AMERICAN Journal of Language, Literacy and Learning in STEM Education
Volume 4, Issue 09, 2026 ISSN (E): 2993-2769

Effectiveness of AI-Generated Study Materials versus Traditional Teaching Methods in Higher Education

Rakesh Kumar Gupta[1], Maumita Sengupta[2], Shirupa Kumari[3], and Shibnath Banerjee[4]
[1]Department of Management, Techno India University, Kolkata, West Bengal, India
[2]Department of Education, Techno India University, Kolkata, West Bengal, India
[3]Department of Management, Techno India University, Kolkata, West Bengal, India
[4]Department of Management, Techno India University, Kolkata, West Bengal, India

Abstract. *The increasing integration of artificial intelligence (AI) in higher education has significantly trans-formed the creation and dissemination of study materials. While AI-generated content offers advantages such as accessibility, scalability, and instant feedback, concerns remain regarding its depth, contex-tual relevance, academic integrity, and ability to support meaningful learning. This study examines student perceptions of AI-generated educational materials and explores faculty perspectives on AI-assisted evaluation, student dependence on AI-generated responses, and future assessment practices in higher education. A mixed-method research design was employed, involving 1,164 undergraduate students from diverse academic disciplines across private universities in India, along with structured interviews with faculty members from different age groups and younger AI experts. Quantitative data were collected using a structured 21-item Likert-scale questionnaire and analyzed using descriptive statistics, exploratory factor analysis (EFA), reliability analysis, and comparisons across academic years. Qualitative interview responses were thematically examined to identify concerns and proposed strategies related to AI-generated evaluation, academic integrity, student copying, and assessment redesign. The results indicate that students moderately acknowledge limitations of AI-generated materials, particu-larly regarding depth, contextual understanding, and human experiential insight. The EFA identifies a multidimensional structure explaining approximately 59% of the total variance, while the overall instru-ment demonstrates high internal consistency (Cronbach's Alpha = 0.917). Furthermore, senior students demonstrate greater skepticism toward AI-generated materials than junior students. Faculty perspectives highlight a contrasting but complementary view: some participants favor limiting AI use and retaining subjective evaluation, whereas others emphasize adapting assessment practices to the inevitability of technological disruption through case-based, practical, analytical, and data-driven examinations. The study concludes that AI-generated materials are better positioned as supplementary learning resources than as standalone instructional tools, and that effective integration requires both human oversight and redesigned assessment strategies that emphasize independent reasoning and practical application..*

Keywords: *Green chemistry, tertiary institutions, chemistry education, environmental sustainability.*

I. INTRODUCTION

The rapid development of Artificial Intelligence (AI), particularly generative artificial intelli-gence (GenAI) and large language models (LLMs), is reshaping higher education by changing how students access information, generate academic content, communicate ideas, and engage with learning resources. Contemporary systems such as ChatGPT and other large language models can generate explanations, summaries, examples, questions, feedback, and problem-solving assistance with minimal delay. Consequently, AI has moved beyond the traditional role of an automated educational technology toward becoming an interactive learning resource capable of supporting students and educators across a wide range of academic activities[1][2][3][4]. Systematic reviews of artificial intelligence in higher education have documented applications in learning support, assessment, prediction, and instructional assistance, while also noting that pedagogical integration and the role of educators require substantially greater attention The educational appeal of GenAI lies in its accessibility, scalability, responsiveness, and capacity to provide personalized assistance. Students can request explanations at different levels of complexity, obtain alternative formulations of difficult concepts, generate examples, and receive immediate responses without being constrained by classroom schedules or instructor availability[5][6]. Experimental and review-based evidence further suggests that ChatGPT and related systems can, under appropriate conditions, contribute to academic performance, affective-motivational outcomes, and some aspects of higher-order thinking[7]. Such capabilities create opportunities for self-directed learning and for expanding educational support beyond conventional teacher-led interactions. However, the educational value of AI-generated content cannot be determined solely by the fluency or apparent coherence of its outputs. Effective learning also depends on accuracy, contextualization, disciplinary depth, critical engagement, and the learner's ability to evaluate and apply information appropriately[8][9].

A major concern therefore relates to the quality and epistemic reliability of AI-generated educational materials. Large language models generate text probabilistically and can produce responses that are coherent and persuasive while containing factual errors, unsupported claims, or fabricated information[10][11]. The problem is particularly important in higher education, where students are expected not only to acquire information but also to distinguish reliable evidence from unsupported assertions and to engage critically with disciplinary knowledge. The absence of appropriate references, real-world contextualization, and domain-specific nuance can further reduce the usefulness of automatically generated study materials. Recent discussions of GenAI in education consequently emphasize the need for human oversight, fact checking, AI literacy, and pedagogical strategies that explicitly account for the limitations of these systems. The effectiveness of GenAI also depends substantially on how users interact with these systems. Prompt-based interaction means that the quality, specificity, and structure of user instructions can influence the resulting output[12]. In educational environments, this introduces an additional dimension of variability because students with different levels of AI literacy may obtain substantially different results from the same technology. Poorly formulated prompts may generate superficial, overly generalized, or inappropriate explanations, whereas carefully structured prompts can produce more targeted content. Thus, the presence of AI technology alone does not guarantee effective learning; its educational value is mediated by users' ability to formulate appropriate requests, assess generated outputs, and integrate them with established academic sources and disciplinary knowledge.

Another fundamental concern is the role of the human element in teaching and learning. Effective higher education extends beyond the transmission of information and includes inter-pretation, mentorship, contextual judgment, discussion, motivation, and connection of theoretical concepts to practical situations. AI-generated explanations can reproduce linguistic patterns and provide broad information, but they do not inherently possess the lived experience, pedagogical judgment, or contextual awareness of a human instructor. The importance of human involvement is also emphasized in international guidance on GenAI, which advocates a human- centred approach in which AI supports rather than displaces educational expertise and human agency. Accordingly, the comparison between AI-generated study materials and traditional teacher-led instruction should not be reduced to a comparison of speed or information avail-ability; it should also consider the depth, relevance, interpretation, interaction, and experiential dimensions of learning. The emergence of GenAI has also created significant challenges for assessment and academic integrity. AI systems can produce essays, explanations, solutions, and other forms of academic output that may resemble student-authored work, creating difficulties for conventional assessment methods that rely heavily on take-home written responses[13][14]. Research examining university teachers' assessment practices has demonstrated that AI-generated responses can perform convincingly against conventional examination prompts, raising concerns about the validity of assessment practices and the ability of instructors to distinguish student understanding from machine-generated production[15]. At the same time, evidence from assessment-focused research suggests that students may primarily use GenAI for writing, paraphrasing, and rephrasing rather than for deeper critical engagement, creating a risk that technology may facilitate completion of academic tasks without proportionate development of independent reasoning[16]. These developments have intensified calls for institutions to reconsider assessment design, academic integrity policies, and the balance between AI-permitted and AI-restricted activities. Importantly, the assessment challenge is not limited to whether AI should be prohibited. A more fundamental question is how educational evaluation should evolve when AI can increasingly generate plausible answers to conventional questions. Recent scholarship argues that assessment may need to place greater emphasis on application, interpretation, critical evaluation, process, and authentic demonstration of learning rather than relying exclusively on easily reproducible written outputs. Similarly, emerging work on AI-assisted assessment indicates that generative models may themselves contribute to feedback and evaluation processes, although issues of reliability, transparency, contextual judgment, and human oversight remain important[17]. This creates a paradox: AI may simultaneously challenge the validity of conventional assessment and provide tools that can assist institutions in developing new assessment practices.

Ethical and institutional considerations further complicate the adoption of AI in higher educa-

tion. Concerns related to bias, transparency, accountability, privacy, student agency, and unequal access have been repeatedly identified in the AI ethics literature[18]. UNESCO similarly emphasizes that educational institutions need coherent policies, appropriate safeguards, data protection, capacity building, and human-centred pedagogical frameworks for the responsible adoption of GenAI. Faculty readiness is particularly important because instructors are re-sponsible not only for deciding whether and where AI may be used, but also for redesigning learning activities and assessment practices in ways that preserve academic rigor. Yet the higher education literature has historically devoted considerable attention to technological capabilities and student-facing applications, while educator perspectives and institutional adaptation have received comparatively less empirical attention. Recent empirical research has begun to address student perceptions of GenAI. Studies have reported that university students recognize both the usefulness and limitations of ChatGPT, with perceived benefits including accessibility, efficiency, and learning support, while concerns include inappropriate use, academic misconduct, and uncertainty regarding effective institutional guidance. However, important gaps remain. First, much of the emerging literature examines specific groups, individual courses, or particular implementations of ChatGPT rather than providing broad evidence across diverse undergraduate disciplines. Second, research has often focused on students' use or acceptance of AI rather than examining the perceived quality and limitations of AI-generated study materials as educational resources. Third, the interaction between academic maturity and perceptions of AI-generated learning materials remains comparatively underexplored. Finally, although assessment and academic integrity are increasingly recognized as major concerns, fewer studies combine student perceptions with faculty perspectives on how evaluation itself should evolve in response to GenAI. To address these gaps, the present study employs a mixed-method research design that combines a large-scale quantitative investigation of undergraduate students with qualitative faculty interviews. The quantitative component examines the perceptions of 1,164 undergrad-uate students from diverse academic disciplines across private universities in India using a structured 21-item Likert-scale instrument. The analysis investigates the perceived limitations of AI-generated educational materials across dimensions including content quality, human con-tribution, prompt dependency, assessment challenges, credibility, and faculty awareness. The qualitative component complements the student evidence through structured interviews with faculty members representing different age groups and younger AI experts, focusing specifically on AI-generated evaluation, the possibility of AI-generated responses encouraging copying, the legitimacy of limiting AI use, academic integrity, and strategies for redesigning assessment practices. This integration provides both a learner-centred and an educator-centred perspective on the evolving role of GenAI in higher education.

The study is guided by the following research questions:

- RQ1: How do students perceive the effectiveness of AI-generated study materials in higher education?
- RQ2: What are the key limitations of AI-generated educational content as perceived by students?
- RQ3: Does academic maturity influence students' perceptions of AI-generated learning materials?
- RQ4: How do faculty members perceive the implications of AI-generated study materials for evaluation, academic integrity, and future assessment practices?

Based on the quantitative component of the study, the following hypotheses are proposed:

- H1: Students perceive AI-generated study materials as lacking depth and contextual under-standing.
- H2: AI-generated materials are perceived as insufficient in supporting higher-order learning.
- H3: Senior students exhibit greater skepticism toward AI-generated materials compared to junior students.

The study contributes to the growing literature on AI in higher education in three ways. First, it provides large-scale empirical evidence on student perceptions of AI-generated study materials across diverse undergraduate disciplines. Second, it examines academic maturity as a potential factor shaping students' evaluation of AI-generated learning resources. Third, it extends the analysis beyond student perceptions by incorporating faculty perspectives on AI-generated evaluation, academic integrity, student dependence on machine-generated responses, and the redesign of assessment practices. By combining these perspectives, the study seeks to clarify how AI can be incorporated into higher education without weakening independent reasoning, academic rigor, and meaningful human engagement.

A. Artificial Intelligence and Generative AI in Higher Education

The integration of Artificial Intelligence (AI) into higher education predates the recent emer-gence of generative AI and large language models (LLMs). Earlier applications primarily focused on intelligent tutoring systems, adaptive learning environments, automated assessment, learning analytics, and student-support systems, with the broader objective of personalizing instruction and improving learning efficiency[19][20]. A systematic review by Zawacki-Richter et al. identified several major applications of AI in higher education, including assessment and evalu-ation, adaptive systems, profiling and prediction, and intelligent tutoring, while also highlighting the comparatively limited involvement of educators in AI-related research and implementation[21]. This observation remains relevant to contemporary generative AI because successful adoption depends not only on technological capability but also on pedagogical decisions made by instructors and institutions. The emergence of ChatGPT and related LLM-based systems has substantially expanded the scope of AI-supported education. Unlike earlier educational AI systems that were generally designed for specific instructional tasks, generative AI can produce open-ended text, explanations, summaries, examples, questions, feedback, and problem-solving assistance through natural-language interaction[22]. Reviews of the emerging literature indicate that generative AI is being explored by students, teachers, researchers, and administrators for a diverse set of academic purposes, including content generation, language support, brain-storming, programming, tutoring, and assessment-related activities[23][24][25]. This versatility has increased the potential of AI to function as an on-demand academic resource, but it has also made questions concerning reliability, pedagogical appropriateness, and responsible use substantially more important. Recent systematic reviews indicate that research on ChatGPT in higher education has expanded rapidly but remains methodologically and conceptually heterogeneous. Baig and Yadegaridehkordi reviewed 57 studies published during 2023–2024 and found that applications of ChatGPT span multiple user groups and educational activities, with survey-based quantitative studies constituting a major proportion of the evidence base[26]. Similarly, Albadarin et al. identified a relatively small body of early empirical studies and noted the need for further research examining how ChatGPT affects actual educational practices rather than merely its potential applications[27]. These reviews suggest that, although the literature is growing rapidly, there remains a need for large-scale empirical research that integrates perceptions of both students and educators.

B. Educational Benefits and Opportunities of AI-Generated Learning Materials

One of the principal arguments for the use of generative AI in education is its ability to provide rapid and flexible access to instructional content. Students can request explanations tailored to their perceived level of understanding, obtain alternative formulations of concepts, receive examples, and interact iteratively with generated content[28]. Such capabilities may be particularly valuable in self-directed and digitally mediated learning environments, where immediate access to assistance can complement conventional classroom instruction. Empirical evidence provides some support for these potential benefits. A systematic review and meta-analysis of experimental studies reported that ChatGPT can positively influence academic perfor-mance, affective-motivational outcomes, and propensities toward higher-order thinking, although the technology may also reduce students' mental effort and does not necessarily improve self-efficacy[29]. This finding is important because educational effectiveness cannot be inferred simply from improved task performance. A student may obtain a correct answer more efficiently while engaging less deeply with the underlying reasoning process. Consequently, the distinction between task completion and learning is central to evaluating AI-generated educational materials. Student-perception studies similarly report a combination of benefits and concerns. Chan and Hu found that students recognized the usefulness of generative AI while simultaneously expressing concerns related to accuracy, ethical issues, and responsible use. Research examining student perspectives has also identified dimensions related to academic benefits, academic concerns, accessibility, and attitudes toward ChatGPT. These findings suggest that students do not per-ceive AI in exclusively positive or negative terms; rather, their evaluations are multidimensional and depend on how AI is incorporated into academic work. The literature therefore supports the view that generative AI can function effectively as a supplementary resource, particularly for explanation, ideation, language support, and personalized assistance. However, these benefits are conditional on students' ability to evaluate generated information and on appropriate pedagogical guidance from instructors. The same characteristics that make AI-generated content accessible and scalable may also make it attractive for superficial learning when students rely on generated answers without independently processing the information.

C. Content Quality, Accuracy, and Contextual Depth

The quality of AI-generated study materials is a major concern in the emerging literature. Unlike human instructors, LLMs do not possess experiential understanding of a subject in the same sense as domain experts; instead, they generate outputs based on patterns learned from large datasets. This distinction becomes important when educational tasks require nuanced in-terpretation, contextual judgment, disciplinary conventions, or integration of practical experience. A particularly important limitation is hallucination, in which language models produce incorrect or fabricated information while maintaining a fluent and confident style. Such outputs can be especially problematic in higher education because students may not possess sufficient subject expertise to distinguish reliable information from plausible-sounding errors. The risk extends beyond isolated factual mistakes: inaccurate explanations, unsupported generalizations, or fabricated references can influence students' subsequent reasoning and become embedded in their understanding of a topic. The literature also highlights the importance of contextualization. Generative AI can produce generalized explanations that may not adequately reflect the specific requirements of a discipline, institution, course, local context, or professional practice. Human instructors, by contrast, can adapt explanations to students' prior knowledge, interpret ambiguous questions, connect theory to practical experience, and emphasize disciplinary conventions. This difference provides an important theoretical basis for examining students' perceptions of content depth and contextual understanding in AI-generated study materials. The issue of credibility further reinforces this concern. AI-generated responses may present information without appropriate supporting evidence, citations, or reference to authoritative textbooks and scholarly sources. Students consequently require sufficient information literacy and AI literacy to verify generated outputs against reliable academic sources. The ability to generate fluent content should therefore not be equated with the ability to generate academically authoritative content.

D. The Human Element and Student Engagement

The pedagogical role of instructors extends beyond content delivery to include explanation, mentorship, interpretation, feedback, motivation, and the contextualization of knowledge. These dimensions are difficult to reduce to automated text generation because effective teaching is also relational and situational. The human element is therefore an important dimension when evaluating whether AI-generated study materials can serve as substitutes for conventional teaching. Research on student engagement reinforces this distinction. A systematic review of 72 empirical studies found evidence of both engagement and disengagement associated with ChatGPT-supported learning, including behavioral engagement with the technology as well as disengagement in the form of academic dishonesty[30]. This dual effect suggests that AI can increase students' interaction with learning resources while simultaneously creating conditions under which students may bypass effortful academic activities. The question is therefore not simply whether students use AI, but how AI changes the nature of their engagement with learning. When AI is used to obtain explanations, compare alternative perspectives, or receive formative feedback, it may support learning. When it is used primarily to generate finished answers, however, it may reduce the need for independent information processing. This distinction is particularly relevant to the present study because several questionnaire items address students' perceptions of the depth, human insight, engagement, and higher-order-learning capabilities of AI-generated materials.

E. Prompt Dependency and AI Literacy

Another emerging dimension of AI-supported learning concerns the dependence of output quality on the formulation of user prompts. Prompting is not merely a technical issue; it represents a new form of interaction between the learner and the educational technology. Liu et al. described prompting as a mechanism through which users influence the behavior of pretrained language models, with prompt formulation affecting task performance across a range of applications. In educational contexts, prompt dependency can create differences in the quality of learning support received by students. A student who provides vague or poorly structured instructions may obtain generic or incomplete material, whereas an AI-literate student may generate more targeted and useful explanations. Kasneci et al. consequently emphasize the need for educa-tors and learners to develop competencies that allow them to understand both the capabilities and limitations of LLMs. UNESCO similarly emphasizes AI literacy and human-centred approaches as essential conditions for responsible educational use of generative AI. Prompt dependency is therefore conceptually connected to the broader issue of educational equity and digital competence. Access to an AI system does not guarantee equivalent educational benefit because students vary in their ability to formulate effective prompts, critically evaluate responses, identify errors, and refine generated content. Faculty members may face similar challenges when attempting to use generative AI for instructional design, feedback, and assessment.

F. Generative AI, Academic Integrity, and Assessment

Among the most extensively discussed concerns surrounding generative AI in higher education is its impact on academic integrity. The ability of LLMs to generate coherent essays, solutions, explanations, and other forms of academic output creates new opportunities for students to outsource tasks traditionally used to demonstrate individual learning. The issue is not limited to conventional plagiarism because AI-generated content may be newly produced rather than copied from a single identifiable source. Consequently, existing approaches to plagiarism detection and academic misconduct are not necessarily sufficient for the generative-AI envi-ronment. Research has shown that ChatGPT can create substantial challenges for conventional assessment practices, particularly assessments based on unsupervised written responses. Farazouli et al. found that university teachers perceived the emergence of AI chatbots as a significant disruption to existing assessment practices and reported changes in assessment design in response to the technology. This reinforces the argument that the challenge presented by generative AI is fundamentally an assessment-design problem rather than only a technology-detection problem. Recent reviews of the assessment literature similarly indicate that generative AI is prompting a reconsideration of how learning should be evaluated. Chiu et al. reported that assessment transformation should place greater emphasis on self-regulated learning, responsible learning, and academic integrity[31]. The increasing availability of AI-generated responses creates pressure to design tasks that require contextual application, reasoning, interpretation, reflection, and demonstration of learning processes rather than relying exclusively on repro-ducible written outputs. This issue is directly relevant to the distinction between AI-generated study materials and AI-generated evaluation. If AI is used to generate learning materials and students subsequently use AI to generate responses to conventional assessments, the assessment may lose its ability to distinguish between authentic student learning and machine-generated performance. Conversely, AI may itself be used to support formative feedback and evaluation. Recent research examining GPT-based feedback generation indicates that large language models have potential for producing useful feedback on open-ended student writing, although questions concerning reliability, consistency, and appropriate human oversight remain. Thus, AI can simultaneously create assessment problems and provide components of a possible solution.

G. Faculty Perspectives and the Need for Assessment Redesign

Although much of the early discussion of educational AI focused on students, increasing attention has been directed toward educators' perspectives. Faculty members play a central role in determining how AI is incorporated into teaching, learning activities, and assessment. Their acceptance of AI cannot be assumed to be homogeneous because perceptions may vary according to experience, discipline, institutional support, and technological familiarity. Empirical research involving educators has demonstrated substantial variation in attitudes toward generative AI. A study of higher education educators found no homogeneous sentiment regarding AI and reported ambiguity concerning appropriate practices. At the same time, many participants were already incorporating AI into their teaching and modifying assessments, while a substantial proportion reported a need for greater institutional support. These findings suggest that educators are not responding to generative AI through a single strategy of either acceptance or rejection. The comparison between educator and student perspectives is particularly important for assessment. A cross-national study involving university educators and students reported that both groups perceived generative AI as affecting assessment, while educators were more likely than students to support adapted assessment approaches incorporating generative AI. Such findings suggest that the educational response to GenAI is increasingly shifting from the question of whether AI should be allowed toward the question of how assessment should be redesigned. Faculty experience may also influence the balance between restriction and adaptation. Some educators emphasize controlled environments, supervised examinations, and limitations on AI use to preserve academic rigor. Others argue that technological disruption cannot realistically be reversed and therefore advocate more authentic, analytical, and application-oriented forms of assessment. This tension is especially relevant in the context of undergraduate education, where institutions must simultaneously protect academic integrity and prepare students to operate in workplaces where AI-enabled tools are increasingly prevalent.

H. Academic Maturity and Differences in Perception

An additional dimension that remains relatively underexplored is the role of academic ma-turity in shaping perceptions of AI-generated learning materials. Students at different stages of undergraduate education differ in their exposure to disciplinary knowledge, conventional teaching practices, assessment formats, and independent learning expectations. These differences may in-fluence both their perceived usefulness of AI and their ability to critically evaluate its limitations. Existing empirical studies have examined variables such as perceived usefulness, performance expectancy, academic integrity, gender, personal innovativeness, and study year in relation to ChatGPT adoption and use. However, much of the literature has emphasized technology acceptance or usage intention rather than whether academic experience changes students' critical evaluation of AI-generated educational content. There is therefore a need to examine whether students with greater exposure to conventional instructor-led education develop more critical perceptions of AI-generated materials. The present study addresses this issue by comparing perceptions across undergraduate academic years. This perspective complements existing research because it considers academic maturity not merely as a demographic characteristic but as a possible explanatory factor in how students judge the depth, credibility, contextual quality, and instructional value of AI-generated materials.

I. Research Gap

The literature demonstrates that generative AI offers substantial opportunities for accessibil-ity, personalization, content generation, and learning support, while simultaneously introducing concerns related to accuracy, contextual depth, human interaction, prompt dependency, aca-demic integrity, and assessment validity. Systematic reviews further show that research in higher education has grown rapidly but remains concentrated in relatively early-stage investigations of use, perceptions, adoption, and individual applications. Four gaps are particularly relevant to the present study. First, although numerous studies investigate ChatGPT adoption and general attitudes, fewer large-scale empirical studies examine students' perceptions of AI-generated study materials across diverse undergraduate disciplines. Second, the multidimensional nature of perceived limitations—including content quality, human contribution, prompt dependency, credibility, and assessment-related concerns—requires further empirical val-idation. Third, the role of academic maturity in shaping perceptions of AI-generated educational resources remains comparatively underexplored. Fourth, although assessment disruption and academic integrity are widely discussed, fewer studies explicitly connect student perceptions of AI-generated learning materials with faculty perspectives on AI-assisted evaluation, student dependence on generated responses, copying behavior, and the redesign of assessment. The present study addresses these gaps through a mixed-method approach combining a large-scale survey of 1,164 undergraduate students with structured interviews involving faculty from different age groups and younger AI experts. The quantitative component investigates the underlying dimensions of student perceptions and their variation across academic years, while the qualitative component examines faculty views concerning AI-generated evaluation, academic integrity, stu-dent copying, restrictions on AI use, and future assessment strategies. The resulting framework provides a more comprehensive perspective on the role of AI-generated study materials by connecting the learner's perception of AI with the educator's challenge of evaluating authentic learning.

II. METHODOLOGY

A. Research Design

The study adopted a mixed-method research design to examine the perceived effectiveness and limitations of AI-generated study materials in higher education and to investigate faculty perspectives on the implications of generative AI for evaluation and assessment. The quantitative component focused on undergraduate students' perceptions of AI-generated educational mate-rials, whereas the qualitative component explored faculty perspectives concerning AI-generated evaluation, AI-generated student responses, copying behavior, the need for restricting AI use, and future assessment strategies. The quantitative and qualitative components were designed to provide complementary perspectives on the changing role of generative AI in higher education. The student survey was used to identify measurable patterns in perceptions and to examine the underlying dimensions associated with AI-generated learning materials. The faculty inter-views were used to explore the practical and pedagogical implications of these perceptions, particularly with respect to academic integrity and the redesign of assessment. The overall methodological framework therefore integrates learner perceptions with educator perspectives to provide a broader understanding of the opportunities and limitations associated with AI-generated educational resources.

B. Student Sample and Data Collection

The quantitative component included a total of 1,164 undergraduate students enrolled in private universities across India. The sample included students from diverse academic disciplines and undergraduate programs, including B.Tech, BBA, BA Psychology, English, Physiotherapy, Nursing, BBA LL.B, and Law. Students from different stages of undergraduate education, ranging from first year to fourth year, were included to facilitate comparison according to academic maturity. The sampling design incorporated diversity across gender, academic discipline, and academic year. A stratified sampling approach was used to obtain representation across gender categories. The inclusion of students from different academic years enabled the study to examine whether prolonged exposure to conventional higher education and teacher-led instruction was associated with differences in perceptions of AI-generated educational materials. Data were collected using a structured questionnaire administered to the participating students. Participation was voluntary, and respondents were informed about the purpose of the study. Responses were collected anonymously to protect participant confidentiality and reduce potential response bias.

C. Research Instrument for Student Survey

A structured 21-item questionnaire was developed to measure students' perceptions of AI-generated educational materials. Each item was measured using a five-point Likert scale ranging from 1 = Strongly Disagree to 5 = Strongly Agree.

The questionnaire covered seven conceptual dimensions:

1) Content Quality and Depth (Q1–Q5)
2) Human Element and Teaching Quality (Q6–Q9)
3) Prompt Dependency (Q10–Q12)
4) Assessment and Learning Challenges (Q13–Q14)
5) Credibility and Supporting Evidence (Q15–Q17)
6) Faculty Awareness and Adoption (Q18–Q19)
7) Overall Perception (Q20–Q21)

The items were formulated to capture commonly reported concerns associated with AI-generated educational content, including insufficient contextual depth, limited experiential in-sight, prompt dependency, assessment challenges, credibility concerns, and the role of faculty in AI-supported education.

D. Faculty Interview Component

To complement the quantitative findings, a structured interview component was conducted with faculty members representing different age groups, together with younger professionals having expertise in Artificial Intelligence. The interview component was designed to obtain educator perspectives on the consequences of AI-generated study materials for evaluation, academic integrity, student learning behavior, and future assessment practices.

The structured interview consisted of five questions:

1) Should AI-generated study materials also have AI-generated evaluation?
2) Will AI-generated evaluation encourage AI-generated responses?
3) Will such practice encourage students to simply copy responses from AI?
4) Is this concern a valid reason for limiting the use of AI-generated materials?
5) What steps should be taken going forward?

The questions were designed to examine the relationship between AI-generated learning resources and AI-assisted assessment, with particular emphasis on whether the increasing use of generative AI may encourage students to outsource cognitive effort or reproduce machine-generated responses. The final question was intentionally open-ended to capture faculty rec-ommendations regarding institutional and pedagogical responses to the changing assessment environment.

For qualitative interpretation, responses were organized thematically around five broad areas:

(i) the appropriate role of AI-generated evaluation, (ii) the prevalence and consequences of AI-generated student responses, (iii) copying and academic integrity, (iv) the desirability of restricting AI use, and (v) strategies for future assessment design.

E. Qualitative Analysis

The faculty interview responses were analyzed using thematic categorization. Responses were first organized according to the five interview questions and subsequently compared across the identified faculty age groups. Particular attention was given to areas of convergence and divergence between the responses. Two broad perspectives emerged from the interview material. Faculty in the 45–50+

age group emphasized controlled use of AI, retention of subjective evaluation, concerns regarding declining academic rigor and increased copying, and greater use of practical and case-based classroom examinations. Faculty in the 30–45 age group placed greater emphasis on adapting to technological disruption, using AI-supported checking mechanisms, permitting AI in selected open-book contexts, and increasing the analytical and practical difficulty of assessment tasks. The qualitative component was not intended to statistically generalize faculty opinion. Rather, it was used to contextualize and extend the quantitative findings by identifying practical concerns and alternative approaches to assessment in an environment where AI-generated content and AI-generated responses are increasingly accessible.

F. Data Analysis Techniques

The quantitative data were analyzed using descriptive and inferential statistical techniques. Descriptive statistics, including mean and standard deviation, were computed for all 21 questionnaire items to summarize the central tendency and dispersion of student responses. Exploratory factor analysis (EFA) was conducted to examine the underlying structure of the questionnaire. Principal component extraction with Varimax rotation was employed to obtain an interpretable factor structure. Factors were retained on the basis of eigenvalues greater than one and their theoretical alignment with the conceptual dimensions of the questionnaire. Factor loadings were examined to identify the association between individual questionnaire items and the extracted factors. Reliability analysis was performed using Cronbach's Alpha. The overall internal consistency of the 21-item instrument was assessed together with the reliability of the individual conceptual constructs. The analysis was used to determine the consistency of the measurement instrument and to identify constructs requiring cautious interpretation because of their limited number of constituent items. In addition, comparisons were conducted across academic years to examine differences in student perceptions according to academic maturity. Particular attention was given to the contrast between students in earlier years of undergraduate study and students in later years. The qualitative findings from the faculty interviews were subsequently considered alongside the quantitative findings. This integration was used to identify points of convergence, particularly concerning the limitations of AI-generated materials and higher-order learning, as well as points of divergence concerning whether AI should be restricted or whether assessment practices should instead be redesigned.

G. Validity and Reliability Considerations

Construct validity of the student questionnaire was examined through factor analysis. The extracted factor structure was interpreted in relation to the conceptual dimensions developed for the study, including content quality, human contribution, prompt dependency, assessment, credibility, and faculty awareness. The overall internal consistency of the instrument was eval-uated using Cronbach's Alpha. The overall scale demonstrated high internal consistency, while lower coefficients were observed for some individual constructs. Because Cronbach's Alpha is sensitive to the number of items included within a construct, constructs represented by only two or three questionnaire items were interpreted with appropriate caution. The qualitative component was used primarily for contextual and explanatory purposes rather than statistical validation of the quantitative instrument. The interview themes were examined in relation to the quantitative findings to provide complementary evidence concerning assessment, academic integrity, and the appropriate role of AI in higher education.

H. Ethical Considerations

Participation in the study was voluntary. Student respondents were informed of the purpose of the research, and responses were collected anonymously. Confidentiality was maintained through-out the data collection and analysis process. The faculty interview component was conducted for academic research purposes, and the responses were considered at the group/theme level rather than being attributed to individual participants. No personally identifying information is reported in the qualitative findings. The study followed standard ethical practices for research involving human participants and maintained confidentiality in the handling and presentation of participant responses.

The overall research methodology adopted in this study is illustrated in Fig. 1. The frame-work integrates two complementary components: a quantitative student-survey component and a qualitative faculty-interview component. The quantitative component examines undergraduate students' perceptions of AI-generated study materials using a structured 21-item Likert-scale questionnaire, followed by descriptive statistical analysis, exploratory factor analysis, reliability analysis, and comparison across academic years. The qualitative component captures faculty perspectives through structured interviews focusing on AI-generated evaluation, AI-generated student responses, copying behavior, the appropriateness of limiting AI use, and future assess-ment practices. The findings from both components are subsequently integrated to provide a comprehensive understanding of the educational opportunities, limitations, academic-integrity concerns, and assessment implications associated with AI-generated study materials.

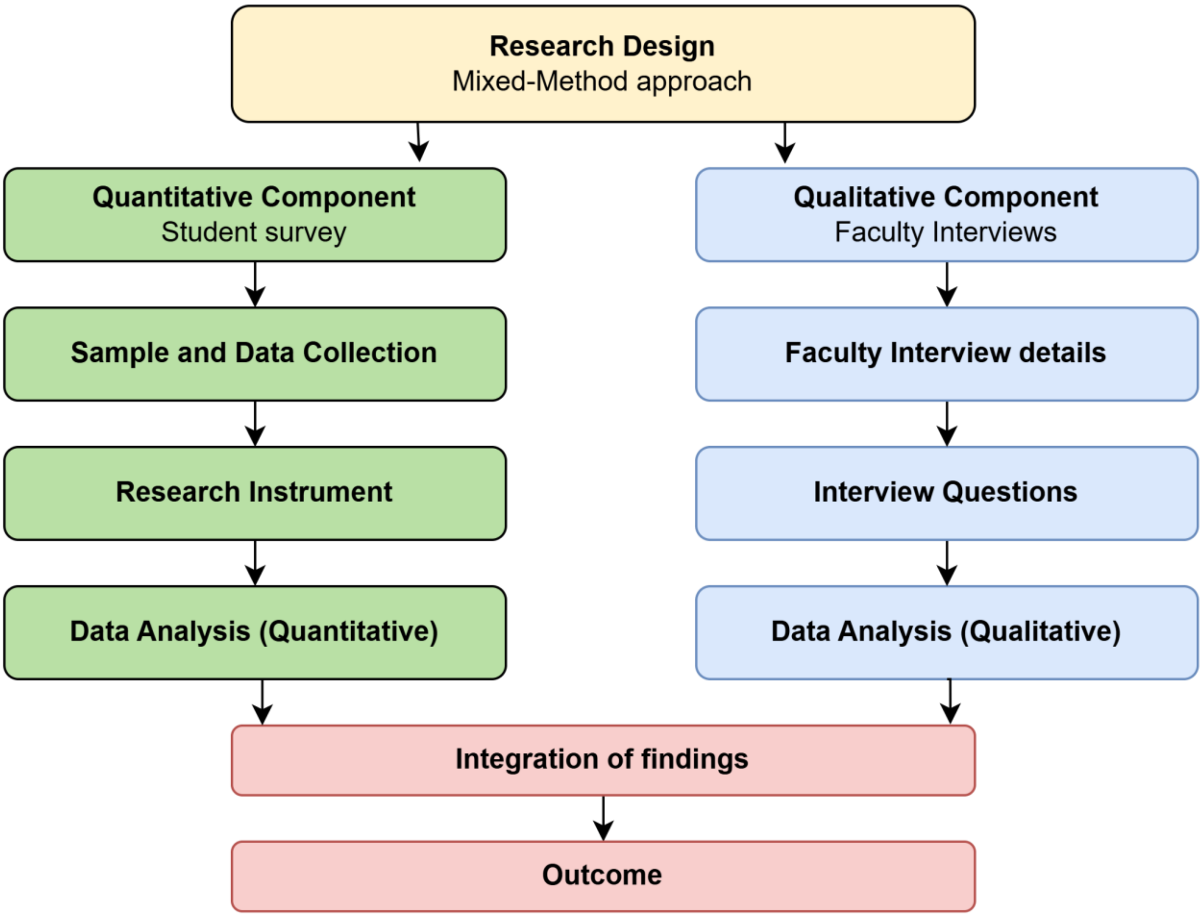

Fig. 1. Mixed-method research framework integrating quantitative student perceptions and qualitative faculty perspectives on AI-generated study materials, evaluation, and assessment practices.

III. RESULTS AND DISCUSSION

A. Descriptive Statistics

Table 1. Descriptive Statistics (Q1–Q10)

Item	Mean	SD
Q1	3.16	1.12
Q2	3.16	1.16
Q3	3.17	1.13
Q4	3.19	1.13
Q5	3.21	1.13
Q6	3.20	1.16

Item	Mean	SD
Q7	3.17	1.11
Q8	3.16	1.14
Q9	3.16	1.13
Q10	3.18	1.11

Table 2. Descriptive Statistics (Q11–Q21)

Item	Mean	SD
Q11	3.17	1.12
Q12	3.20	1.15
Q13	3.23	1.14
Q14	3.18	1.12
Q15	3.13	1.16
Q16	3.17	1.13
Q17	3.19	1.12
Q18	3.17	1.13
Q19	3.16	1.12
Q20	3.18	1.12
Q21	3.12	1.11

Tables I and II present the descriptive statistics for all 21 questionnaire items. The mean values range from 3.12 to 3.23, indicating a moderate level of agreement among respondents regarding the limitations of AI-generated educational materials. The standard deviation values range from 1.10 to 1.16, indicating a moderate dispersion of responses across the questionnaire items.

The highest mean score is observed for Q13 (3.23), indicating the strongest agreement among the questionnaire items with the statement that AI-generated materials make open-ended or critical evaluation more difficult. Relatively high mean scores are also observed for Q5 (3.21), Q6 (3.20), and Q12 (3.20), which relate to the depth of AI-generated explanations, human insight, and the influence of prompt formulation, respectively. These results indicate that students perceive AI-generated materials as useful but recognize limitations associated with depth, contextualization, human experience, and the conditions under which AI-generated content is produced.

The lowest mean score is observed for Q21 (3.12), which concerns students' preference for teacher-created materials over AI-generated materials. Although this remains above the neutral midpoint of the five-point scale, the relatively lower value suggests that students' perceptions are not uniformly negative toward AI-generated resources. Instead, the responses indicate a moderate and differentiated assessment of their usefulness and limitations.

B. Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) was conducted to examine the underlying structure of the 21-item questionnaire. The sample size of N = 1164 was considered adequate for multivariate analysis. Principal component extraction with Varimax rotation was employed to obtain an interpretable factor structure. Seven factors were retained based on eigenvalues greater than one and their conceptual alignment with the dimensions examined in the study.

Table III presents the rotated factor loadings. Loadings greater than 0.50 were treated as relatively strong associations for interpretation. The resulting structure indicates that perceptions of AI-generated educational materials are multidimensional rather than being represented by a single underlying construct.

The factor structure broadly reflects several theoretically meaningful dimensions of student perceptions. The first factor is associated primarily with concerns regarding content reliability and accuracy; the second captures assessment and credibility concerns; the third reflects higher-order learning limitations; the fourth represents content depth and contextual considerations; the fifth reflects prompt dependency; the sixth represents deficiencies in the human element; and the seventh relates to faculty awareness and practical exposure.

The extracted structure therefore provides empirical support for treating perceptions of AI-generated educational materials as multidimensional. In particular, the results indicate that stu-dents distinguish between the quality and reliability of generated content, the contribution of human instruction, prompt-related limitations, assessment implications, and faculty preparedness.

Table 3. Rotated Factor Loadings (Varimax Rotation)

Item	F1	F2	F3	F4	F5	F6	F7
Q1				0.58			
Q2	0.42						
Q3						0.34	
Q4			0.40				
Q5						0.40	
Q6						0.58	
Q7				0.36	-0.52		
Q8							0.55
Q10	0.47						
Q11					-0.69		
Q12				0.49			
Q13		0.62					
Q14			0.49				
Q15	0.49						
Q16						0.43	
Q17		0.53					
Q18							0.57
Q20			0.61				
Q21							0.39

C. Variance Explained

Table IV presents the proportion of variance associated with the seven extracted factors. The seven-factor solution accounts for approximately 59.1% of the total variance. Factor 1 accounts for 37.6% of the variance, while the remaining six factors contribute between 3.4% and 3.8% each. The cumulative variance reaches 59.1% after the seventh factor.

The cumulative variance indicates that the extracted factors capture a substantial propor-tion of the observed variation in student responses. The dominance of Factor 1 suggests that content-related concerns represent an important component of students' overall evaluation of AI-generated study materials, while the remaining factors capture additional pedagogical, as-sessment, and implementation-related dimensions.

D. Interpretation of Factors

The rotated factor structure reveals the following conceptual groupings:

Table 4. Total Variance Explained

Factor	Proportion of Variance	Cumulative Variance
Factor 1	0.376	0.376
Factor 2	0.038	0.414

Factor	Proportion of Variance	Cumulative Variance
Factor 3	0.037	0.452
Factor 4	0.036	0.488
Factor 5	0.035	0.523
Factor 6	0.035	0.558
Factor 7	0.034	0.591

- **Factor 1: Content Reliability and Accuracy** – primarily associated with Q2, Q10, and Q15.
- **Factor 2: Assessment and Credibility Concerns** – primarily associated with Q13 and Q17.
- **Factor 3: Higher-Order Learning Limitations** – primarily associated with Q14 and Q20.
- **Factor 4: Content Depth and Context** – primarily associated with Q1 and Q12.
- **Factor 5: Prompt Dependency** – primarily associated with Q11 and Q7.
- **Factor 6: Human Element Deficiency** – primarily associated with Q5, Q6, and Q16.
- **Factor 7: Faculty Awareness and Practical Exposure** – primarily associated with Q8, Q18, and Q21.

Taken together, the factor solution indicates that students' perceptions of AI-generated study materials encompass multiple dimensions extending beyond simple acceptance or rejection of the technology. These dimensions include the reliability and depth of generated content, the contribution of human instructors, prompt-related variability, assessment challenges, credibility, and faculty readiness.

E. Reliability Analysis

To assess the internal consistency of the measurement instrument, Cronbach's Alpha was calculated for the overall questionnaire and for the individual conceptual constructs.

Table 5. Reliability Analysis (Cronbach's Alpha)

Construct	Cronbach's Alpha
Overall Scale	0.917
Content Quality	0.738
Human Element	0.678
Prompt Dependency	0.616
Assessment	0.505
Credibility	0.616
Awareness	0.475
Overall Perception	0.488

The overall questionnaire demonstrates high internal consistency, with a Cronbach's Alpha of 0.917 across the 21 items. Content Quality shows the highest construct-level reliability at 0.738, followed by Human Element at 0.678 and Prompt Dependency and Credibility at 0.616 each. Lower coefficients are observed for Assessment (0.505), Awareness (0.475), and Overall Perception (0.488).

The lower construct-level coefficients should be interpreted cautiously because several con-structs contain only a small number of items. Cronbach's Alpha is sensitive to scale length, and constructs represented by two or three items may produce comparatively lower coefficients even when the items address a related conceptual domain. Accordingly, the strong overall reliability supports the internal consistency of the complete instrument, whereas individual construct-level results should be interpreted with appropriate caution.

F. Graphical Analysis

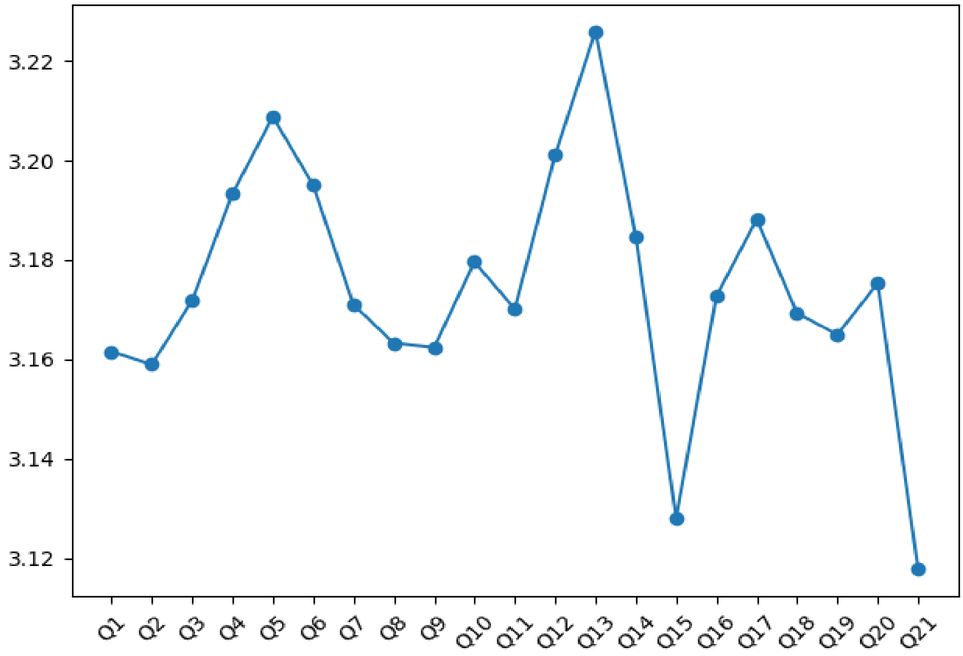

Fig. 2. Mean Scores of Questionnaire Items

Figure 2 presents the mean scores across all questionnaire items. The relatively narrow range of means indicates that student responses are broadly distributed around the moderate-agreement level. Q13 represents the highest mean score, while Q21 represents the lowest. The relatively uniform distribution suggests that concerns regarding AI-generated educational materials are not restricted to a single dimension but occur across content, human contribution, prompt dependency, assessment, credibility, and faculty-related issues.

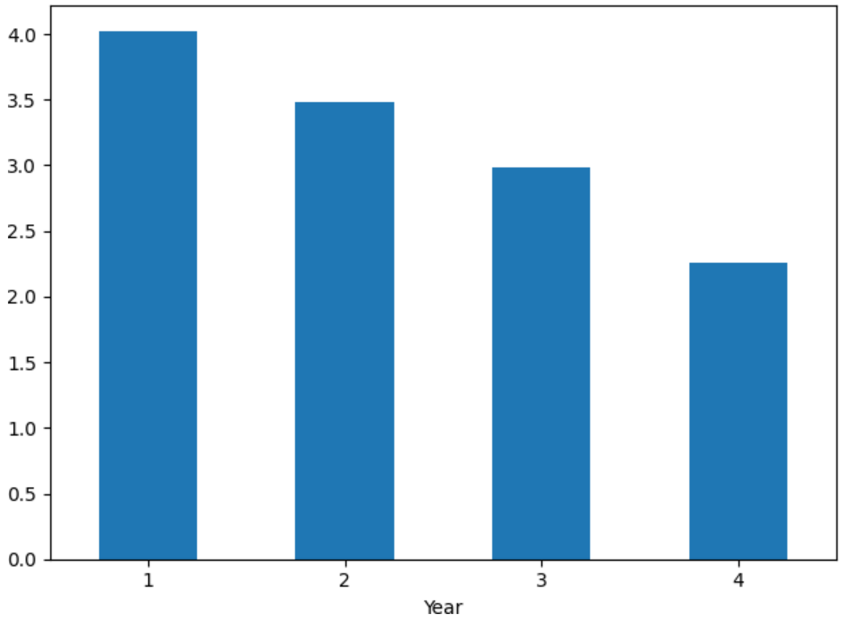

Fig. 3. Year-wise Comparison of Student Perceptions

Figure 3 presents the year-wise comparison of student perceptions. The results indicate a progressive difference in overall perception across academic years. First-year students exhibit the most favorable perception, whereas students in later academic years demonstrate increasingly skeptical perceptions of AI-generated study materials. The fourth-year group exhibits the lowest overall perception score in the comparison.

The year-wise pattern suggests that academic maturity may be associated with greater critical evaluation of AI-generated educational resources. Students with greater exposure to conventional instructor-led learning may therefore be more likely to identify limitations relating to contextualization, depth, experiential knowledge, and the authenticity of AI-generated learning support.

G. Qualitative Findings from Faculty Interviews

To complement the quantitative findings, structured interviews were conducted with faculty members from different age groups. The interviews focused on five issues: the use of AI-generated evaluation alongside AI-generated study materials, the possibility that AI-generated evaluation could encourage AI-generated responses, the tendency of students to copy AI-generated responses, whether such concerns justify restricting AI use, and the measures required for future educational practice.

The qualitative responses revealed both areas of convergence and differences in emphasis between the faculty groups. The responses were organized into five themes: (i) the role of AI-generated evaluation, (ii) AI-generated student responses and academic rigor, (iii) copying and academic integrity, (iv) restriction versus adaptation of AI use, and (v) redesign of assessment practices.

Table 6. Thematic Summary Of Faculty Interview Responses

Theme	45–50+ Faculty	30–45 Faculty
AI-generated evaluation	Limited use alongside subjective evaluation	Can support ethical examination and reduce paper-leak risks
AI-generated responses	Increasingly predominant and potentially reducing academic rigor	Can potentially be checked using AI-based tools
Student copying through responses	AI access has made copying easier	Challenge should be addressed through technological and assessment measures
AI restriction realism	Supports limiting AI use	Technological disruption cannot realistically be limited
Future assessment	Practical and case-based classroom examinations under controlled conditions	Higher-difficulty problems, classroom problem solving, data analysis, and interpretation

1) AI-Generated Evaluation: Faculty members in the 45–50+ age group expressed the view that AI-generated evaluation should be used only as a limited option and should be accompanied by subjective or human evaluation. Their responses emphasized the continuing importance of human judgment in evaluating student learning.

Faculty members in the 30–45 age group adopted a more technology-oriented perspective. They indicated that AI-generated evaluation could contribute to more ethical examination pro-cesses and potentially reduce risks associated with examination paper leaks. Thus, while both groups recognized a possible role for AI in evaluation, they differed in the degree of reliance they considered appropriate.

2) AI-Generated Responses and Academic Rigor: The 45–50+ faculty group expressed con-cern that AI-generated responses are becoming increasingly predominant among undergraduate students. According to these responses, students may increasingly obtain answers directly from AI systems rather than processing and interpreting information independently. The faculty group associated this behavior with a potential reduction in the academic rigor of undergraduate study. The 30–45 faculty group acknowledged the prevalence of AI-generated responses but empha-sized the possibility of using AI-based tools to examine or check AI-generated student responses.

Their perspective therefore focused more strongly on adapting evaluation mechanisms to the technological environment rather than attempting to eliminate AI-generated responses entirely.

3) Copying and Academic Integrity: Both faculty groups identified copying as an important concern associated with widespread access to generative AI. Faculty in the 45–50+ group specifically reported that the ease of obtaining AI-generated answers has made copying easier and has contributed to concerns regarding students' independent engagement with academic material.

The responses therefore complement the quantitative finding that students perceive limitations in open-ended and critical evaluation. The qualitative evidence suggests that the concern extends beyond the generation of learning materials to the possibility that students may use the same technology to generate or reproduce responses in assessed activities.

4) Restriction Versus Adaptation of AI Use: A notable difference emerged between the faculty groups regarding whether AI use should be restricted. The 45–50+ faculty group considered limiting the use of AI-generated materials to be an appropriate response to concerns regarding copying and declining academic rigor.

In contrast, the 30–45 faculty group emphasized that technological disruption cannot realis-tically be limited. Their responses favored adapting educational practices to the presence of AI rather than attempting to prevent its use altogether.

This divergence indicates that faculty responses are not characterized by a simple acceptance–rejection distinction. Instead, they reflect two broad approaches: a restriction-oriented approach, emphasizing controlled AI use and preservation of conventional academic rigor, and an adaptation-oriented approach, emphasizing institutional and assessment redesign in response to an increas-ingly AI-enabled educational environment.

5) Assessment Redesign and Future Practice: The strongest area of convergence across the faculty responses concerns the need to modify assessment practices. Faculty in the 45–50+ group recommended greater use of practical and case-based examinations conducted in controlled classroom environments, including situations in which students are required to leave their phones outside the examination setting. The objective is to create assessment situations in which students must independently apply knowledge rather than reproduce AI-generated responses.

Faculty in the 30–45 group similarly recommended increasing the difficulty and application-orientation of assessment tasks. Their suggestions included classroom-based problem solving, data analysis, and interpretation. They also indicated that open-book examinations could incor-porate AI when the assessment is designed to evaluate the student's ability to use and interpret information rather than merely reproduce an answer.

Taken together, the responses indicate a common movement away from assessments that can be easily completed through direct AI-generated responses and toward tasks requiring contextual application, analytical reasoning, interpretation, and practical problem solving.

H. Integration of Quantitative and Qualitative Findings

The quantitative and qualitative findings provide complementary perspectives on the role of AI-generated educational materials in higher education. The student survey indicates moderate agreement that AI-generated materials have limitations related to content depth, human insight, prompt dependency, higher-order learning, assessment, and credibility. The qualitative faculty responses extend these findings by illustrating how such limitations may influence actual assess-ment practices and student behavior.

A particularly important point of convergence concerns higher-order learning and independent reasoning. Students report limitations in the ability of AI-generated materials to support critical and open-ended evaluation, while faculty members express concern that direct reliance on AI-generated responses may reduce students' independent processing of information. These findings collectively suggest that the educational challenge is not merely the availability of AI-generated content but the possibility that excessive reliance on generated responses may reduce opportunities for students to demonstrate independent reasoning.

The two data sources also converge on the importance of human involvement. The quantitative findings identify limitations in human insight, experiential knowledge, and contextualization, while the older faculty group emphasizes the continued importance of subjective evaluation and controlled examination conditions. This supports the interpretation that AI is more appropriately positioned as a support mechanism within a broader human-led educational framework.

At the same time, the qualitative findings reveal an important divergence in how faculty mem-bers believe institutions should respond to AI. The 45–50+ faculty group emphasizes restricting or controlling AI use, whereas the 30–45 group emphasizes adaptation and assessment redesign. Despite this difference, both perspectives converge on the need to strengthen assessment practices so that students are required to demonstrate understanding, interpretation, and application rather than merely reproduce generated responses.

The combined findings therefore point toward a hybrid approach in which AI-generated materials may be used to supplement learning while assessment increasingly emphasizes practical application, case analysis, data interpretation, problem solving, and other forms of learning that require meaningful student engagement.

I. Key Findings

The findings of the study can be summarized across five principal observations. First, students demonstrate a moderate level of agreement regarding the limitations of AI-generated educational materials, with the strongest concern observed for open-ended and critical evaluation. Second, the EFA indicates that student perceptions are multidimensional, encompassing content reliability, assessment and credibility, higher-order learning, content depth, prompt dependency, human contribution, and faculty awareness. Third, the overall questionnaire demonstrates high internal consistency, although several short constructs exhibit lower individual reliability and should therefore be interpreted cautiously.

Fourth, the year-wise comparison indicates that academic maturity is associated with in-creasingly skeptical perceptions of AI-generated educational materials, with final-year students demonstrating less favorable perceptions than students in earlier academic years. Fifth, the faculty interviews reveal that concerns regarding AI extend from the generation of study materials to the evaluation of student learning. Faculty members differ in their preferred response—with some emphasizing restrictions and controlled assessment and others emphasizing technologi-cal adaptation—but both groups support greater use of practical, analytical, case-based, and application-oriented assessment.

Overall, the integrated findings suggest that the central challenge is not whether AI should be completely accepted or rejected, but how its use can be structured so that accessibility and technological assistance do not displace independent reasoning, human judgment, contextual learning, and authentic assessment.

IV. CONCLUSION

This study examined the perceived effectiveness and limitations of AI-generated study mate-rials in higher education by integrating quantitative evidence from 1,164 undergraduate students with qualitative perspectives from faculty members. The findings indicate that, although AI-generated materials provide important benefits in terms of accessibility, flexibility, and rapid content generation, students perceive meaningful limitations related to content depth, contextual understanding, human insight, credibility, prompt dependency, and support for higher-order learning. The exploratory factor analysis further indicates that these perceptions are multidi-mensional, encompassing content-related, human, prompt-related, assessment, credibility, and faculty-awareness dimensions. The overall reliability of the 21-item instrument was high, while the comparatively lower reliability of some short individual constructs suggests the need for further refinement of the measurement instrument in future research. A key contribution of the quantitative component is the identification of academic maturity as an important dimension of student perception. Senior students demonstrate greater skepticism toward AI-generated study materials than students in earlier academic years. This finding suggests that greater exposure to conventional teacher-led education and disciplinary learning may be associated with a more critical evaluation of AI-generated educational resources. The results therefore indicate that student acceptance of AI should not be interpreted as uniform across academic stages. The qualitative faculty findings extend the student-based results by demonstrating that con-cerns surrounding AI-generated study materials also have direct implications for evaluation and academic integrity. Faculty members in the 45–50+ age group emphasized that AI-generated evaluation should be used only as a limited option alongside subjective evaluation. They also expressed concern that the increasing availability of AI-generated responses may reduce academic rigor and make copying easier for undergraduate students. Their proposed response emphasized greater control over AI use and practical, case-based examinations conducted in classroom environments where students are required to work independently. Faculty members in the 30–45 age group adopted a comparatively more adaptive perspective. They suggested that AI-generated evaluation could contribute to more ethical examination processes and help address issues such as examination paper leaks. They also indicated that AI-generated responses could potentially be checked using AI-based tools and argued that technological disruption cannot realistically be limited. Instead, they recommended increasing the difficulty and application-orientation of assessment through classroom problem solving, data analysis, and interpretation.

REFERENCES

[1] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – Where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, Art. no. 39, 2019.

[2] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *IEEE Access*, vol. 8, pp. 75,264–75,278, 2020.

[3] E. Kasneci, K. Sessler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, E. Hüllermeier, S. Krusche, G. Kutyniok, T. Michaeli, and C. Nerdel, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learning and Individual Differences*, vol. 103, Art. no. 102274, 2023.

[4] C. K. Lo, "What is the impact of ChatGPT on education? A rapid review of the literature," *Education Sciences*, vol. 13, no. 4, Art. no. 410, 2023.

[5] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, Art. no. 43, 2023.

[6] H. Singh, M.-H. Tayarani-Najaran, and M. Yaqoob, "Exploring computer science students' perception of ChatGPT in higher education: A descriptive and correlation study," *Education Sciences*, vol. 13, no. 9, Art. no. 924, 2023.

[7] R. Deng, M. Jiang, X. Yu, Y. Lu, and S. Liu, "Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies," *Computers & Education*, 2024.

[8] W. Holmes, K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. Buckingham Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, I. I. Bittencourt, and K. R. Koedinger, "Ethics of AI in education: Towards a community-wide framework," *International Journal of Artificial Intelligence in Education*, vol. 32, no. 3, pp. 504–526, 2022.

[9] UNESCO, *Guidance for Generative AI in Education and Research*. Paris, France: United Nations Educational, Scientific and Cultural Organization, 2023.

[10] E. M. Bender, T. Gebru, A. McMillan-Major, and M. Mitchell, "On the dangers of stochastic parrots: Can language models be too big?" in *Proc. 2021 ACM Conf. Fairness, Accountability, and Transparency*, 2021, pp. 610–623.

[11] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, "Survey of hallucination in natural language generation," *ACM Computing Surveys*, vol. 55, no. 12, pp. 1–38, 2023.

[12] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," *ACM Computing Surveys*, vol. 55, no. 9, pp. 1–35, 2023.

[13] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, “Chatting and cheating: Ensuring academic integrity in the era of ChatGPT,” *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228–239, 2024.

[14] T. Susnjak, “ChatGPT: The end of online exam integrity?” *arXiv preprint arXiv:2212.09292*, 2022.

[15] A. Farazouli, T. Cerratto-Pargman, K. Bolander-Laksov, and C. McGrath, “Hello GPT! Goodbye home examination? An exploratory study of AI chatbots’ impact on university teachers’ assessment practices,” *Assessment & Evaluation in Higher Education*, vol. 49, no. 3, pp. 363–375, 2024.

[16] I. Fischer, S. Sweeney, M. Lucas, and N. Gupta, “Making sense of generative AI for assessments: Contrasting student claims and assessor evaluations,” *The International Journal of Management Education*, vol. 22, no. 3, Art. no. 101081, 2024.

[17] O. Wang, “Assessing the proficiency of large language models in automatic feedback generation: An evaluation study,” *Computers and Education: Artificial Intelligence*, vol. 7, Art. no. 100299, 2024.

[18] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena, “AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations,” *Minds and Machines*, vol. 28, no. 4, pp. 689–707, 2018.

[19] B. P. Woolf, *Building Intelligent Interactive Tutors*. Burlington, MA, USA: Morgan Kaufmann, 2010.

[20] G.-J. Hwang, H. Xie, B. W. Wah, and D. Gašević, “Vision, challenges, roles and research issues of artificial intelligence in education,” *Computers and Education: Artificial Intelligence*, vol. 1, Art. no. 100001, 2020.

[21] A. Dahal, O. P. Sarker, J. Kashyap, R. K. Gupta, S. S. Chaudhuri, and S. Moulik, “Detection and classification of waste materials using deep learning techniques,” in *Proc. Int. Conf. Machine Learning, Image Processing, Network Security and Data Sciences*. Cham, Switzerland: Springer, 2023, pp. 75–85.

[22] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, and R. Huang, “What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education,” *Smart Learning Environments*, vol. 10, no. 1, Art. no. 15, 2023.

[23] M. I. Baig and E. Yadegaridehkordi, “ChatGPT in the higher education: A systematic literature review and research challenges,” *International Journal of Educational Research*, vol. 127, Art. no. 102411, 2024.

[24] B. Samanta, S. Banerjee, R. K. Gupta, S. Moulik, and S. S. Chaudhuri, “In-depth survey of automated shoreline alteration identification methods,” in *Proc. 2024 5th Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. IEEE, 2024, pp. 61–67.

[25] Y. Albadarin, M. Saqr, N. Pope, and M. Tukiainen, “A systematic literature review of empirical research on ChatGPT in education,” *Discover Education*, vol. 3, no. 1, Art. no. 60, 2024.

[26] R. K. Gupta, S. Moulik, and S. S. Chaudhuri, “High-resolution mangrove land cover mapping in the Sundarbans using U-Net and multi-source satellite data,” in *Proc. 2025 IEEE Space, Aerospace and Defence Conference (SPACE)*. IEEE, 2025, pp. 1–4.

[27] A. Bhattacharjee, S. S. Ebrahim, S. Banerjee, R. K. Gupta, S. S. Chaudhuri, and S. Moulik, “A critical review of methods for automated detection of mangrove deforestation,” in *Proc. 2023 7th Int. Conf. Electronics, Materials Engineering & Nano-Technology (IEMENTech)*. IEEE, 2023, pp. 1–6.

[28] L. Mohebi, “Empowering learners with ChatGPT: Insights from a systematic literature exploration,” *Discover Education*, vol. 3, no. 1, Art. no. 36, 2024.

[29] R. K. Gupta, S. Moulik, S. S. Chaudhuri, and A. Misra, “Multi-scale contextual learning for mangrove detection: A TransUNet model with ViT encoder and pyramid pooling,” in *Proc. 2025 IEEE India Geoscience and Remote Sensing Symposium (InGARSS)*, 2025, pp. 618–622.

[30] C. K. Lo, K. F. Hew, and M. S.-y. Jong, “The influence of ChatGPT on student engagement: A systematic review and future research agenda,” *Computers & Education*, vol. 219, Art. no. 105100, 2024.

[31] T. K. F. Chiu *et al.*, “A scoping review on how generative artificial intelligence transforms assessment in higher education,” *International Journal of Educational Technology in Higher Education*, vol. 21, 2024.