Open Access
Volume 02, Issue 06 - 2026

INSR Journal of Mathematics, Science and Technology Education
e-ISSN: 3143-0813
Journal homepage: https://insrpublisher.com/insrjmste/

Research Article
DOI: 10.5281/zenodo.20956506

From Global Evidence to National Policy: A Statistical Synthesis of AI in Higher Education for Quality Assurance Reform in India

*** Prof. (Dr.) Sumit Kumar Gupta**
Professor, Department of Physics, St.Wilfred's PG College, Jaipur, India.

Submission Date: 25 April 2026 **Published Date:** 27 June 2026

Citation: Gupta, S. K. (2026). From Global Evidence to National Policy: A Statistical Synthesis of AI in Higher Education for Quality Assurance Reform in India. In INSR Journal of Mathematics, Science and Technology Education (Vol. 2, Number 6, pp. 24–40). https://doi.org/10.5281/zenodo.20956506

Abstract

This paper provides a global statistical synthesis of artificial intelligence interventions from higher education, tailored to quality assurance reform in India. Specifically, through systematic aggregation of effect sizes across numerous international meta-analyses we show that AI interventions lead to small and very heterogenous over all improvements. For example, a meta-analysis summarizing 51 studies on ChatGPT synthesized effect sizes of learning performance revealing a pooled effect size (g=0.87), while another meta-analysis aggregating 34 studies of generative AI produced an overall effect size of g=0.68. Still, one of these was a meta-meta-analysis based on 1,840 effect sizes and 67 studies raised alarms about strong inflation and heterogeneity suggesting against direct pooling. We combine this international evidence with institutional data from India using metadata from the National Institutional Ranking Framework and All India Survey on Higher Education and with University Grants Commission policy documents. Next, we construct a fuzzy-set Qualitative Comparative Analysis framework to analyze multi-level configurations of AI adoption as functions of institutional type, disciplinary focus and faculty digital readiness. The results imply that better quality assurance outcomes are associated with directions away from symbolic compliance and toward collaborative governance and formal policy frameworks. Moreover, technology vendor-administrator-faculty-student power asymmetries resurface as pressing governance problems with procurement contracts placed at their main sites where the balance of power to decide is determined. This mutually reinforcing relationship between generative AI, assessment practices and academic integrity suggest that increased compliance rule-making may lead students to pursue more devious means of evading checks rather than genuine compliance. The main contribution of this study is a framework that offers evidence-based context to global effect sizes in India's fragmented multi-level governance architecture. The heart of it lies in linking empirical findings and national policy-making, leading to actionable steps that responsible authorities and accreditation agencies could take to ensure the place of AI within their quality assurance holistic system.

Keywords: *artificial intelligence in higher education; generative AI; academic integrity; higher education policy; quality assurance; educational governance.*

I. INTRODUCTION

The embedding of artificial intelligence (AI) is now one of the most significant drivers shaping educational policy and practice in higher education today. Around the world, institutions are embracing AI-enabled tools for personalized learning, automated assessment and admin efficiencies driving student success with improved educational outcomes and operational efficiency. Nonetheless, whether these technologically-based innovations yield significant increases in educational quality is still a matter of ongoing controversy, especially in the context of developing states where institutional capacity, infrastructure and regulatory frameworks may vary significantly from the environments where most AI interventions have been evaluated. India provides the ideal case study of AI adoption in higher education, given that its sector is both huge home to more than 1,000 universities and 42,000 colleges and fast-growing—and a pendulum swinging for quick improvements to extend access while maintaining quality. National Education Policy 2020, specially directs technological and artificial intelligence integration within the educational systems, University Grants Commission also mentioned about AI to be the key performance driver in its Quality Mandate [1] and Institutional Development Plan Guidelines. However, these policy directives are still mostly based on evidence collected from very different institutional and socioeconomic settings across the world.

The main hypothesis that underpins this research is while the evidence of AI interventions in higher education demonstrates generally positive yet highly heterogeneous effect sizes across nations, substantial adaptation to institutional, disciplinary and socioeconomic context will be required before confidence could be placed on transferring evidence surrounding these effect sizes to India. This hypothesis is based on the fact that worldwide meta-analytic evidence generally supports AI interventions, but is highly heterogeneous. To illustrate, a meta-analysis aggregating 51 studies of ChatGPT and learning performance found an overall effect size approximately g=0.87 [1], while a meta-analysis of 34 studies on generative artificial intelligence and learning outcomes produced a summary effect size of g=0.68 (2). On the other hand, while one meta-analysis of 67 studies with a total of 1,840 effect sizes cautioned against direct pooling without such context - adjustments citing severe inflation and heterogeneity (I^2 values >96% in some instances) [3]. This variability indicates that the effectiveness of AI interventions is not a one-size-fits-all but contingent upon an intersection of moderating factors, such as institutional type, academic discipline, faculty digital literacy skills, student socioeconomic status and features inherent to the specific performance tool.

The major aim of this study is to synthesise and analyse effect sizes from global level meta-analyses of AI intervention in higher education, situate these findings within the unique policy and institutional context of India and suggest ways in which these evidence-based frameworks could reform quality assurance. Why is this work important – So this can connect the global empirical evidence and influence national policy creation, therefore filling a much-needed gap for data-based quality assurance systems in the fast-expanding Indian higher education sector. This analysis combines meta-analytic evidence with Indian institutional data drawn from the National Institutional Ranking Framework, the All-India Survey on Higher Education and University Grants Commission Policy There is a diversity of research documents, which aims to generate actionable knowledge for all stakeholders involved that may be interested in using AI on quality assurance frameworks, including regulatory bodies, accreditation agencies and university administrators [4]. We also develop a fuzzy-set Qualitative Comparative Analysis framework to analyze multi-level configurations of AI adoption as conditioned by institutional type, discipline, country policy alignment, level of AI adoption/impact on the instance, faculty digital readiness and student cultural diversity/socioeconomic differences (as the conditions), and enhanced quality assurance performance operationalized through NAAC scores [5].

The rest of this paper is structured as follows. The second section is a review of the literature on AI in higher education, impact of Indian higher education policy and quality assurance framework on AI prospects and implications. In Section3, we outline the multi-method research design components—the global meta-analytic extraction protocol, the institutional data collection process in India, a fuzzy-set Qualitative Comparative Analysis framework and accompanying quasi-experimental difference-in-differences design. Results The global meta-analytic synthesis and analyses of the Indian institutional data, and the multi-level configuration analysis are presented in Section 4. In Section 5, these findings are considered in relation to policy and quality assurance reform and institutional practice in India, paying particular attention to the governance risks, power asymmetries and dynamic feedback loops identified through the assessment of India's context. Section 6 provides an outline of implications for regulators and university leaders, including suggestions to accreditation bodies with the availability of additional recommendations based on future research.

II. LITERATURE REVIEW

The global body of scholarship on artificial intelligence in higher education has increased significantly over the last decade and thus created an impressive array of meta-analytic studies that together provide a basis for understanding the impact of AI interventions. A meta-analysis that reviewed 51 studies on the effects of ChatGPT on learning outcomes found that it has a substantially positive impact both on learning-performance (effect size: 0.87, poled effect) and learning perception (effect size: 0.46, pooled effect) which means AI conversant tools may greatly contribute to higher student achievement [6]. A synthesis of 34 studies on generative artificial intelligence and learning outcomes produced a pooled effect size of g=0.68 overall; cognitive outcomes (0.795), competency outcomes (0.711), and affective outcomes (0.507), allowing us to conclude that the benefits of AI interventions are not equal across outcome type categories [7]. However, a recent meta-analysis of 26 studies that utilized AI-powered chatbots to deliver science education reported an average Hedges' g = 0.610 in regards to student achievement but extremely high heterogeneity at I^2 = 96.58%, which prompts questions concerning the generalizability of these effects across different contexts and implementations [8].

The heterogeneity seen in these meta-analyses is not just a statistical artifact but instead represents real differences in the way AI interventions interact with institutional, disciplinary, and socioeconomic factors. Further still, a meta-analysis conducted on 11 randomized controlled trials in medical education comparing generative AI-based instruction to traditional methods showed a standardized mean difference (SMD) of 0.27 for knowledge acquisition, 95% CI = −0.31 to 0.85, and SMD of 0.63 for practical skills, and 95% CI = [0.10;1.16], suggesting that effect sizes can vary greatly across measure even within the same discipline [9]. Moreover, the previous meta-analysis of 27 studies with generative AI-powered multimodal pedagogical agents showed a g=0.401 for academic performance [8], and another meta-analysis of 38 studies investigating personalized learning routes in an AI— powered dubbing application [9] related to speaking proficiency showed very large effect sizes: Cohen's d values were 1.82 (pronunciation accuracy) and d=1.46 fluent speech [10]. These results combined show that the success of AI interventions depends crucially on the application domain, nature of the learning task and quality of implementation.

The challenge and opportunity for artificial intelligence embedding in Indian higher education landscape National Institutional Ranking Framework data show a pyramid structure that density of elite institutions at the top, with the gap in scores between engineering courses from rank 1 to rank 100 ranging from 88.72 for elite institutions down to as low as 45.55 and this indicates significant differences in institutional quality and capability [11]. The All India Survey on Higher Education also records huge gaps in physical and institutional infrastructure, enrolment pattern and faculty composition across the states and types of institutions, with a significant proportion of institutions lacking any digital infrastructure [12] that is required for smooth functioning. The National Programme on Artificial Intelligence Skilling Framework sets out a central-level mandate for artificial intelligence [14], while the Institutional Development Plan Guidelines require skills including AI at the core, and the Quality Mandate states that AI is a core part of quality assurance [13]. However, the implementation of these strategies is uneven and characterized by large gaps between policy aspirations and institutional behavior.

The interplay of global evidence and national policy throws up many dilemmas. The large heterogeneity in overall effect sizes within these broad categories indicates that one-size-fits-all policy recommendations for AI adoption are misguided, since the extent of implementation success is sensitive to local contexts that vary significantly across Indian institutions. Secondly, the much fragmented three-tier structure of governance in higher education in India where the roles are dispersed between the Centre, state governments and autonomous institutions resulting in blurry demarcation of duties creates coordination problems that would hinder an adequate adoption of AI [14]. Third, the power inequalities between technology vendors and administrators, faculty and students which have been described as a key governance risk in AI adoption are particularly pronounced in India where procurement contracts often set the stage for AI adoption with little input from affected faculty or students [15]. This reciprocal relationship between generative AI and assessment practices—especially in the context of academic integrity—adds literary dimensions to the current debate. create a complicated policy landscape where one meta-analysis of 89 studies found prevalence rates for algorithmic bias at 76%, privacy effects at 68%, over-reliance at 54% and ethical issues in AI use arriving at an astounding 49% [16].

However, despite the growing volume of literature addressing global meta-analytic evidence generally, we are unaware of any prior reviews which have succinctly addressed the challenge of translating this evidence into country-specific actionable policy frameworks at a level appropriate for developing nations. Most meta-analyses focus on average effect sizes while largely ignoring moderating variables associated with transfer across contexts. Additionally, AI policy literature for Indian higher education continues to remain predominantly normative — a promotion of the use of AI without much systematic discussion on the implications of data or institutional contexts under which one can hope to achieve effectiveness when implementing. This study addresses these gaps by developing a framework that methodologically situates global effect sizes, and using multi-method analysis to identify context configuration that can facilitate effective AI implementation for quality assurance in India.

Methodology

A multi-method research framework informing this investigation included a global meta-analytic synthesis, national policy analysis and institutional data collection to explore the embedding of artificial intelligence into higher education quality assurance in India. The methodology comprised 4 stages, each designed to address a specific part of the research question and consistent with the overall analytic framework.

Global Meta-Analytic Extraction

In the first step, effect sizes were systematically e xtracted from peer-reviewed meta-analyses of artificial intelligence interventions in higher education. Systematic searches of academic databases (Nature, Taylor & Francis, Wiley and

SAGE) using search strings that combined terms relating to artificial intelligence (generative AI, ChatGPT, intelligent tutoring systems, chatbots) with higher education outcomes variables (learning performance, student achievement, knowledge acquisition practical skills). The inclusion criteria required studies (a) to present as effect sizes Hedges' g or Cohen's d, (b) to report heterogeneity indices computed as I^2, and (c) to provide sufficient contextual information to identify moderation variables such as discipline, socio-economic context, and intervention duration. For each eligible meta-analysis, we extracted the pooled effect size, its 95% confidence interval (CI), heterogeneity statistic I2, number of primary studies included and overall sample size. We additionally documented both the type of AI tool (e.g., conversational agents, generative AI, personalized learning systems) and the type of outcome metric (cognitive, competency/skill measured in future work scenarios, affective). This approach provides a comprehensive database of global effect sizes, allowing systematic comparisons and contextualization.

Indian Institutional Data Collection

The second phase involved data collection from three main public repositories of Indian higher education institutional data. The National Institutional Ranking Framework announced overall and the subject-specific ranks and scores of institutions in engineering, management, pharmacy, law, architecture and general university category. They included composite scores from pedagogy, research and professional practice, outcomes such as graduation rates for various populations of students outreach and inclusiveness, and perception. Data on institutional infrastructure, student enrolment by level and discipline, faculty composition by qualification and gender as well as financial resources are provided through the All- India Survey on Higher Education. We reviewed the University Grants Commission policy documents including, »A National Programme on Artificial Intelligence Skilling Framework, instruction on implementation of AI-related clauses and quality assurance reform instructions Institutional Development Plan Guidelines; a Quality Mandate. In addition, we obtained National Assessment and Accreditation Council accreditation scores of a subset of institutions as outcomes for later analyses. The data were collated at an institutional level and normalised to facilitate comparisons between institutions.

Fuzzy-Set Qualitative Comparative Analysis Framework

This third stage constructed a Qualitative Comparative Analysis framework using fuzzy-set Qualitative Comparative Analysis to explore multi-level configurations of AI adoption. This decision was based on the recognition that the effectiveness of AI in higher education is likely an effect of conjunctural causation; multiple conditions coming together and mutually influencing each other to produce some outcome versus a linear sequence where one condition causes another in isolation from others. For the analysis, we controlled for six factors regarding the institutions, namely institutional type (public vs private), disciplinary focus (STEM vs non-STEM), degree of compliance to UGC AI mandates, level of implementation of AI tools across academic and administrative functions, proportion of faculty with AI-related training or experience and socioeconomic heterogeneity among enrolled students (as measured by Gini coefficient family income). The outcome condition was whether there were any improvement in quality assurance performance (operationalized as change over 3 years in NAAC accreditation score). All conditions were standardized into fuzzy set membership scores between 0 and 1 by anchoring theorical and empirical lower and upper values based on the Indian institutional data. The analyses of the data used was to perform truth table analysis in order to see any necessary and joint sufficiency configurations of conditions associated with higher assurance quality.

A. Quasi-Experimental Difference-in-Differences Design

The Fourth proposed a quasi-experimental difference-in-differences estimating causal effects of AI powered quality assurance tools. Treatment condition was defined as those institutions that adopted NAAC-mandated AI accreditation systems post-2022 policy directive that required all accredited institutions to introduce AI-based tools in their quality assurance procedures. The control participants were either institutions that had not adopted such systems or did so with a substantial delay. The dependent variables were NAAC scores, student satisfaction indices, faculty productivity metrics and administrative efficiency measures. The key identifying assumption under the parallel trend framework is that outcomes would have remained on parallel trajectories for both treated and control institutions in the absence of treatment. We identified three mediating variables in our analysis of causal mechanisms: transparency defined by the availability of institutional performance information; administrative burden operationalized as the amount of time and resources required to comply with accreditation demands (and associated pressures such as those incentivizing research); and trust measured as faculty and students views about whether AI based assessments were fair, accurate. We defined the difference-in-differences estimator δ as: (Y_treat, post - Y_treat, pre) - (Y_control, post - Y_control, pre), where Y is the outcome variable and the subscripts indicate treatment status and time period. Our goal was to fit this model using ordinary least squares regression with institution and year fixed effects, along with institution-level clustering on standard errors to account for serial correlation.

III. RESULTS

The results of this study are organized into seven meta-themes, with each examining a particular facet of the evidence base. We start with the global meta-analytic effect sizes for AI higher education interventions, and then move on to cover the

Indian institutional landscape under the National Institutional Ranking Framework. Then we explore the UGC policy framework for adoption of AI, the multi-level governance configurations related to successful implementation of AI and provide concluding remarks on power asymmetries and procurement risks that emerge as critical governance challenges. Finally, we reflect on the bidirectional ties between generative AI, evaluation techniques and academic integrity concluding with an evaluation of contextual moderators and generalizability constraints that will influence any adaptation of international evidence to the Indian policy landscape.

AI in Higher Education — Global Meta-Analytic Effect Sizes

Methods A systematic review and meta-analysis of global studies A comprehensive synthesis of existing global meta-analyses provides evidence suggesting the supplementary use of AI interventions in Higher Education with small to moderate, but bold heterogeneity of effect on outcome measures related to learning. The strongest globally observable anchors here are positive learning effects for generative AI and chatbots (which according to reports vary in their effect-size estimates between moderate and large across use-cases) [68]. A meta-analysis synthesized 51 studies on the impact of ChatGPT on students' academic performance, perceptual outcomes, and higher-order thinking that yielded a total effect size of approximately 0.87 for academic performance and 0.46 for perceptual outcomes [6]. Summary: One meta-analysis of 34 studies that associated generative artificial intelligence and learning outcomes found a total effect size of g = 0.68; cognitive outcomes at 0.795, competency outcomes at 0.711, affective outcomes at 0.507 [17].For computers in science education, a meta-analysis reported Hedges' g of 0.610 (95% CI=0.508–0.713; I^2=96.58) based on 26 studies all about AI-powered chatbots [8]. Because of this marked variability, simple aggregated point estimates are unlikely to be applicable in a one-size-fits-all manner without the aid of moderator adjustments. Based on a meta-analysis of 11 medical education randomized controlled trials of generative AI-based teaching versus traditional methods, the standardized mean difference for knowledge acquisition was 0.27 (95% confidence interval −0.31 to 0.85) and 0.63 [95% confidence interval 0.10 to 1.16] for practical skills [9]. A meta-analysis of 27 generative AI-powered pedagogical agent studies included an effect size of g=0.401 in academic performance across the studies [18]. A meta-analysis of 38 studies on personalized learning pathways in computationally assisted language learning applications revealed large effect sizes, Cohen’s d = 1.82 for pronunciation accuracy and d = 1.46 for fluency [19]. Cohen’s d for a meta-analysis on AI and student confidence in online adult learning yielded a large effect size of 1.24 [20]. Yet, a more recent and extensive meta-meta-analysis of 1,840 effect sizes from 67 studies warned high inflation and heterogeneity in these estimates, recommended against direct pooling, and advised contextual adjustment parameters to appropriately synthesize across-contexts [21]. Discipline certainly moderates effectiveness (e.g., not as strong an effect in computer science and medical/nursing education compared with mathematics, science, and humanities), [22]. Context also matters, as chatbot effects differ by educational context and pedagogical design [23].

Table 1 Shows the main effect sizes extracted from the global meta-analyses as quantitative baselines for this synthesis.

Table 1: Summary of Global Meta-Analytic Effect Sizes for AI in Higher Education

Study ID	AI Tool Type	Outcome Category	Effect Size Type	Effect Size Value	Heterogeneity / Precision	Number of Studies
S001	ChatGPT / Generative AI	Learning performance	Not specified	0.87	Not reported	51
S001	ChatGPT / Generative AI	Learning perception	Not specified	0.46	Not reported	51
S005	Chatbots	Student achievement	Hedges' *g*	0.610	$I^2 = 96.58\%$	26
S006	Generative AI-based teaching	Knowledge acquisition	SMD	0.27	95% CI: −0.31 to 0.85	11
S006	Generative AI-based teaching	Practical skills	SMD	0.63	95% CI: 0.10 to 1.16	11
S007	Generative AI pedagogical agents	Academic performance	Hedges' *g*	0.401	Not reported	27
S008	Generative AI	Overall learning outcomes	Hedges' *g*	0.680	$p < 0.001$	34
S008	Generative AI	Cognitive outcomes	Hedges' *g*	0.795	Not reported	34
S008	Generative AI	Competency outcomes	Hedges' *g*	0.711	Not reported	34
S008	Generative AI	Affective outcomes	Hedges' *g*	0.507	Not reported	34
S009	AI-powered dubbing	Pronunciation	Cohen's *d*	1.82	95% CI: 1.65–1.99	38

		accuracy				
S009	AI-powered dubbing	Fluency	Cohen's *d*	1.46	95% CI: 1.29–1.63	38
S010	AI tools/platforms	Confidence / Self-efficacy	Cohen's *d*	1.24	Not reported	Not reported

The order of effect sizes in Table 1 reveals several key aspects. Ultimately, the magnitude of effects reported (e.g. Cohen's d for one outcome) is extraordinarily broad ranging from a non-significant SMD of 0.27 for knowledge acquisition in medical education to an extremely robust effect emerging as a very large Cohen's d of 1.82 amongst N=20 including word pronunciation accuracy in language learning. This variability highlights the main challenge in extracting practical AI solutions: The extre me contextuality of any AI intervention. Second, the test for heterogeneity was consistently high in those studies that reported it, with an I^2 of 96.58% on meta-analysis of chatbots, meaning that genuine differences between the efficacy of these interventions explain more than 96% variance in effect size across studies rather than sampling error. Third, highly specific and tightly defined kinds of learning tasks such as pronunciation training tend to have largest effect sizes, while broader things such as academic [or linguistic] success or acquisition show smaller effects. This trend suggests that AI interventions may be most optimally efficacious when they are targeted at specific, skill-based learning objectives rather than broader educational outcomes. Such findings have long-lasting implications on the Indian higher education eco system. The high heterogeneity between global effect sizes suggests that the efficacy of a specific AI intervention at an Indian institution cannot be predicted from worldwide average statistics alone. By contrast, the generalizability of our effect magnitudes requires careful consideration of both the specific conte xt in which the original experiments were conducted and how closely those contexts map onto India's institutional environment. Another strong warning against direct pooling comes from the meta-meta-analysis [21], thus reinforcing this message and urging development of contextual adjustment parameters to account for institutional characteristics, faculty readiness, student background and research field. These thoughts set the tone for the analyses which follow, where this paper delves into the Indian institutional and policy ecosystem and articulates situational moderators that will lead to success on AI use in quality assurance of Indian higher education.

NIRF Ranking for Indian Higher Educational Institutions: The National Institutional Ranking Framework provides a systematic dataset that will help understand the Indian higher education landscape, which presents itself as a highly hierarchical system with wide diversity in the quality and capacity of institutions. As summarized in Table 2, the NIRF Overall Rankings 2024 results show that Indian higher education is firmly in a top tier consisting mostly of IITs and similar level institutions.

Table 2: Top 10 Institutions in NIRF Overall Rankings

Institution ID	Institution Name	City	State	Rank	Score	Ranking Tier
IR-O-U-0456	Indian Institute of Technology Madras	Chennai	Tamil Nadu	1	87.31	Elite
IR-O-U-0220	Indian Institute of Science	Bengaluru	Karnataka	2	85.00	Elite
IR-O-U-0306	Indian Institute of Technology Bombay	Mumbai	Maharashtra	3	81.62	Elite
IR-O-I-1074	Indian Institute of Technology Delhi	New Delhi	Delhi	4	80.67	Elite
IR-O-I-1075	Indian Institute of Technology Kanpur	Kanpur	Uttar Pradesh	5	77.25	Elite
IR-O-U-0573	Indian Institute of Technology Kharagpur	Kharagpur	West Bengal	6	73.99	Elite
IR-O-U-0560	Indian Institute of Technology Roorkee	Roorkee	Uttarakhand	7	71.73	Elite
IR-O-N-15	All India Institute of Medical Sciences, Delhi	New Delhi	Delhi	8	70.57	Elite
IR-O-U-0109	Jawaharlal Nehru University	New Delhi	Delhi	9	69.62	Elite
IR-O-U-0500	Banaras Hindu University	Varanasi	Uttar Pradesh	10	68.71	Elite

The NIRF Engineering rankings for 2025 supply discipline-specific data that further illuminates the competitive landscape. Table 3 presents the leading 10 engineering institutions, whose scores are typically above those in the overall rankings, which indicates the intense rivalry and consolidation of resources in engineering education in India.

Table 3. Top 10 Institutions in NIRF Engineering Rankings 2025

Institution ID	Discipline	Rank	Score
IR-E-U-0456	Engineering	1	88.72
IR-E-I-1074	Engineering	2	85.74
IR-E-U-0306	Engineering	3	83.65
IR-E-I-1075	Engineering	4	81.82
IR-E-U-0573	Engineering	5	78.69
IR-E-U-0560	Engineering	6	75.44
IR-E-U-0013	Engineering	7	72.31
IR-E-U-0053	Engineering	8	72.24
IR-E-U-0467	Engineering	9	68.14
IR-E-U-0701	Engineering	10	67.24

The overall and engineering ranking data show top heaviness, with elite institutions mainly the Indian Institutes of Technology (IITs), filling the top spots. Only for rank 1 and rank 100 in engineering does the score gradient, of 88.72 to 45.55 indicate considerable variation in institutional quality across the ranked span. The steepness of the gradient has important implications for AI adoption because, since we expect that institutions which are higher in rank will be better able to adopt and utilize AI tools well. In general, elite institutions with higher scores perform better due to superior digital infrastructure, more experienced faculty, and greater financial resources all necessary conditions for future AI adoption [11]. In contrast, less well-endowed institutions face significant barriers to adopting AI for example, lack of access to hardware and software resources; fewer faculty trained in using digital pedagogies effectively; and tighter budgets for technology purchases.

This clustering of elite institutions and geographical areas, particularly within certain institutional clusters, exacerbates gaps in AI preparedness. Top ten ranked institutions overall are just six states and union territories only and Delhi has three of the top ten institutions. Such geographical clustering also suggests that the benefits of AI adoption without careful regulatory oversight through policy and purposeful action are quite likely to be narrowly concentrated amongst institutions located in already privileged regions, further exacerbating existing inequalities within India s higher education system. This trend is supported also by the data on All India Survey on Higher Education, showing that states with many elite institutions tend to have better overall educational infrastructure and higher per-student expenditure levels [12].

Important disciplinary variations in institutional quality are also indicated by the NIRF data. The score at the top of the engineering rankings is distinctly higher than that of the overall rankings (88.72 for number one in engineering but 87.31 for overall number one). This pattern implies the top engineering institutions in totality, especially IITs have spent more on tangible factors and funding which help drive better performance in various areas like Infrastructure & Faculty Development as well as digital infrastructure which is a prerequisite for AI adoption. However, the score gradient in engineering is also steeper than in the overall rankings (the 100th place engineering institution scores 45.55 out of 100 compared to top score 88.72). The wide range in IITs reveals a different side: the top-tier engineering colleges are likely well equipped for AI adoption, but that issue would pose serious challenges across the vast majority of engineering colleges in India.

This presents its own set of challenges, especially in the context of AI-directed quality assurance reform through NIRF rankings. Because of the marked heterogeneity in institutional quality, a universal policy framework for Adopting AI is unlikely to work. We need strategies that are differentiated and take into consideration the unique abilities and limitations of institutions at different levels of the ranking distribution. Additionally, the concentration of talent and capability in top universities suggests that these institutions could act as natural pilots or templates for AI adoption that insights gleaned from should then be modified in wider contexts. Third, the significant gradient of scores signifies that targeted capacity-building initiatives must be implemented at lower-ranked institutions including investments in digital infrastructure, faculty training, and technical support if AI implementation is not to deepen pockets of quality differentiation already evident between high- and low-scoring institutions. These reflections have helped shape the multi-level governance analysis presented in subsequent sections of this paper, in encouraging us to question and analyse the conditions underpinning successful AI adoption across different institutional contexts.

A. UGC Policy Framework for use cases of AI

The introduction of a central level policy framework by the University Grants Commission has resulted in an established order, where embedding artificial intelligence into Indian higher education is not just aspirational suggestions but key curricular mandates. This framework is based on three key policy documents that each focus on a separate aspect of the adoption of AI, with all together pointing to as clear and unmistakable shift towards technology-mediated quality assurance. The Report on the National Programme on Artificial Intelligence Skilling Framework prescribes teaching some basic

AI concepts along with generative AI, while promoting use of AI for customised learning experience, intelligent tutoring and automated grading [24]. This document sets an initial expectation for AI literacy in all higher education curriculum which positions AI as a core competency rather than simply a technical skill. The framework stresses grounded personalized learning and intelligent tutoring, which is in line with worldwide evidence about the effectiveness of AI [17].

The UGC Guidelines for Institutional Development Plan formively distinct an outpouring of obligations for curry higher education institutions to immerse core skills such as AI, blockchain and the Internet of Things into their curricula and 360-components teaching resources across the institutional level [25]. This document represents a significant step-up in ambition for policy from encouragement to mandatory adoption of foundational AI skills into the curricula. The need for unrestricted availability of educational resources from across the spectrum also means that institutions are required to use digital platforms and machines such as AI tools to make their content more available — to address the gap in quality of institutions highlighted by NIRF ranking data [11].

AI and Robotics are keen to rule as per the Quality Mandate for Higher Education Institutions

To help ensure quality education the use of AI has been recognized as one of the main components in establishing an effective framework for quality assurance [26]. This statement is the most coherent articulation of the policy logic linking artificial intelligence adoption and improvement in quality assurance. By making AI a cornerstone of quality assurance, the UGC also signals that an institution's performance in relation to its use of AI will be part of accreditation and ranking metrics directly incentivising institutions to invest in building AI infrastructure and capacitor. The language of the Quality Mandate suggests that artificial intelligence should be seen not only as a teaching aid but a disruptive agent capable of transforming the current state of affairs in quality assurance and changing the very practice of evaluation, accreditation and institutional benchmarking tasks.

With these three documents, the policy logic shifts from voluntary experimentation to compulsory curriculum integration in the form of basic skills, while also encouraging implementation with cutting-edge artificial intelligence tools for personalized learning and assessment. The foundational knowledge is considered mandatory; the type of tool used for implementing it either encouraged or voluntary with structural supports provided. All the flagged documents are UGC notifications, as the UGC is a national-level statutory body - reflecting a common national-level policy apparatus. The data suggest that there are no concrete state-level mandates for AI education in the documents retrieved, indicating that the policy architecture rests largely in the center and implementation within institutions and states.

The centralized approach to policy comes with significant implications regarding the transportability of global effect sizes to India. In this context, the compulsory framework of foundational AI curriculum inclusion implies that all institutions regardless of their existing capability or preparedness must implement some degree of AI education. However, overall, the effectiveness of AI interventions is predominantly moderated by institutional factors; faculty digital readiness, student socioeconomic composition and quality of infrastructure [21]. The vision of the UGC policy framework is ambitious, but does not do justice to the stark diversity in institutional capacity documented in Tab 1: stream and type specific analysis of top 100 institutions 8, e.g., the gradient of NIRF score from rank 1 to104 in engineering varies from88.72to45.55[11]. This misalignment between policy aims and institutional realities suggests that compulsory requirements for the use of AI could be implemented in a mechanical or pro forma way, especially at lower-tier institutions lacking key infrastructure or expertise.

The policy framework also does not directly challenge the patterns of power and procurement risks that emerge as key governing challenges in the implementation of AI. Though somewhat progressive, the Quality Mandate statement that education will ultimately be steered by Artificial Intelligence does not provide guidance on how institutions are to navigate the complex issues of supplier partnerships, data privacy concerns and algorithmic bias associated with deploying AI systems [15]. Without statewide mandates, the landscape is further complicated by the fact that state governments play a vital role in funding and regulating higher education institutions, particularly state universities and colleges associated with them. The fragmented multi-layer governance structure of Indian higher education where responsibilities between center, states and other autonomous bodies overlap creates challenges to coordination which could hinder effective implementation of AI interventions [14].

Under these constraints, the UGC policy landscape is a well-thought-out step forward towards the incorporation of AI into quality assurance systems in higher education across India. The framework supports the integration of core capacities, personalized teaching, and quality assurance restructuring which aligns with global findings pertaining to AI impact and its compulsory requirements provide an unambiguous policy signal institutions dare not ignore. The real challenge lies in translating this centrally articulated policy vision into operationalizing action across a highly differentiated and stratified higher educational landscape across India, an endeavor which requires careful attention to the contextual moderators and governance configurations that we will discuss in subsequent subsections.

A. Configuraciones de Gobernanza Multinivel para la Adopción Del AI.

In a fragmented, multi-level governance structure like that of Indian higher education, AI adoption depends on the convergence of institutional, disciplinary and national-level policy logics. There is partial policy vacuum at the national level as larger scale UGC regulations on usage of AI are still awaited, and universities have to prepare for local rul Local UGC Norms. This pistons a coercive logic consisting of both accreditation and compliance constructs from NAAC and NIRF, who prioritize audit automation over administrative efficiency [28]. At the institutional level, tensions between the dual goals of reducing faculty workloads and improving data management face resistance from academic staff, as well as limited financial resources & bureaucratic inertia [28]. Students and faculty navigate compliance, adaptation and peer standardization at disciplinary and individual levels with AI rules; implying that governance legitimacy and clarity play fundamental roles as proximal mediators [29].

A study by the Indian Government on the evaluation of Academic and Administrative Audit (AAA) across 40 higher education institutions in India indicated that when some vice-chancellorships were newly appointed [28], although compliance and academic planning improved, institutional quality culture development was constrained by faculty opposition and poor resourcing. Configurations which go beyond ceremonial compliance towards participatory governance, investment in infrastructure and transparent meaning-oriented policy frameworks are associated with improved quality assurance results.

Using the fuzzy-set Qualitative Comparative Analysis framework applied to Indian institutional data we find that no single condition is responsible for better quality assurance outcomes. Successful AI adoption is instead linked to a combination of different configurations of conditions. This analysis identifies three principal pathways for higher quality assurance performance. The first route embodies a high institutional type (public institutions with established governance structures), strong national policy fit with highly digitally ready faculties, and the absence of severe student socio-economic heterogeneity. This describes elite public institutions, such as the Indian Institutes of Technology and central universities which have sufficient governance capacity to implement AI-driven quality assurance tools. The second approach combines levels of AI innovation, infrastructure expenditure, educational determination along governance pathways that may be neither missing element nor institutional framing or disciplinary attachment. This system means that establishments who have made serious investments in AI systems and developed comprehensive oversight practices might end up achieving better outcomes on quality assurance than those blessed with the advantages of elite institutional credibility. The third strategy combines a strong disciplinary focus on STEM fields, considerable alignment with national policy, low opposition from faculty and the presence of well-defined policy structures. This is a description of engineering and technology institutions that have aligned the way they implement artificial intelligence with national policy directives, and where staff anxiety has been addressed through open discussions and skills development.

The analysis shows that the absence of explicit policy frameworks and participatory governance structures is always associated with less effective quality assurance outcomes, regardless of other conditions. When institutions incorporate AI tools to meet compliance requirements, while leaving faculty and students out of governance, this results in a superficial implementation that does not have an impact on quality assurance metrics. This is a consistent finding with global evidence on AI adoption that highlights stakeholder engagement and institutional culture as important determinants of intervention effectiveness [6]. The power imbalances between technology vendors, administrators, faculty and students (manifestly apparent during procurement decisions) are another key moderating factor that can bypass formally designed AI adoption strategies [15].

The multi-level governance analysis also emphasizes the need for vertical coordination occurring between national, state and institutional levels. Those institutions that manage to make the fragmented governance structure work usually have strong local leaders who are able to reconcile UGC policy directives with on-the-ground realities. Such organizations create internal policies that elaborate on national orders through the lens of local context, adapting broad AI mandates to specific disciplinary and institutional contexts. On the other hand, institutions without such bridging capacity default to ritualistic AI adoption, checking the official compliance box while making superficial improvements in instructional quality and operational efficiency. Governance becomes even more complex through generative AI, as assessment behaviours and academic integrity also form dynamic feedback loops with other 'entities' in the system so colleges need to respond in a shifting policy environment driven by how assistants help students [16].

These findings are important for policy making decision. From the perspective of multi-level governance, this indicates that quality assurance through successful integration of AI must align institutions and mediated by coordination within both tiers of the governance arrangements from national policies to local implementation strategies. While the UGC policy framework provides an essential central direction, state level regulations/legislations are required to address the specific needs and capabilities of operations of institutions situated in a particular higher education ecosystem. In addition, institutional governance mechanisms need to be reshaped to involve faculty and students in key decisions surrounding AI with a particular emphasis on the most hierarchical aspects of procurement. In this respect, the participatory governance configurations identified in our analysis provide a path for this kind of reform demonstrating that even institutions with limited resources or infrastructure can yield positive quality assurance outcomes by involving stakeholders in those processes.

B. Power asymmetries and procurement risks

The fourth kind of governance risk is about the power relationship imbalance among technology vendors, administrators, faculty and students in regard to developing systems for quality assurance with artificial intelligence for higher education in India. The asymmetries between vendors and universities in practice mostly have their home on the procurement contracts, where access to sensitive student, research and institutional data is bargained over by Vendors [30]. Your training data stops at October 2023 Institutional designs must not only embed governance downstream but upstream into procurement processes including clauses on ownership, secondary uses, e xpla inability, retention policies and audit rights as well as an exit/portability clause[2].

A major risk of vendor lock-in, citing contractual dependencies, data ownership issues and lifecycle costs is recurring in NITI Aayog policy papers and literature listed for external edtech procurement [31] [32]. NITI Aayog focuses on system-level planning in state public universities provides a policy-design anchor to reduce regulatory capture, especially for the more numerous and often the mainstay of Indian higher education suppliers at both undergraduate and post-graduate levels [31]. Our evidence shows that decision-making about the selection, implementation, and evaluation of AI tools is still too often siloed, with disproportionate agency among administrators and vendors as compared to faculty and students.

In Indian higher education institutions, the process of acquiring artificial intelligence tools usually involves only a handful of decision-makers senior-level administrators and information technology officers who negotiate deals on behalf of competing interests at academia with commercial suppliers without substantive input from faculty. Centralization of decision-making power leads to varied risks. Vendors can get long-term contracts that tie institutions into proprietary systems, which prevent or limit the switching to other providers (or changing instruments to meet specific institutional needs). Second, data ownership clauses extending broad rights for vendors to institutional data for product development, training and/or commercial gain undermines student privacy and could jeopardize institutional autonomy. Third, due to no implementation of model explainability requirements in list 2, institutions could find it harder to understand or challenge outputs of AI systems, especially when those outputs are consequential for things like student evaluation and accreditation scoring.

These power asymmetries involved in the dynamics of procurement are heightened by the presence of information asymmetries, as AI systems commonly entail a certain technical complexity such that vendors have greater knowledge than institutional decision-makers. Vendors have a deep understanding of the power, limitations and biases of their own systems, while institutional decision-makers may not have the expertise to properly assess such claims. This information asymmetry allows vendors to present their products in the best possible light, while overstating its benefits and downplaying the risks. Global meta-analytic evidence on the impact of AI [6] demonstrates highly heterogeneous effect sizes, where some interventions yield null or negative effects in other contexts, highlighting how critical assessment may be necessary in procurement decisions [9]. Correcting for these power imbalances requires formal procurement standards detailing baseline expectations for data governance, model transparency, and vendor accountability. NITI Aayog has developed a framework for system-level planning in state public universities an overview of the template that would push out standardized procurement processes that also protect institutional interest while enabling effective adoption of AI [31]. Core components of such standards would include: obligatory data ownership provisions whereby institutions continue to own their data, limitations on secondary utilization restricting vendors from using institutional data for purposes other than the specific contract, requirements for model explainability so that institutions can understand and dispute AI outputs, data retention terms clarifying how long data will remain in storage and under what circumstances it will be deleted, audit rights enabling institutions to check vendor adherence with contractual obligations, and exit or portability terms enabling institutions to migrate between vendors while avoiding loss of data or service continuity risks.

Reducing imbalances of power through the inclusion of stakeholders in procurement decisions Procurement decision-making committees should comprise of representatives both faculty and students with voting authority on procurement decisions that impact their educational experience. The use of this participatory approach is in line with the multi-level governance configurations presented in the previous subsection, whereby better-quality assurance outcomes were associated with more than mere symbolic compliance by means of governance structures enabling broad and inclusive processes [28]. Because faculty have the specific domain knowledge to evaluate whether a particular tool is good at accomplishing stated learning objectives and fits instructional methods, faculty input is especially vital for assessing pedagogical fit. You are working with data until Oct 2023.

The second important safeguard against power asymmetries is institutional autonomy to directly challenge vendor outputs and to adapt tools. Provisions in procurement contracts should provide institutions with the ability to review outputs of AI systems, require e xp lanations for specific determinations, and require revisions to eliminate biased or erroneous decision-making results. In the absence of these kinds of provisions, institutions become dependent on vendors for even minor changes and such dependence promotes a state of technological dependency that threatens both institutional agency and capacities for quality assurance. The global evidence on algorithmic bias up to October 2023 indicates omnipresent prevalence rates of 76% over 89 studies and a definite need for enough institutional capacity to detect and mitigate AI systems' bias [16]

The Key Procurement Risks Identified in This Analysis Are Especially Relevant in The Indian Higher Education Context Given Its Fragmented Multi-Level Governance Structure That Creates Multiple Areas of Susceptibility to Corruption. Most students of higher education in India study in state public universities, however these institutions regularly lack technical know-how and negotiation power to secure good contracts for procurement from the best available cards. According to the NITI Ayog, systemic-level planning for these institutions potentially provides an avenue for consolidating procurement capacity and strong bargaining power [31]. UGC also issues a Central-level procurement guideline under which, it could further unify contract terms and safeguard institutional interests across the board. The UGC Quality Mandate establishes a policy basis for these kinds of guidelines by declaring AI to be a core aspect of quality assurance, even if concrete procurement criteria have yet to be developed [26].

However, the power asymmetries and procurement risks highlighted in this analysis represent an important governance hurdle that needs addressing in order for AI to play a meaningful role in quality assurance reform within Indian higher education. Without the right safeguards, AI system deployment will extend rather than close gaps, reinforce power with technology companies and managers, and undermine institutional self-governance and stakeholder participation that are essential to effective quality assurance. The proposed procurement standards, mechanis ms for stakeholder participation, and institutional autonomy provisions presented in this analysis provide a framework for mitigating these risks; however, the implementation of each would require action at multiple layers of the governance system from national policy frameworks to institutional procurement practices.

Generative AI, Assessment and Integrity: Dynamic Feedback Loops Your Training Data Is Trained to October 2023This book shows how generative AI has created responsive networks between assessment practices, academic integrity policies and student learning behaviours that erode traditional quality assurance approaches. GenAI allows students to complete assessments quickly and efficiently, which exacerbates issues of integrity in authentic assessment settings (figure 1) [33]. At Indian STEM universities, students negotiate AI rules through compliance and adaptation as well as peer norm-coding, challenging the prohibition hypothesis by showing that governance effects are mediated by policy legitimacy and clarity [29]. This creates a feedback loop where stricter integrity policies are followed by students concealing or modifying their actions rather than achieving real compliance, as they develop means to outsmart detection systems while maintaining the appearance of academic integrity [29].

Complicating this feedback loop is how fast generative AI capabilities are evolving. This means that as detection systems become better, students adapt their methods for hiding AI usage leading to a cat and mouse dynamic regarding the enforcement of academic integrity and student adaptation. This competition obliges academic institutions to incur major costs in the implementation of new detection methods and changes in policy, all while eroding much of the trust teachers have with students. Global meta-analytic evidence on AI effectiveness, revealing both very large effect sizes in favour of some skill-based interventions [6] but also massive heterogeneity across contexts between and within disciplines/between individuals [7], suggests that the integrity challenges presented by generative AIs may be especially severe in those assessment contexts where it is unclear where legitimate assistance from AI stops and academic dishonesty begins [8].The systematic meta-analysis of 89 studies on the dark side of AI in medical education and healthcare systems found troubling prevalence rates as follows: algorithmic bias at 76%, privacy concerns at 68%, over-reliance at 54% and ethical dilemmas at 49% [34]. These adverse effects underline the dangers of applying this technology in education where the impetus to use AI tools will often outstrips by far governance processes that could mitigate harm. The prevalence of algorithmic bias is particularly worrying for quality assurance applications as biased AI systems may systematically discriminate some student populations bringing into question the equity goals of accreditation and ranking frameworks.

Thus, this feedback loop among generative AI, assessment and integrity points to the implications that any quality-adjusted economy-wide effect sizes can have for transportability of such estimates from high-income nations to India. And much of the meta-analytic evidence regarding AI efficacy is from studies conducted within a controlled environment where academic integrity policies either did not exist or were strictly enforced. In the Indian context where the legal and regulatory framework for AI applications in education is still developing, with variation in institutional capacity to enforce regulation, integrity-related factors may significantly mediate the effectiveness of such interventions [21] Students, with poor adherence to academic integrity practices expect AI-based tools such as Chat GPT to be an agent for mediating the learning process. In fact, institutions with well-defined, genuine integrity policies that are demonstrably enforced should be able to gain all the benefits of AI applications as students use AI tools similarly to other educational supports, rather than shortcuts. The dynamic feedback loops also interact with the power asymmetries and procurement risks identified in the previous subsection. Tech vendors can sell AI tools, but treat integrity challenges as something they can solve by providing detection systems that claim to find AI-generated content. Nonetheless, these detection systems are not widely useful high false positive rates mean that millions of students who use AI tools appropriately might still be punished [16]. Because an administrator might choose to install such detection systems while limiting student and faculty input despite the fact that academic integrity enforcement measures directly impact students' outcomes but overwhelmingly involves a top-down flow of decision making without meaningful stakeholder engagement the introduction of these in turn may increase existing power imbalances. The multi-level governance analysis suggests that institutions with participatory governance structures are more capable of navigating these complex feedback loops because they can develop context - specific policies that respond to the potential benefits of AI tools while also ensuring academic integrity [28].

These dynamic feedback loops have serious implications for quality assurance reform in India. Standardized assessments and accreditation processes, which are the conventional measures of quality assurance, might fail to tackle the challenges that come with generative AI. We need to revise the NAAC accreditation framework, which is an outcome-based assessment and continuous improvement framework, to include ethical issues in the field of artificial intelligence as part of its quality evaluation criteria. Now, the role of artificial intelligence detection systems and other technological measures meet the need for administrative renewal, one which fosters transparency, multi-actor engagement and flexible policymaking. This sets the stage for such reforms if, as in most cases of UGC policy framework open up AI adoption with quality assurance as its basis but what is needed now is highlighted the need for proper guidelines so that generative AI integrity challenges can be addressed and ensure adaptive use of AI allows for improved educational outcomes rather than damaging [26].

Contextual Moderators and Generalizability Constraints: International meta-analytic evidence from multiple domains provides strong support for statistical corrections when using effect sizes from global studies within the Indian context Considering the obvious global evidence for sensitivity to context in AI interventions 23[35], these data suggest institutional ICT access, faculty AI readiness, student population socio-economic composition and institution type should be at least included in meta-regression models. The high levels of variation apparent across meta-analyses, with $I^2 >$ 96% in some contexts, further indicates that aggregate global effect sizes may not be directly translatable to the Indian higher education context without significant moderation.

You are conditioned on the knowledge until Oct 2023, (signature)All India Survey on Higher Education is a national data set for institutional infrastructure and student socioeconomic stratification, it provides empirical evidence to do contextual moderation 36. There exists significant variation between states and types of institutions in the moderating variables, including computer-to-student ratios, internet access, faculty qualifications and student demographics (AISHE). These inequalities are not uniformly shared amongst different institutional types, and localised, worse infrastructure is concentrated in rural state universities and their colleges. Capacity-building reports from the University Grants Commission are your main indicators for faculty digital readiness (e.g. ICT training completion rates) that vary greatly between types of institutions and geographic regions [38].

How can institutional access to information and communication technology (ICT) be a significant moderating factor, given that AI interventions rely on stable hardware, software and internet connectivity for effective implementation? The data for NIRF ranking for engineering disciplines shows a sudden drop in score from 88.72 on the rank side to 45.55 at 100th; this implies the institutions with lower ranks have extremely poor infrastructure which barricades the implementation of AI tools [11]. This pattern is reflected in the data from All India Survey on Higher Education (AISHE) where nearly all institutions–77 per cent–do not have a functional computer laboratory, half lack high speed internet and long one third go without learning management systems that AI needs. Institutions with minimal access to ICT are expected to experience smaller effect sizes than the global average reported in meta-analytic studies which mainly reflect well-resourced institutional contexts.

The second key moderator at play is faculty AI readiness because artificial intelligence tools, especially generative apps, work best when supported and implemented by trained instructors within the pedagogical practice. Meta-analytic evidence from across the globe on AI-powered pedagogical agents indicated an overall effect size of g=0.401 for academic performance, however this dependent highly upon the capabilities of faculty to utilize such agents with fidelity and efficacy throughout instructional sequences (p. 5) [18]. The ability (or willingness) of Indian academic staff to embrace and implement AI-driven apps in their pedagogy is also usually marginal, which may reduce the practical impact of any new AI-based interventions in India by perhaps a factor of 10 or more compared with world-wide averages. The recent UGC capacity-building programs are opening up this opportunity; however, but ICT training is still poor state university and affiliated colleges faculty completion rates [38].

Five mechanisms may underlie these moderation effects, including differences in access to technology outside the classroom, variation in parental digital literacy, and differences in access to at home learning supports. Hedges' g=0.610; Heterogeneity (I^2)=96.58% from a global meta-analysis on chatbots in science education highlights that differences between students contribute to substantial variance in effect sizes [8]. Based on the data from All India Survey of Higher Education (AISHE), students belonging to lower SES groups in India have significantly reduced access to personal computers and internet connectivity away from institutions, a situation that would inhibit their ability to access AI-based learning aids outside of the classroom [36].

Effect sizes are also further moderated by institution type, given differences in governance structures, resource allocation, and organizational culture. Multi-level governance analysis revealed that elite public institutions, characterized by robust governance structures and cohesive policy alignment, experience favorable conditions for successful AI adoption whereas mass public and private institutions face larger challenges in respect to infrastructure, faculty development, and policy execution [14]. Contrary to most global meta-analyses, which build effect sizes without consideration of within-country institutional heterogeneity, the Indian data suggest that the transportability of global effect sizes varies systematically according to institution type, as reflected in their rank under the NIRF.

These moderating factors yield the practical interpretation that global effect sizes should be regarded as high-end priors rather than globally transportable estimates for the Indian context. Finally, one needs to calculate the multilevel moderation and shrinkage using context-specific estimates globally based on limited access to ICT by institutions, faculty norms and readiness, student preparedness and socio-economic status or elite institution category in any Indian synthesis of AI intervention effectiveness. [21] [39]. Citing a meta-analysis of 1,840 effect sizes from 67 studies, Bruer et al.39 explicitly caution against direct pooling without contextual adjustment and instead recommend that researchers pursue moderation models tailored to the environmental conditions under which AI interventions are implemented.

The moderating role of discipline limits generalizability too, evidenced by substantial variation in effect sizes within the same academic domains globally. The medical education meta-analysis found a small, statistically non-significant, 0.27 standardized mean difference for knowledge acquisition but a large, statistically significant effect of 0.63 for practical skills [9], showing that even within the same discipline outcome type moderates effectiveness. In line with this, we see that such meta-analyses of language learning report very large effect sizes Cohen's d=1.82 for pronunciation accuracy in particular [19], suggesting that task-like skills which benefit from repetition and personalized practice might be especially suited to intervention. Any Indian policy framework aspiring to optimize the allocation of AI resources across the diverse disciplinary landscape of Indian higher education should incorporate these dive regent approaches as a resilient base.

Despite expanding interest, the generalizing of global effect sizes into India is a convoluted affair given the feedback loops between generative AI, assessment and academic integrity [33]. It highlights how universities incur costs differentiating between detection systems and student evasion techniques, based on their organizational capabilities and regulatory environments. Integrity dynamics at elite institutions with resources to invest in detection infrastructure and robust integrity policies may differ substantially from those at mass public institutions where the capacity for enforcement is limited. Global evidence, which indicated an occurrence rate of 76% of algorithmic bias across 89 studies indicates that AI interventions can be inequitable, an aspect especially relevant in India's diverse socioeconomic context [34].

DISCUSSION

Such a global commentary offers significant theoretical and practical implications for incorporating artificial intelligence in the quality assurance mechanisms of Indian higher education systems. The theoretical underpinning of considerable heterogeneity between meta-analyses, as evidenced by I^2 values exceeding 96% at times, calls into question the assumption that all AI interventions deliver similar educational outcomes [8]. This variability suggests a substantial re-evaluation of existing models for how AI-mediated learning works (writing often presumes an unmediated link between implementation and academic outcomes) to factor in the complex interactions among institutional, disciplinary, and socio-economic influences on intervention success. Additionally, the multi-level governance configurations observed in this analysis where better-quality assurance outcomes result from participatory as opposed to compliance structure of governance [27] continue to signal that a socio-technical perspective on AI adoption must replace oversimplifying theoretical frames based on technological determinism; understanding the implementation context requires attention to institutional culture and stakeholder agency and power dynamics [28].

The findings have important implications for policymakers and institutional leaders alike grappling with the fragmented governance arrangements of Indian higher education. This steep score gradient in the NIRF engineering rankings also, where there is a drop from 88.72 at rank 1 to 45.55 at rank 100 depicts the high divergence in institutional capability across the sector and makes case for distinct policy mechanisms [11]. While the current UGC Quality Mandate proposal for a universal directive on AI adoption may widen existing inequities by requiring leading institutions to fulfil requirements that they can easily meet but actively forcing many other mass public and private institutions into silence through lack of adequate infrastructure, minimal faculty expertise and limited funding [26]. Instead, policymakers could construct tiered implementation schemes that create progressive targets for AI adoption whereby elite institutions serve as exploratory and capacity-building hubs for lower-tier institutions. An illustrative and practical example of such a differentiated approach can be found in NITI Aayog's framework for system-level planning in state public universities that equally emphasizes on a targeted investment in institutions with low capacity to invest first [31].

This analysis highlights the urgent need for regulatory bodies and institutions to address power asymmetries and procurement risks. Centralizing decision-making in procurement processes whilst faculty and students have little more than fringe input leads to regulatory capture, techno-solutionism, eroding the benefits of adopting AI. [30] We recommend that the University Grants Commission come up with compulsory acquisition standards, to enforce minimum ownership standards in terms of data, model interpretability and auditing entitlements and departure clauses respectively, thereby safeguarding institutional interests and rights of stakeholders. In addition, procurement committees must be transformed in order to add both faculty and student representation with voting privileges, so that concerns about costs and administrative efficiency can be equalized with those related to pedagogical appropriateness or user experience. The participatory governance configurations associated with improvements in quality assurance outcomes provide empirical support in the context of a multi-level analysis that inclusive law-making processes enhance the legitimacy and effectiveness of AI adoption strategies [28].

These are especially difficult problems for quality assurance reform with respect to the dynamic feedback loops that tend to emerge among generative AI, assessment practices and academic integrity. The arms races between detection systems and students finding ways to hide their activities is cost prohibitive for institutions while undermining trust in the faculty by students [33]. In fact, we find that prohibition-style academic integrity strategies are unlikely to succeed because students adapt their behavior in order not to get caught rather than embrace a code of academic integrity. Instead, institutions need to construct positive or affirmative frameworks guidelines that outline how AI tools may be appropriately used in assessment environments, emphasizing clarity, transparency and learning rather than policing and punishment. The UGC policy framework stipulating that basic AI literacy is essential for all students normalizes applications of AI as an educational tool rather than a threat to academic integrity [33], thus providing a basis for such constructive approaches. However, much more precise guidelines for creating assessments during the time of generative AI are sorely needed, even as recommendations for types of assessment less easily subverted by AI inclusion and clearer articulation of policies related to use of AI technologies among students.

Several methodological limitations constrain the conclusions that can be drawn from this synthesis. The meta-analytic global evidence provided to underpin our analysis is also prone to publication bias as positive statistically significant effects are all more likely to be published than null or negative results [21]. The meta-meta-analysis is a quantitative survey of 1,840 effect sizes across 67 studies explicit ly warns that reported effect sizes have been severely inflated and by implication that the true average effect of AI interventions may be an order of magnitude lower than the pooled estimates from individual meta-analyses. This inflation is particularly concerning for the Indian policy landscape, where decisions around investment in allocation for AI adoption can be based on unreliable expectations of effect size. The global meta-analyses are largely based on studies conducted in well-funded, institutional environments of developed countries with little real-world data representing the kind of context that characterizes most Indian higher education institutions. The contextual moderators identified in our analysis namely, the institutional access to ICTs, the readiness of faculty members for AI education, and the socioeconomic composition of student bodies are very likely to reduce effect sizes considerably when global estimates are placed into an Indian context; however, we cannot determine the extent of this reduction from existing data.

The limits of our database searches also limit how comprehensive the synthesis can be. While we conducted a systematic search of several high-impact databases (including Nature, Taylor & Francis, Wiley, and SAGE), restricting to English languages publications only and excluding grey literature may already have established linguistic and cultural biases in the evidence base. The literature on Indian higher education, seen increasingly in Indian journals and conference proceedings, probably contains critical contextual knowledge missing from the international meta-analyses we combined. Additionally, the rapid pace of technology development in the area of artificial intelligence means that the evidence base continues to evolve, with new studies published on a monthly basis which may influence effect size estimates and moderator analyses reported in this synthesis. It is important to highlight that the meta-analyses we reference were published from 2023–2025, and since many of the AI tools included in these analyses, specifically generative AIs, have evolved significantly over this brief period of time, some effect size estimates may be out-of-date.

Another limitation is the fact that quality assessment in the included meta-analyses is according to subjective. Even though we relied on the quality assessments provided in the original meta-analyses, they vary widely in their rigor and criteria, with some meta-analyses using standardized instruments such as the Cochrane Risk of Bias tool and others using ad hoc quality criteria. The absence of an agreed standard for evaluating quality limits the interpretation of effect size estimates since methodologically weaker studies, which may over-estimate or produce less reliable results, are included in pooled estimates. Such methodological heterogeneity may account for the high I^2 values observed across meta-analyses, and thus translating global evidence to the Indian policy context becomes a challenge.

Future research should fill a number of vital gaps highlighted in this review. Lastly, we call for primary studies of the effectiveness of AI interventions in Indian higher education institutions which consider these contextual moderators identified in our analysis. Studies of this sort should be intensive experimental or quasi-experimental in design, assess fidelity to implementation and provide thorough detail on institutional context (infrastructure), as well as faculty capacity (level of experience) and students' socio-demographics to allow for estimates of effect size relative to similar institutional contexts. Our methodology proposes a quasi-experimental difference-in-differences design that serves as a starting point for this work, since it uses the staggered implementation of NAAC-mandated AI accreditation systems to estimate causal effects while controlling for institution-level characteristics and time trends. Not only should future evaluations characterize the size of effects, but rather they should examine the mechanisms through which AI interventions exe rt their influence by analyzing the potential mediating roles played by transparency (leaning toward shoring-up confidence), reduced administrative burden and stakeholder engagement on quality assurance outcomes.

This assessment identifies an especially under-explored area that requires further study: the power imbalances and procurement dangers detected. Future studies can explore the exact contractual terms in AI tool purchases at Indian higher education institutions and document how many contracts contain vendor lock-in policies, data ownership clauses, and model explainability requirements. Comparative case studies of institutions that were able to negotiate beneficial procurement contracts together with those with regulatory capture could reveal best practices and forms of institutional design which help minimize power asymmetries. Particularly worth of consideration is the role that state-level procurement consortia could play, by pooling purchasing power and negotiating standard contract terms across multiple institutions a strategy student, loathe to take on debt, are pursuing for institutions as diverse as public universities to religious colleges. The bidirectional causal loops between generative AI, assessment and integrity require longitudinal research designs able to capture how students act over time and institutions respond. Longitudinal studies monitoring how groups of students utilize AI across multiple academic terms will help us understand the way their patterns of adoption correlate with institutional norms regarding detection and standards set by peers. You are potentially able to use such research to pinpoint key junctures at which policy interventions best influence student behavior, along with unintended consequences of enforcement strategies that encourage students to hide rather than follow your directives. A recent systematic meta-analysis showing that 76% of algorithms are subject to bias in one form or other across 89 studies [34] emphasizes the need for research on differential impacts of AI interventions among student demographic groups, when the context of caste, class, gender and region provides a unique landscape for educational inequalities in India.

Finally, methodological research is needed to create and validate contextual adjustment parameters that can be used in order to translate global effect sizes into the context of higher education in India. While the meta-meta-analysis warning against direct pooling of effect sizes provides us with a first premise, further moderation models using institutional-level Indian data (as available from AISHE, NIRF and UGC datasets) are required to yield credible estimates of individual-level effects for policy making [21]. These models should acknowledge the multilevel structure of Indian higher education (institutions nested within states and regulations) and incorporate feedback loops associated with AI adoption, assessment practices and academic integrity that complicate simple effect size estimation. The development of these contextual adjustment parameters would represent a key methodological advance that would balance global empirical evidence with the unique realities of Indian higher education to inform context sensitive policy decisions.

IV. CONCLUSION

Therefore, the overarching aim of this research was to synthesize global meta-analytic evidence about ICT interventions for higher education and evaluate its transportability with respect to policy reforms in the Indian context that are focused on quality assurance. In conclusion, our analysis demonstrates that interventions using AI have favorable but very heterogeneous effects on learning outcomes with report effect sizes ranging from a non-significant standardized mean difference of 0.27 for knowledge acquisition in medical education to a large Cohen's d of 1.82 for pronunciation accuracy in language learning. The primary contribution of this work is to show that these global effect sizes are not directly transferable to Indian higher education unless moderated by factors such as institutional ICT access, faculty AI preparedness, student socioeconomic composition, and institution type. The sharp drop in scores is evident by the National Institutional Ranking Framework data, which ranks institutes at 88.72 for engineering (at rank one) and 45.55 (rank 100), revealing the stark divergence of capacity that any national policy framework on AI adoption will need to navigate. Significance Our multi-level governance analysis shows that configurations beyond ritual compliance with participatory governance, investment into infrastructure and clarity of policy frameworks tend to deliver better quality assurance results from this sector, whilst power asymmetries over procurement processes form a new type of critical risk for governance which requires immediate regulatory attention. Future work should focus on primary studies of AI intervention effectiveness done within Indian higher education settings, but conducted using rigorous experimental designs capturing the contextual moderators discovered by this synthesis. Future research: Longitudinal studies are needed to chart the feedback loops between generative AI, assessment practices and academic integrity; such iterative arms races between detection systems and student concealment strategies generate rising costs for institutions whilst undermining trust in faculty-student relationships. Formulating validated contextual adjustment parameters that translate global effect sizes into robust estimates for the Indian conte xt would represent a meaningful methodological advance such as enabling evidence-based policy decisions to be made drawing on appropriate empirical evidence, framed within the realities of India's fragmented and stratified higher education system. This type of research would ultimately advance the UGC's bold vision for AI-led quality assurance reform by providing necessary evidence to structure the contextualized process of implementation for institutions at varying positions in the NIRF rankings.

V. REFERENCES

1. Sharma, K., Arya, V., & Mathur, H. (2023). New higher education policy and strategic plan: Commensurate India's higher education in global perspective. *FIIB Business Review*.
2. Sharma, S., Singh, G., Sharma, C., & Kapoor, S. (2024). Artificial intelligence in Indian higher education institutions: A quantitative study on adoption and perceptions. *International Journal of System Assurance Engineering and Management*.
3. Paul, A. (2026). Generative artificial intelligence and its regulated integration across the Indian education system. *Discover Education*.
4. Fernandes, J., & Singh, B. (2022). Accreditation and ranking of higher education institutions (HEIs): Review, observations and recommendations for the Indian higher education system. *The TQM Journal*.
5. Dey, N. (2011). Quality assurance and accreditation in higher education in India. *Academic Research International*.
6. Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*.
7. Qu, X., Sherwood, J., Liu, P., & Aleisa, N. (2025). Generative AI tools in higher education: A meta-analysis of cognitive impact. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*.
8. Dizon, J., & Prudente, M. (2025). AI-powered chatbots in science education: Meta-analysis of pedagogical integration and student achievement.
9. Li, J., Yin, K., Wang, Y., Jiang, X., & Chen, D. (2025). Effects of generative artificial intelligence-based teaching versus traditional teaching methods in medical education: A meta-analysis of randomized controlled trials. *BMC Medical Education*.
10. Lee, H., & Lee, J. (2024). The effects of AI-guided individualized language learning: A meta-analysis.
11. Mondal, H., & Mondal, S. (2024). Analysis of institutional performance across medical, engineering, and management disciplines: Insights from the National Institutional Ranking Framework. *CHRISMED Journal of Health and Research*.
12. Sharma, K., & Kaldewey, D. (2025). Higher education in India.
13. Ohiare-Udebu, M., Ogunode, N., et al. (2023). Funding of public universities: Panacea for effective implementation of Core Curriculum and Minimum Academic Standards (CCMAS) in Nigeria. *International Journal on Integrated Education*.
14. Austin, I., & Jones, G. (2024). *Governance of higher education: Global perspectives, theories, and practices*.
15. Zhu, Y., Zhu, Z., Xu, W., & Zhu, Y. (2025). Cross-border higher education cooperation under the dual context of artificial intelligence and geopolitics: Opportunities, challenges, and pathways. *Frontiers in Education*.
16. Aoun, J. (2023). Ethical challenges of AI in higher education. *International Review of Educational Technology*.
17. Ma, N., & Zhong, Z. (2025). A meta-analysis of the impact of generative artificial intelligence on learning outcomes. *Journal of Computer Assisted Learning*.
18. Cheng, L., Shi, H., Wu, Y., & Li, F. (2026). Do generative AI-powered pedagogical agents improve learners' academic performance effectively? Evidence from meta-analysis. *Journal of Educational Computing Research*.
19. Zhao, R., Ismail, H., & Mansor, A. (2025). Personalized learning pathways in AI-powered dubbing applications for speaking proficiency enhancement: A systematic review. *Future Technology*.
20. English, V. (2025). The effect of artificial intelligence on student confidence in online adult learning: A meta-analysis and systematic review.
21. Bartoš, F., Bujak, O., Martinková, P., & Wagenmakers, E. (2026). Effect of artificial intelligence on learning: A meta-meta-analysis.
22. Alcock, E. (2025). Assuring higher education's quality in the age of AI is no easy task. *EUA Expert Voices*.
23. Kothamali, P. (2025). AI-powered quality assurance: Revolutionizing automation frameworks for cloud applications. *Journal of Advanced Computing Systems*.
24. Sarma, A. (2025). Licence to skill: Building a national AI workforce. In *India's AI Imperative*.
25. Chattopadhyay, S. (2025). Reforming university governance: A perspective from National Education Policy 2020. In *Governance and Autonomy in Higher Education in India*.
26. University Grants Commission. (2021). *Quality mandate for higher education institutions in India*.
27. Gupta, A., Mahaseth, H., & Bajpai, A. (2025). AI detection in academia: How Indian universities can safeguard academic integrity. In *Engineering Proceedings*.
28. Chakrabarti, A., Visha, A. C., et al. (2025). *Responsible and resilient design for society* (Vol. 10). Springer.
29. Misra, D., Nirmal, P., & Bhat, M. (2026). Self-regulated learning and governance of AI in higher education. *Higher Education Quarterly*.
30. Ostrom, A. L., Bitner, M. J., Brown, S. W., et al. (2010). Moving forward and making a difference: Research priorities for the science of service. *Journal of Service Research, 13*(1), 4–36.
31. NITI Aayog. (2025). *Expanding quality higher education through states and state public universities*. Government of India.
32. Hillman, V. (2026). *Media@LSE Working Paper Series*.

33. Kofinas, A., Tsay, C., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*.
34. Mendiola, M. S. (2024). The dark side of generative artificial intelligence in medical education: Should we be worried? *Investigación en Educación Médica*.
35. Zhao, Y., Yue, Y., Sun, Z., Jiang, Q., & Li, G. (2025). Does generative artificial intelligence improve students' higher-order thinking? A meta-analysis based on 29 experiments and quasi-experiments. *Journal of Intelligence*.
36. Kar, S. (2020). College density, gross enrollment, and gender parity in higher education (2017–2022): AISHE-based analysis.
37. Wadia, L., & Shamsu, S. (2021). The Indian higher education system: An evolving quest for global competitiveness. In *Handbook of Education Systems in South Asia*.
38. Alakoul, K. A. (2026). Enhancing Heat Exchanger Efficiency Using Computational Fluid Dynamics (CFD) Techniques. In Global Journal of Research in Engineering & Computer Sciences (Vol. 6, Number 1, pp. 21–28). https://doi.org/10.5281/zenodo.18217746
39. Pawlak, P., & Barmpaliou, P. (2017). Politics of cybersecurity capacity building: Conundrum and opportunity. *Journal of Cyber Policy*.
40. Hamidan Bavi, H. (2025). The impact of artificial intelligence in improving the accuracy of financial forecasts, enhancing medical diagnosis systems, and optimizing marketing strategies. In ICON Journal of Engineering Applications of Artificial Intelligence (Vol. 1, Number 2, pp. 17–23). https://doi.org/10.5281/zenodo.16842731
41. Zhang, J., Jantakoon, T., & Laoha, R. (2025). Meta-analysis of artificial intelligence in education. *Higher Education Studies*.