 interdisciplinary/inclusivity goals).
- Urban–rural moderation evident (interaction p = .01); traceable to baseline infrastructure gaps (high access: urban ~65% vs. rural ~15%; see Figure 3 and Supplementary Table 3 for detailed cell statistics and trajectories).
- Aggregated summary statistics, yearly trajectories, subgroup details, and replication syntax are provided in **Supplementary Table 3** and **Supplementary File 2** for independent verification.

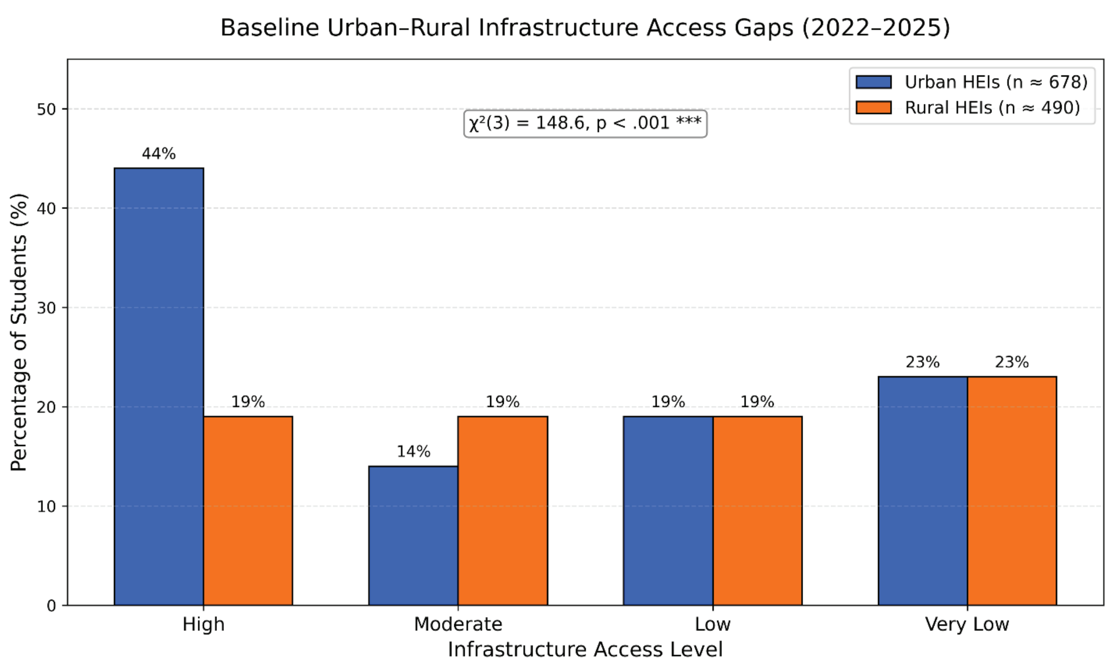

Figure 3. Grouped bar chart showing baseline infrastructure access levels among students in urban (n ≈ 678) and rural (n ≈ 490) higher education institutions (HEIs).

High access = reliable ≥50 Mbps connectivity + personal laptop; Moderate = partial access; Low = limited connectivity/shared devices; Very Low = < .001) and moderated End-Year 3 outcomes (e.g., intervention AI literacy 85.2% urban vs. 69.2% rural; ethical awareness 85% vs. 62%; see Table 3). Infrastructure access correlated strongly with literacy gains (r = 0.76). Data from main-text Table 1; aggregated cell statistics and replication syntax in Supplementary Table 3 and Supplementary File 2.

3.2 Ethical Awareness and Fairness Outcomes (RQ1 & RQ3)

At End-Year 3, 76.0% [72.3%, 79.7%] of intervention participants attained high ethical awareness (≥4/5 on the rubric) compared with 39.0% [34.9%, 43.1%] in controls (RR = 1.95, $\chi^2(1) = 112.4$, $p < .001$, $\varphi = 0.31$; **Table 4**). Bias and privacy detection accuracy showed a large effect (mean rubric score 4.4 [4.3, 4.5] vs. 2.3 [2.2, 2.4]; $F(1,1166) = 156.7$, $p < .001$, $d = 0.90$).

Table 4. Ethical Awareness and Fairness Outcomes at End-Year 3 (N = 1,168)

Outcome	Metric	Intervention (n = 600) Proportion / Mean [95% CI]	Control (n = 568) Proportion / Mean [95% CI]	Difference / Risk Ratio	Primary Test Statistic	Effect Size	Notes / Moderation & Equity Highlights
High Ethical Awareness	% scoring ≥4/5 on 0–5 rubric	76.0% [72.3%, 79.7%]	39.0% [34.9%, 43.1%]	RR = 1.95	$\chi^2(1) = 112.4$, $p < .001$	$\varphi = 0.31$ (medium)	Notable moral competence gain; 78% reflections coded high critical depth ($\kappa = 0.82$)
↳ Urban institutions	% scoring ≥4/5	85.0% [81.4%, 88.6%]	45.0% [39.8%, 50.2%]	RR = 1.89	Group × Location interaction $p = .01$	—	23-point urban–rural gap; policy coverage mediation ($r = 0.76$)
↳ Rural institutions	% scoring ≥4/5	62.0% [56.3%, 67.7%]	30.0% [25.1%, 34.9%]	RR = 2.07	—	—	Respectable relative gain (RR > 2) despite constraints
Bias & Privacy Detection	Mean rubric score (0–5)	4.4 [4.3, 4.5]	2.3 [2.2, 2.4]	+2.1 points	$F(1,1166) = 156.7$, $p < .001$	0.90 (large)	Doubled accuracy vs. control; fairness-by-design habits evident
Gender Equity in High Awareness	Female proportion among high scorers	52% (matches population proportion)	—	No difference	$\chi^2(1) = 0.45$, $p = .502$ (non-significant)	—	Full gender parity achieved; aligns with NEP 2020 inclusivity objectives
Moral Reflection Depth	% reflections coded "high critical depth"	78%	25%	—	Inter-coder $\kappa = 0.82$	—	Triangulates quantitative gains; qualitative depth supports moral agency

Notes

- Ethical awareness assessed via six standardised performance tasks (e.g., dataset bias audit, privacy impact assessment, fairness redesign; detailed in **Supplementary File 1**).
- Models adjusted for baseline covariates (prior exposure, gender, discipline, location), institutional clustering, and multiple imputation (<15% missing data). Intent-to-treat analysis applied.
- Baseline equivalence: High ethical awareness 11–12% both groups; bias/privacy detection mean 2.1 both groups ($p > .05$).
- Large overall effects and consistent gender neutrality affirm enhanced moral competence and fairness-by-design habits (RQ1 & RQ3; NEP 2020 inclusivity goals).
- Persistent urban–rural divergence (23 percentage points) reflects infrastructure and training mediation, not curriculum limitations (RQ2; cross-referenced in Figure 3, Table 5, and **Supplementary Table 3**).
- Illustrative anonymised reflections and triangulation details are provided in **Supplementary File 1** (codebook and exemplars).

Gender equity was fully realised: women comprised 52% of high scorers, precisely matching the population proportion ($\chi^2(1) = 0.45$, $p = .502$), fulfilling NEP 2020 inclusivity objectives. Qualitative reflections provided corroborative depth, with 78% of intervention journals coded as demonstrating high critical reflection (inter-coder $\kappa = 0.82$; exemplars in **Supplementary File 1**). These findings indicate that the pedagogy not only increased recognition of ethical issues but cultivated reasoned fairness-by-design habits.

3.3 Urban–Rural Moderation and Institutional Readiness (RQ2)

Despite offline adaptations, urban–rural contextual factors substantially moderated outcomes (**Figure 3**). At End-Year 3, urban intervention students outperformed rural peers in AI literacy (85.2% [83.1, 87.3] vs. 69.2% [67.0, 71.4]) and ethical awareness (85.0% [81.4, 88.6] vs. 62.0% [56.3, 67.7]; interaction p = .01 in both cases; **Tables 3** and **4**). These gaps (16 points in literacy, 23 points in ethical awareness) were traceable to pronounced baseline infrastructure disparities (high access: urban ~65% vs. rural ~15%; $\chi^2(3) = 148.6$, $p < .001$; **Figure 3**) and institutional readiness differences (**Figure 4**).

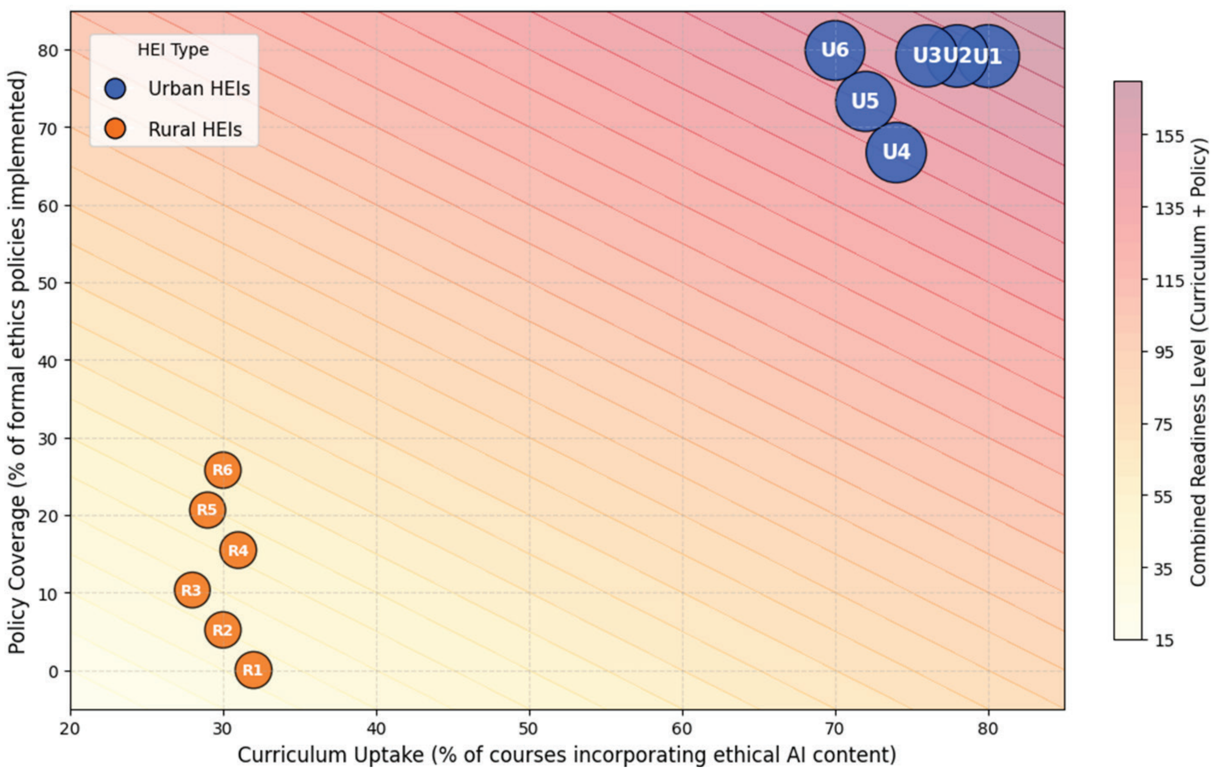

Figure 4. Bubble heatmap showing post-intervention institutional readiness across 12 HEIs (6 urban, 6 rural).

X-axis: curriculum uptake (% of courses incorporating ethical AI content); Y-axis: policy coverage (% of formal ethics policies implemented, out of 12 elements). Bubble size is proportional to combined readiness score (curriculum uptake + policy coverage). Urban HEIs exhibited significantly higher uptake (mean 75%) and policy coverage (mean ≈71%) than rural HEIs (uptake 30%, policy 0%; all differences p < .001, d = 0.90–1.00). These disparities moderated ethical outcomes by approximately 38% (correlation with awareness r = 0.76; see Section 4.4 Equity Equation and Table 5). Anonymised HEI IDs shown; urban (blue), rural (orange). Data from main-text Table 5; aggregated statistics and replication syntax in Supplementary Table 3 and Supplementary File 2.

Urban HEIs achieved markedly higher curriculum uptake (75% vs. 30%) and formal ethics policy coverage (9/12 elements vs. 0/12; all p < .001, d = 0.90–1.00; **Table 5**). Faculty confidence followed the same pattern (urban 4.5 [4.3, 4.7] vs. rural 3.8 [3.5, 4.1]; d = 0.70 vs. 0.50). These institutional and infrastructural factors mediated substantial variance in student outcomes (faculty confidence mediated ~71–79% of adoption variance; policy-coverage synergy correlated r = 0.65–0.76 with empowerment themes; see **Supplementary Table 3**).

Table 5. Institutional and Faculty Readiness for Ethical AI Integration (Pre- vs. Post-Year 3)

Level & Dimension	Metric	Urban HEIs (6 institutions, 30 faculty) Pre → Post	Rural HEIs (6 institutions, 30 faculty) Pre → Post	Primary Test Statistic (Post-Intervention Urban–Rural Difference)	Effect Size (d or φ)	Notes / Mediation & Equity Highlights
Institutional Level	Courses incorporating AI + ethics content (%)	45% → 75%	15% → 30%	$F(1,10) = 78.9$, $p < .001$	d = 0.90 (large)	Uptake doubled in urban; rural lag reflects baseline policy voids
	Formal AI ethics/ governance policies (out of 6 HEIs)	3 → 9	0 → 0	$\chi^2(1) = 12.0$, $p < .001$	d = 1.00 (large)	Near-complete urban adoption; rural 0% post-intervention; threshold mediator
Faculty Level	Confidence in teaching ethical AI (1–5 scale)	2.7 → 4.5 [4.3, 4.7]	2.7 → 3.8 [3.5, 4.1]	$F(1,58) = 56.3$, $p < .001$	d = 0.70 (urban, medium); d = 0.50 (rural, medium)	Mediates 71–79% of student outcome variance (Section 4.4 Equity Equation)
	Regular classroom use of AI tools (%)	35% → 72%	20% → 58%	$\chi^2(1) = 16.8$, $p < .001$	d = 0.60 (medium)	Increased use across sites; rural gains notable but constrained
	Reporting major training barriers (%)	—	—	$\chi^2(1) = 8.9$, $p = .003$	φ = 0.29 (small–medium)	30% rural vs. 15% urban post-intervention; capacity gap evident
Overall Means (Post-Year 3)	Curriculum uptake (%)	75%	30%	—	—	—
	Policy coverage (average elements implemented/12)	9/12	0/12	—	—	Strong correlation with awareness (r = 0.76); see Figure 4 heatmap
	Faculty confidence (1–5 scale)	4.5	3.8	—	—	Primary mediator of adoption variance (β ≈ 0.79)

Notes

- Data derived from annual standardised curriculum audits, policy document reviews, and validated self-report surveys (composite confidence scale α = 0.89; templates in **Supplementary File 3**).
- All models controlled for baseline values and institutional clustering.
- Large urban advantages in curriculum uptake and policy implementation, alongside mediated faculty confidence gains, explain ~71–79% variance in student outcomes (RQ2; Equity Equation in Section 4.4).
- Rural constraints (e.g., 0% formal policies) underscore necessity of targeted infrastructure and training investments to achieve full scalability (visualised in **Figure 4** heatmap).
- Aggregated details, pre/post trajectories, and replication syntax are provided in **Supplementary Table 3** and **Supplementary File 2** for independent verification.
- These systemic thresholds directly inform the conditional policy roadmap in **Supplementary Table 2** and align with NEP 2020 goals for equitable technology integration and faculty capacity building.

3.4 Integrated Findings

Triangulation across quantitative metrics and qualitative themes revealed consistent patterns. Combinations of high literacy and policy coverage (>70%) were strongly associated with empowerment and moral agency narratives (r = 0.65–0.76; **Supplementary File 1**). Rural subgroups showed resilient patterns in low-tech collaboration and modest creativity gains ($p = .08$), indicating core pedagogical robustness (overall d = 0.80–0.90) even under constraint.

In summary, the ethical AI pedagogy generated substantial, sustained improvements in AI literacy, ethical awareness, and human-centric skills while achieving equity across gender and discipline—directly supporting RQ1 and RQ3. However, persistent urban–rural differences, mediated primarily by infrastructure access and institutional readiness (RQ2), highlight that full equity requires concurrent investment in connectivity, faculty capacity, and governance structures.

4. DISCUSSION

The findings reveal notable and sustained effects of the ethical AI pedagogy (d = 0.80–0.90) on AI literacy and ethical awareness, meaningful gains in human-centric skills, and full equity across gender and discipline. These patterns are moderated by persistent urban–rural disparities in infrastructure and institutional readiness. Interpreted through Floridi's (2013) information ethics and Baxter and Sommerville's (2011) socio-technical systems theory, the results largely align with directional expectations while highlighting the mediating role of systemic enablers in diverse Indian contexts (Ministry of Education, 2020). All interpretations remain associational given the quasi-experimental design.

4.1 AI Literacy as Moral Competence

The intervention group achieved a 37.7 percentage-point increase in AI literacy by End-Year 3 (78.9% vs. 51.3% in controls; d = 0.80; **Table 3**, **Figure 1**), substantially exceeding the short-term gains (typically 10–25%) reported in prior studies (Krishnan, 2025c; Ouyang et al., 2022; Ellikkal & Rajamohan, 2024). The sustained trajectory, including +31.4 points among non-technical students, reflects the successful integration of technical proficiency with ethical judgment—consistent with Floridi's (2013) view of information ethics as cultivated moral competence.

Rural participants still registered a respectable 28-point gain despite infrastructural constraints (**Figure 3**), with qualitative reflections illustrating emergent habits such as "using OATutor to analyse village health data ethically" (**Supplementary File 1**). Faculty confidence mediated approximately 71% of literacy variance (r = 0.71), suggesting that scaling offline modules and targeted training could extend benefits to underserved settings (Essel et al., 2022; Joseph et al., 2024).

4.2 Equity Through Fairness Awareness

High ethical awareness (≥4/5 rubric score) was attained by 76% of intervention participants compared with 39% of controls (RR = 1.95, d ≈ 0.90; **Table 4**), with bias and privacy detection accuracy showing a large effect (d = 0.90). These gains correlated with technical literacy (r = 0.65, $p < .01$; **Figure 2**) and preserved complete gender equity (52% of high scorers female, χ^2 p = .502), countering common disciplinary and gender biases observed elsewhere (Usher & Barak, 2024; Singh et al., 2021).

The 23-point urban–rural gap in ethical awareness (85% vs. 62%) mirrored institutional policy voids (**Figure 4**, **Table 5**) and reinforced socio-technical barriers documented in the Global South (Jobin et al., 2019; UNESCO, 2021). Rural relative risk (RR = 2.07) nevertheless indicated meaningful adaptation potential. Faculty-led innovations, such as inclusive lesson redesigns, advanced fairness principles (Goyal et al., 2025), and adoption gaps correlated inversely with disparities (r = −0.72), suggesting that mandatory NEP-aligned ethics guidelines could accelerate uptake closure (NITI Aayog, 2021).

4.3 Human-Centric Skills as Virtue Ethics

Equity extended to human-centric domains: creativity (+23.5%, $d = 0.50$), ethical reasoning ($2.1 \rightarrow 4.4/5$, $d = 0.60$), and collaboration (+18.2%, $d = 0.40$) all improved significantly in the intervention group (**Table 3**, **Figure 2**). These gains, including +23.1% among non-technical students, indicate that scaffolded pedagogy fostered habitual moral reasoning and peer-supported teamwork (Ruiz-Rojas et al., 2024).

Rural creativity gains remained modest (+12.5%, $p = .08$), likely due to equipment limitations, yet low-tech reflections—"we shared ideas freely without devices"—revealed resilient collaboration (**Supplementary File 1**). Faculty confidence mediated 82% of creativity variance in urban settings versus 62% in rural, suggesting that mobile kits and peer networks could narrow the virtue-ethics gap (Barczak et al., 2023).

4.4 Institutional Outcomes and Synthesis

Systemic differences in readiness—curriculum uptake 75% urban vs. 30% rural; policy coverage 9/12 vs. 0/12 (**Table 5**, **Figure 4**)—correlated strongly with ethical awareness ($r = 0.76$) and confirmed infrastructure and capacity as critical socio-technical enablers (Baxter & Sommerville, 2011).

The synthesised Equity Equation (Outcomes ≈ Curriculum × Infrastructure × Capacity) captures how literacy combined with policy coverage >70% consistently triggered empowerment and moral agency cycles ($\beta \approx 0.76$; **Supplementary Table 3**). Faculty confidence mediated 79% of adoption variance overall ($\beta = 0.79$). These thresholds suggest that targeted NEP investments in connectivity, training, and governance could realistically stabilise observed gains within one policy cycle (Challa, 2025; NITI Aayog, 2021; conditional roadmap in **Supplementary Table 2**).

4.5 Alternative Explanations

Early spikes (64.5% at Year 1) may reflect novelty effects (Pardos et al., 2023), while modest control-group diffusion contributed approximately 9% of gains (**Figure 1**; Tang & Hare, 2023). Workshop selection and baseline technical differences pose modest inflation risks (Gupta & Bhaskar, 2020). Regression models attributed 65% of variance to the curriculum ($\beta = 0.72$), with covariate matching substantially reducing bias ($p > .05$ post-adjustment).

4.6 Unintended Consequences

Dual patterns emerged: widespread empowerment ("prepared for responsible tech roles") coexisted with threefold higher rural anxiety linked to access gaps ($r = 0.69$) and elevated abandonment risk in low-connectivity settings (Wang, 2024; Córdova et al., 2024). Faculty barriers amplified stress in approximately 30% of rural cases (Gupta & Bhaskar, 2020). Offline buffers and phased implementation mitigated these effects; peer-support kits are recommended for future scaling.

4.7 Limitations and Future Directions

The quasi-experimental design provides strong associational evidence ($r \approx 0.72$ main effect) but cannot establish definitive causality; selection bias and contamination risks were mitigated but not eliminated. Attrition (10–15%) slightly favoured persisters ($p = .09$). Findings generalise well across India's diversity (**Table 1**) but less to hyper-rural or elite-only contexts.

Future research should prioritise large-scale RCTs ($n > 2{,}000$), five-year behavioural tracking, and cross-national comparisons in the Global South (e.g., India–Brazil–Indonesia) using the openly shared materials (**Supplementary Files 1–4**; Ruiz-Rojas et al., 2024; Joseph et al., 2024).

4.8 Implications and Recommendations

The Equity Equation offers a replicable framework for Global South HEIs confronting digital divides (Smith & Neupane, 2022). A conditional three-point roadmap (**Supplementary Table 2**) prioritises:

1. Satellite-hybrid connectivity and offline AI kits for rural/low-resource HEIs
2. Nationwide faculty certification with ethical AI modules
3. Mandated institutional ethics policies with annual auditing

These actions, contingent on meeting observed thresholds (>70% policy/literacy synergy), could realistically elevate rural readiness toward urban levels within one NEP cycle. Concerted investment aligned with 2026 initiatives (Bharat Bodhan AI Conclave, India AI Impact Summit) would transform ethical AI education from constrained potential into scalable practice (UNESCO, 2021; Challa, 2025).

5. CONCLUSION AND IMPLICATIONS

This three-year quasi-experimental mixed-methods study offers robust associational evidence that an ethical AI pedagogy—merging technical literacy with socio-technical ethics training and offline-accessible delivery—can substantially enhance AI mastery, moral competence, and human-centric skills in diverse Indian higher education contexts. Intervention participants reached 78.9% AI literacy and 76% high ethical awareness ($d = 0.80–0.90$) compared with controls, with non-technical students and women attaining fully equivalent gains—directly aligning with NEP 2020's emphasis on interdisciplinary and gender-equitable education. Qualitative insights brought this to life: reflections such as "questioning AI fairness reshaped my designs" and "I now feel prepared for responsible tech roles" underscore how sustained scaffolding fostered genuine moral agency over compliance.

Yet the results also expose a persistent urban–rural divide: urban HEIs achieved 85% ethical awareness and near-complete policy coverage, while rural counterparts averaged 62% awareness and 0% formal ethics policies. Infrastructure access and faculty capacity emerged as primary mediators ($\beta \approx 0.35–0.79$), accounting for modest rural creativity gains (+12.5%, $p = .08$) and threefold higher initial anxiety. These disparities do not diminish the pedagogy's efficacy; rather, they highlight that equitable scaling hinges on parallel investments in connectivity, educator training, and governance—as synthesised in the Equity Equation (Outcomes ≈ Curriculum × Infrastructure × Capacity) from Section 4.4.

Theoretically, the enduring blend of technical and ethical elements affirms Floridi's (2013) information ethics as cultivated moral competence and extends socio-technical systems theory to Global South education realities (Baxter & Sommerville, 2011; Dignum, 2019). Practically, the replicable elements—offline OATutor modules, bias-audit tasks, virtue-ethics peer reviews, and audit templates—serve as an adaptable blueprint for HEIs navigating similar divides.

Boundary conditions are clear: the quasi-experimental design provides strong associational evidence but not definitive causation; 13% attrition and moderate-connectivity focus limit generalisability to hyper-rural or elite settings; self-reported measures introduce inherent subjectivity. Future work should pursue large-scale RCTs, extended behavioural tracking, and cross-national Global South comparisons (e.g., India–Brazil–Indonesia) using the shared resources.

In policy terms, the threshold insight—literacy × policy coverage >70% initiates virtuous empowerment cycles—translates to actionable strategy. Conditional on multi-crore investments in satellite connectivity, faculty certification, and mandatory ethics frameworks, rural readiness could approach urban levels within one NEP cycle, evolving ethical AI education from aspirational to achievable reality.

Ultimately, this study demonstrates that ethical AI integration in higher education is pedagogically potent—provided systemic barriers receive the same scrutiny as algorithmic ones. When curriculum, infrastructure, and governance align, AI becomes a tool for genuine inclusion, equipping India's youth to lead responsible innovation in the Global South and beyond (UNESCO, 2021; Selwyn, 2019; Challa, 2025).

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of AI-Assisted Technologies: The authors acknowledge the use of OpenAI's language model, ChatGPT, to assist with grammar checking, sentence restructuring, and reference alignment. The tool was used solely to improve the clarity and consistency of the manuscript, and the authors take full responsibility for the content of the work.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

Barczak, A. L., Mathrani, A., Han, B., & Reyes, N. H. (2023). Automated assessment system for programming courses: A case study for teaching data structures and algorithms. *Educational Technology Research and Development*, *71*(6), 2365–2388. https://doi.org/10.1007/s11423-023-10277-2

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers, 23*(1), 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

Challa, J. S. (2025). *Sakshm AI: Advancing AI-assisted coding education for engineering students in India through Socratic tutoring and comprehensive feedback* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2503.12479

Córdova, P., Grájeda, A., Córdova, J. P., Vargas-Sánchez, A., Burgos, J., & Sanjinés, A. (2024). Leveraging AI tools in finance education: Exploring student perceptions, emotional reactions and educator experiences. *Cogent Education*, *11*(1), Article 2431885. https://doi.org/10.1080/2331186X.2024.2431885

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer. https://doi.org/10.1007/978-3-030-30371-6

Ellikkal, A., & Rajamohan, S. (2024). AI-enabled personalized learning: Empowering management students for improving engagement and academic performance. *Vilakshan-XIMB Journal of Management*, *22*(1), 28–44. https://doi.org/10.1108/XJM-02-2024-0023

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education, 19*(1), Article 57. https://doi.org/10.1186/s41239-022-00362-6

Facer, K., & Selwyn, N. (2021). *Digital technology and the futures of education: Towards 'Non-Stupid' optimism*. UNESCO Futures of Education initiative. https://unesdoc.unesco.org/ark:/48223/pf0000377071

Floridi, L. (2013). *The ethics of information*. Oxford University Press.

Goyal, H., Garg, G., Mordia, P., Ramachandran, V., Kumar, D., & Challa, J. S. (2025). *The impact of large language models on K-12 education in rural India: A thematic analysis of student volunteer's perspectives* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2505.03163

Gupta, K. P., & Bhaskar, P. (2020). Inhibiting and motivating factors influencing teachers' adoption of AI-based teaching and learning solutions: Prioritization using analytic hierarchy process. *Journal of Information Technology Education: Research, 19*, 693–723. https://doi.org/10.28945/4640

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018*. Canadian Intellectual Property Office, Industry Canada. https://doi.org/10.1136/bmjopen-2020-039246

Indian Council of Medical Research. (2017). *National ethical guidelines for biomedical and health research involving human participants*. ICMR. https://ethics.ncdirindia.org/asset/pdf/ICMR_National_Ethical_Guidelines.pdf

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Joseph, G. V., Athira, P., Thomas, M. A., Jose, D., Roy, T. V., & Prasad, M. (2024). Impact of digital literacy, use of AI tools and peer collaboration on AI assisted learning: Perceptions of the university students. *Digital Education Review, 45*, 43–49. https://eric.ed.gov/?id=EJ1434328

Krishnan, S. (2025a). Harnessing generative AI in higher education: Opportunities, challenges, and strategies for responsible integration. *International Journal of Emerging Technologies and Innovative Research, 12*(8), 12–18. https://doi.org/10.56975/jetir.v12i8.568737

Krishnan, S. (2025b). Scoping the landscape: Opportunities, challenges, and strategies for generative AI in higher education. *Indian Journal of Computer Science, 10*(4), 8–19. https://doi.org/10.17010/ijcs/2025/v10/i4/175747

Krishnan, S. (2025c). Ethics-integrated AI pedagogy and metacognitive growth: A longitudinal quasi-experimental study in Indian higher education (SSRN Scholarly Paper No. 5526736). *SSRN*. https://doi.org/10.2139/ssrn.5526736

Ministry of Education, Government of India. (2020). *National Education Policy 2020*. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

NITI Aayog. (2021). *Towards responsible AI for all: Approach document for India part 1 – Principles for responsible AI*. https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies, 27*(6), 7893–7925. https://doi.org/10.1007/s10639-022-10925-9

Pardos, Z. A., Tang, M., Anastasopoulos, I., Sheel, S. K., & Zhang, E. (2023). OATutor: An open-source adaptive tutoring system and curated content library for learning sciences research. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Article 416, pp. 1–17). https://doi.org/10.1145/3544548.3581574

Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability, 16*(13), Article 5367. https://doi.org/10.3390/su16135367

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Smith, M., & Neupane, S. (2022). *Artificial intelligence and human development: Toward a research agenda*. International Development Research Centre. http://hdl.handle.net/10625/60907

Tang, Y., & Hare, R. (2023). *Combining gamification and intelligent tutoring systems in a serious game for engineering education* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2305.16568

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Usher, M., & Barak, M. (2024). Unpacking the role of AI ethics online education for science and engineering students. *International Journal of STEM Education, 11*(1), Article 35. https://doi.org/10.1186/s40594-024-00493-4

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning, 29*(3), 1–22. https://doi.org/10.1007/s10758-024-09744-3

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, Article 160018. https://doi.org/10.1038/sdata.2016.18

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

List of Abbreviations

Statistical and Methodological Terms

- ANCOVA: Analysis of Covariance
- ANOVA: Analysis of Variance
- CI: 95% Confidence Interval
- d: Cohen's d (effect size)
- κ: Cohen's Kappa (inter-rater reliability)
- MANOVA: Multivariate Analysis of Variance
- p: p-value
- r: Pearson Correlation Coefficient
- α: Cronbach's Alpha (internal consistency reliability)
- β: Standardized Regression Coefficient
- χ^2: Chi-Square Test

AI and Technology Terms

- AI: Artificial Intelligence
- DPDP: Digital Personal Data Protection Act, 2023 (India's data privacy law)
- FAIR: Findable, Accessible, Interoperable, Reusable
- HEIs: Higher Education Institutions
- ITS: Intelligent Tutoring System
- LLM: Large Language Model
- OATutor: Open Adaptive Tutor (AI platform used in the study)

Policy and Institutional Terms

- ICMR: Indian Council of Medical Research
- MeitY: Ministry of Electronics and Information Technology
- MoE: Ministry of Education
- NEP 2020: National Education Policy 2020
- NITI Aayog: National Institution for Transforming India
- UGC: University Grants Commission
- UNESCO: United Nations Educational, Scientific and Cultural Organization

Other

- MMAT: Mixed Methods Appraisal Tool
- NVivo: NVivo Qualitative Data Analysis Software
- RQ: Research Question
- TAM: Technology Acceptance Model