Ethical AI Use: Understanding Perception of Learning Outcomes and Equity

Farhana Shahzad
English Department at LUMS, Lahore University of Management Sciences (LUMS)
Lahore, Pakistan
Email: farhana.shahzad [AT] lums.edu.pk

ABSTRACT—*The integration of artificial intelligence (AI) in classrooms has transformed pedagogical and learning practices alike, allowing for greater student academic achievement. However, its ethical implications remain a critical concern, with educational institutions attempting to address (and often circumvent) the use of generative AI in assessment. By analysing existing literature and conducting semi-structured interviews and focus groups discussions with freshmen at a private university in Pakistan, this research qualitatively examines how generative AI tools impact student perceptions and positionality in debates about educational equity, learning outcomes, and ethical engagement with AI. To triangulate the findings, first year students were divided into experimental and control groups, with the former exposed to monthly AI training. Firstly, the findings showed that the use of generative AI, particularly ChatGPT, and subsequent discussions posit AI as both an obstacle and a path to educational equity. Additionally, AI use in universities is negotiated based on one's positionality, with varying concerns for learning outcomes. There was a difference in trust in AI between the groups, with many vocalising and citing concerns such as algorithmic bias and an uneraseable digital footprint. Moreover, generative AI seems to be positioned differently in ethical understandings, as a personal and institutional ethical problem. This study provides a discussion about the appropriate practices for AI use in education, emphasising the need for clearer guidelines. It also highlights the fast-paced progress of generative AI and how debates and perceptions about these tools are still in their nascent phase.*

Keywords---- Artificial Intelligence (AI), learning outcomes, educational equity, and ethical considerations

1. INTRODUCTION

The increased integration of generative artificial intelligence (AI) in education has severely impacted both the teaching and learning process by offering tools like ChatGPT and Grammarly to help enhance academic support. However, due to its highly transformative potential, foresight is required to understand the challenges it may bring. This research also aims to understand the impact of AI integration on equity, learning outcomes and ethical decision-making. This study examines how exposure to formal AI training to first-year university students at the Lahore University of Management Sciences (LUMS) affects students' learning outcomes, ethical awareness, and perceptions of equity. A mixed-method research design was incorporated, and this research aims to compare two groups: the experimental group, which received formal AI training from the pedagogical partner, and the control group, which did not. Online questionnaires, individual semi-structured interviews, and focus group discussions were facilitated to gather deeper-level insights.

In exploring these issues, this study places a great emphasis not only on the growing influence and the rapidly evolving nature of AI in education but also significantly contributes to the discourses on creating equitable and ethical learning environments. Finally, this research presents the existing gaps and suggests practical strategies to integrate the usage of AI, which serves as a guide for educational institutions and policymakers in addressing the potential challenges and possibilities brought in by generative AI.

2. BACKGROUND

2.1 AI: Defined

In recent years, generative artificial intelligence (AI) has experienced exponential growth, profoundly impacting various sectors. This surge is attributed to advancements in computational power, the availability of vast datasets, and innovative algorithmic developments. The proliferation of digital data and enhanced processing capabilities have enabled the training of complex AI models, leading to significant improvements in generative tasks [1].

The academic and pedagogical implications of generative AI are substantial. Tools such as Grammarly and ChatGPT are becoming increasingly crucial for educational assessment. Grammarly, for instance, utilises AI to assist students in refining their writing by providing real-time grammar and style suggestions, thereby enhancing the clarity and

coherence of their work [2]. Similarly, ChatGPT, an AI language model, offers instant responses to queries and prompts, serving as a valuable resource for idea generation and understanding complex topics [3]. The integration of these tools into educational settings has the potential to transform traditional pedagogical approaches, promoting personalised learning and fostering the development of critical thinking skills. However, this shift also raises concerns regarding equity among students and the impact on learning outcomes, as the ease of access to AI-generated content may lead to misuse by students. Educators are thus challenged by developing pedagogical strategies that ethically incorporate AI tools, ensuring they effectively complement the learning process [4].

2.2 Educational Equity & AI

Educational equity has long been a central concern in academic discourse, emphasising the need for fair and inclusive access to learning opportunities regardless of students' socio-economic backgrounds, geographic locations, or institutional resources. As AI becomes increasingly integrated into higher education, questions arise about its role in either bridging or exacerbating existing disparities. Generative AI holds the potential to personalise learning and enhance accessibility. However, these benefits are not distributed equally, as issues such as algorithmic bias, unequal access to AI tools, and varying levels of AI literacy may deepen learning outcome disparities [5]. The ethical implications of AI in education extend beyond its technical capabilities, requiring critical evaluation of how it influences students' perceptions of fairness, academic integrity, and agency in the learning process.

Firstly, Roshanaei et al. provide a theoretical framework on equity in education, discussing socio-economic, gender, geographical, and cultural factors that influence access to AI-driven learning. The study examines predictive analytics and AI-driven personalised learning, which ties into concerns about whether generative AI creates fair learning opportunities or exacerbates existing inequalities [6]. Additionally, the policy recommendations section offers insights into responsible AI integration, emphasising the need for inclusive design and data protection, which aligns with our ethical AI focus. Overall, this study provides a critical foundation for analysing both the promises and risks of AI in university learning environments, making it a valuable source for this research.

Additionally, Chai et al.'s research on the ethical use of AI, equity, and learning outcome perception in university students because it directly examines how students perceive fairness in AI-driven educational evaluation. The findings that students generally view AI algorithms as fairer evaluators than human teachers due to increased transparency align with broader discussions on AI equity, particularly regarding whether AI fosters or undermines trust in academic settings [7]. The study highlights information transparency and explanation of evaluation outcomes as critical factors in shaping perceptions of fairness, which ties into concerns about whether AI creates equitable learning environments or reinforces existing biases. By examining fairness perceptions in AI-driven assessment, this study provides empirical insights into how generative AI might contribute to or mitigate learning outcome disparities in university students and how perceptions of equity and "fairness" might differ.

Hence, a key research question this study answers is *whether access to formal AI training affects equity perceptions among Pakistani students compared to those without formal AI training*. While Roshanaei et al. discuss broad socio-economic, gender, and geographical disparities in AI-driven learning, they do not provide region-specific insights, particularly for countries like Pakistan, where AI access is unevenly distributed. Similarly, Chai et al. examine students' perceptions of AI fairness but do not consider AI literacy—especially formal AI training—as a factor shaping these views. This study bridges these gaps by investigating whether access to formal AI training affects students' equity perceptions in a Pakistani context, while also addressing how disparities in AI literacy may influence fairness perceptions, particularly in regions (like Pakistan) with stark digital divides [8].

2.3 Learning Outcomes & AI

The increasing integration of generative artificial intelligence (AI) in education has sparked discussions about its impact on student learning outcomes. While AI-powered tools offer personalised learning experiences, automate assessments, and enhance accessibility, they also raise concerns about disparities in learning outcomes [9]. Factors such as varying levels of AI literacy, differences in access to advanced AI tools, and potential biases in AI-generated content may contribute to unequal educational benefits among students.

Wu and Yu's study of AI chatbots provided a foundation for the impact on student learning outcomes. The finding that AI chatbots significantly affect higher education students suggests that generative AI may not benefit all learners equally, potentially widening disparities [10]. Additionally, the study's discussion of short-term novelty effects versus long-term engagement challenges raises essential questions about whether generative AI tools provide lasting academic benefits or primary advantages to students with prior AI proficiency. By identifying variations in AI effectiveness, this study supports an analysis of how generative AI might contribute to uneven learning outcomes among university students, making it directly relevant to our understanding of (non-)ethical AI use.

Furthermore, Alshehri and Alotaibi's research on AI integration in the Kingdom of Saudi Arabia (KSA) is relevant to the present study on learning outcome disparity through generative AI. By examining the integration of AI-based learning outcomes in universities, the study highlights both the opportunities and challenges associated with AI adoption in an academic context. The findings suggest that while AI has the potential to revolutionise learning methodologies and address educational challenges, disparities may arise due to varying levels of AI literacy and access to technological resources [11]. Furthermore, the study underscores educators' need to acquire new technological skills, which is critical when considering the uneven impact of AI on student learning outcomes. Identifying AI as a nascent but transformative force aligns with concerns regarding whether generative AI tools contribute to equitable academic progress or exacerbate existing learning disparities. Thus, Alshehri and Alotaibi's research provides a valuable framework for understanding how institutional factors, technological accessibility, and pedagogical adaptation influence the efficacy of AI in higher education, reinforcing the need for further exploration of learning outcome disparities in the context of generative AI.

Additionally, Zheng et al. provide a quantitative meta-analysis of AI's effectiveness in education, specifically examining disparities in learning achievement and perception [12]. The finding that AI has a high effect size on learning achievement but a small effect on learning perception suggests that while AI may improve objective academic performance, students' subjective experiences with AI-assisted learning may vary, raising concerns about engagement, trust, and accessibility. This is an idea addressed in semi-structured interviews. Moreover, the study identifies multiple moderating factors—sample level, AI role, and hardware availability—critical for understanding equity in AI adoption [13]. Differences in how students at various educational levels and learning environments experience AI point to potential disparities in access, effectiveness, and perceived value, which are central to ethical discussions surrounding AI in higher education. By highlighting both the benefits and limitations of AI in shaping student outcomes, this study informs the ethical considerations of AI integration.

While Wu and Yu's study highlights the significant impact of AI chatbots on higher education students [10], it does not address the sustainability of these benefits over the long term or explore whether formal AI training can mitigate learning outcome disparities. Similarly, Alshehri and Alotaibi's research in Saudi Arabia demonstrates that AI has the potential to revolutionize learning methodologies [11], yet it does not fully explore how differing levels of AI literacy and unequal access to advanced tools might affect students' learning outcomes. Moreover, Zheng et al.'s meta-analysis reveals that although AI significantly boosts objective learning achievement [12], it has only a modest effect on students' subjective learning perceptions, leaving open the question of whether integrating generative AI and formal training can meaningfully enhance students' perceptions of their learning outcomes compared to traditional instructional methods. This study addresses these gaps by investigating if formal AI training not only sustains and equitably enhances academic performance but also improves how students perceive and experience their learning in an AI-integrated environment. Hence, another key research question, this study addresses is whether *the integration of generative AI and training in educational settings influences students' perception of their learning outcomes compared to traditional methods of instruction*.

2.4 Ethical Considerations in the Academic Use of AI

AI (especially generative AI) is becoming profoundly significant to society and productive work. Experts in the field are understanding and debating the boundaries of ethical behaviour when it comes to creating and implementing new AI tools. While there is no current wide-scale governing body to write and enforce these ethics, many technology companies have adopted their own version of an AI code of conduct [13]. Similarly, academic circles are also negotiating (and re-negotiating) what AI use in pedagogical practice looks like [14].

The ethical framework for this research's methodology was based on the questions raised in "Moral AI & How We Get There" by Jana Schaich Borg, Walter Sinnott-Armstrong, and Vincent Conitzer [15]. It explores the possibility of AI being free of bias and used safely, while incorporating other aspects of human morality. It provided a foundational framework for evaluating AI's moral implications, what "ethical" AI use could look like, and what key concerns punctuate the field today.

Additionally, Tang & Yu's systematic literature review of AI integration in education raises five main ethical implications of AI usage: "namely algorithmic bias and discrimination, data privacy leakage, lack of transparency, decreased autonomy, and academic misconduct, with algorithmic bias being the most prominent" [16]. We used this to inform our semi-structured interviews and expanded on their work by having "further substantive discussion" and elucidating "the precise mechanism of ethical principles of using AI models in the classroom" into the implementation phase in the experimental group [16].

Furthermore, Akgun and Greenhow's description of the ethical challenges and benefits of AI integration in K-12 education provides a necessary context for incoming first-year students [17]. While it focuses on K-12 settings, many of the AI applications and ethical challenges it discusses—such as personalised learning, automated assessments, and behavioural tracking—are also present in university environments. First-year students may already have prior exposure to AI-driven educational tools, so examining how these technologies shape their academic experiences and ethical decision-

making is crucial. By addressing how AI is introduced and taught in educational settings, this source also provides insight into how first-year students develop their understanding of AI ethics, making it a valuable reference for analysing their engagement with AI in higher education.

The literature on ethical considerations in the academic use of AI reflects an ongoing debate about the proper boundaries of AI-enabled tools, with many experts and institutions still in the process of defining what constitutes "ethical" AI use. While technology companies have begun to establish their own codes of conduct and academic circles continue to negotiate best practices for integrating AI in pedagogy, there remains no universally accepted framework. For instance, the ethical framework adopted for this study draws on the work of Borg, Sinnott-Armstrong, and Conitzer [15], which offers a useful—but not definitive—starting point for evaluating AI's moral implications. Similarly, Tang and Yu's [16] systematic review identifies several critical issues such as algorithmic bias, data privacy leakage, and decreased autonomy; however, it may not fully capture how these challenges manifest specifically in the classroom setting when AI is applied. Furthermore, while Akgun and Greenhow's [17] research on K-12 environments provides valuable context regarding personalized learning and automated assessments, its direct applicability to university settings, where first-year students are already exposed to various AI-driven tools, remains underexplored. Notably, these studies have not examined these issues within the Pakistani context, where unique technological factors may further shape ethical perceptions and practices. This study, therefore, seeks to build on these findings by examining whether formal AI training can enhance students' ethical awareness and responsible use of AI in higher education—addressing the third research question regarding the impact of AI training on ethical perceptions among Pakistani students. Thus, the study's third and final research question is *whether AI training impacts students' ethical awareness and responsible use in academia compared to educational settings with limited or no AI training*.

3. METHODOLOGY

This study explores the impact of integrating generative AI in academic settings and studies its implications on students' learning outcomes, ethical awareness, and perceptions of equity. Qualitative research method design was employed to provide an understanding of the research questions. A mixed method design was adopted to understand the impact of AI training on experimental groups and compare them with the control group, which did not receive such training. The participants of this study were first-year university students studying at the Lahore University of Management Sciences (LUMS) enrolled in the Writing and Communication course, which is part of the university's core curriculum and focused on honing writing skills. The study consisted of the following two groups:

- Experimental Group: Students who received training on AI from their pedagogical partner.
- Control Group: Students who were not exposed to formal training on AI usage.

Purposive and volunteer sampling techniques were employed to select participants from various academic disciplines and based on AI usage level while ensuring students willingness to participate in the study.

The data for the study was collected through semi-structured interviews, which allowed for individuals' perspectives on AI's role in learning, ethics and equity. Moreover, in order to cement the analysis of these interviews, we incorporated focus group discussions to facilitate group-level insights regarding the implications of AI. The structure was such that separate discussions for experimental and control group students were held to compare perspectives. Initially, questionnaires were rolled out to gather small-scale quantitative data on AI usage, which looked for the frequency of AI use, perceived fairness and effect on academic performance. These questionnaires were distributed online to ensure accessibility. Ethical measures were rigorously followed throughout by acquiring informed consent as all participants were briefed about the study's purpose and their right to withdraw at any given point. Personal identifiers and descriptors were excluded for the sake of maintaining the participant's confidentiality.

4. DISCUSSION

The existing literature on AI in education underscores both its transformative potential and critical gaps in understanding its equitable, effective, and ethical integration into academic settings. Studies such as Wu and Yu's [10] suggest that while AI chatbots may enhance learning outcomes, questions remain about the long-term benefits and potential disparities arising from uneven AI literacy. Similarly, research by Alshehri and Alotaibi [11] and the meta-analysis by Zheng et al. [12] reveal that although AI can boost objective academic performance, its impact on subjective learning experiences and the role of formal training in mitigating disparities are less clear. In the realm of ethics, frameworks by Borg, Sinnott-Armstrong, and Conitzer [15], along with reviews by Tang and Yu [16] and insights from Akgun and

Greenhow [17], provide valuable perspectives on AI's moral implications; however, they do not fully examine whether structured AI training can enhance ethical awareness and responsible use, particularly within contexts like Pakistan, where digital divides are pronounced. Collectively, these findings inform the following three key research questions:

1. Does access to formal AI training affect equity perceptions among Pakistani students compared to those without formal AI training?
2. Does the integration of generative AI and training in educational settings influence students' perception of their learning outcomes compared to traditional methods of instruction?
3. Does AI training impact students' ethical awareness compare to educational settings with limited or no AI training?

The following Section 4 and subsequent sub-headings analyse the data from the semi-structured interviews considering these questions.

4.1 Equity Perceptions

Both groups acknowledge AI's potential to create inequities. Still, the experimental group—having received formal AI training—demonstrates a more nuanced understanding of how AI biases operate, whether they help or hinder students from different socio-cultural backgrounds and who, ultimately, is responsible for AI equity.

The control group homed in on AI's limitations in cultural representation, stating that AI often reflects Western perspectives and may misrepresent global issues (e.g., the Israel-Palestine conflict). However, they do not discuss mitigation strategies, or the possibility of AI being intentionally trained to reduce biases. For example, a participant from the control group focus group stated, *"AI is feeding you narratives, so instead of researching further into it, I feel like they just accept it in a way."* This shows a passive consumption of AI-generated content, reinforcing biases without critique. There is an acceptance of "AI narratives".

Additionally, another participant in the focus group stated, *"It plays safe and goes on by terming it as a conflict. So, consider biases in terms of cultural, racial, gender, or linguistic factors and how they might influence the educational outcomes for different groups of students."* While identifying equity problems through AI is strong, there is no discussion of mitigation. The students in the control group frame AI as an external force they can observe, not something they work with or whose adverse effects they can mitigate.

On the other hand, the experimental group with AI training are more analytical about AI bias, citing direct experiences of ideological bias in AI-generated content (e.g., anti-feminist slant, differential framing of conflicts). They acknowledge Western dominance in AI training data but also recognise how user knowledge of AI can help navigate these biases. A focus group participant stated that *"AI tools like ChatGPT and all, they are more Western biased. They are more biased towards the Western side... it has more content about the Western side, like the US and UK, and it limits the information on our South Asian history."* Similar to the control group, this shows a recognition of algorithmic bias.

However, the experimental group is more personalised and inclusive of their own experiences, and they are more agentive in their understanding of equity and AI. *"I was talking to one of our classmates, [and] her research topic is about how feminism is affected... and she said when she was asking AI to generate a couple of points that would help her research, it had a lot of very gender-biased views, like it was very anti-feminism and very pro the other side."* This quote from a participant in the experimental group's focus group shows how understanding ideological biases and spotting how generative AI makes mistakes shows how training can make users part of the process. Another focus group participant stated, *"I think that is because, first of all, it is developed by the Western side… So, I mean, it's sad to say, unfortunately, Pakistan or South Asian countries have not done so much that AI would have a personal bias towards us. It has to be against us."* This framing, while fatalistic, suggests a realisation that AI bias is not neutral but shaped by global power structures. A student in the experimental group also expressed similar ideas in a semi-structured interview: *"Algorithmic bias is more towards Western-centric perspectives—general opinions tend to favour the Western world over South Asian perspectives."* This reflects a more nuanced understanding of AI biases in one-on-one conversations, too, recognising not just individual unfairness but also structural inequities in how AI generates and privileges certain viewpoints. Unlike the control group, these students move beyond merely identifying bias and discuss its origins. Hence, students without training tend to view bias as an obstacle they cannot control, while those with training see AI bias as a structural issue but acknowledge user agency in navigating it.

Students' understanding of equity in the educational context was also slightly different. For the control group, AI was a barrier to equity for some, with no concrete solutions. *"If you are someone from a background where technology was not that easily available, or you're not that good with computers, I would say it might affect how effectively you can use AI."* This focus group student acknowledges the digital divide but does not suggest ways to address it, showing how the control group recognises that students from underprivileged backgrounds may struggle to use AI effectively but does

not explore how AI training could bridge this gap. Another student said, *"During the time that they take to do that, they're going to be disadvantaged compared to the other students."* The control group sees a lack of familiarity with AI as a permanent disadvantage rather than a skill gap that can be closed.

Conversely, the experimental group's understanding of educational equity and AI viewed it as a tool to bridge learning gaps, if used properly. *"I think it can benefit some underprivileged students because it could help them formulate their essays or their assignments because they are new to this or they are not much proficient in English."* This quote from a focus group participant shows that students with AI training recognise that AI can serve as an equaliser by providing support for those struggling with academic writing. AI as a way to be more meritocratic. For example, *"It can help students who are not so proficient in English… but it can also make them dependent. Those students might not try to learn English and will just ask ChatGPT for everything."* This student in question identifies a double-edged effect: AI can help students from non-English-speaking backgrounds learn better, but without guidance, it may hinder their language acquisition instead of fostering growth. The view of AI as a compensatory tool is also shown when a focus group participant says, *"For instance, I was reading something in my economics book, and I just couldn't understand it, so I asked ChatGPT to explain it to me like I'm a fourth grader. It broke everything down into simple terms with examples."* For those without tutors or a lack of personalised instruction, students in the experimental group could see how AI may make difficult concepts more accessible. A student expressed in an interview: "*"A fair and equitable classroom means everyone has access to resources—no one should be deprived, AI or not, paid or unpaid."* This solo response, while acknowledging inequity, focuses on AI as a tool that either expands or limits access. It lacks a deeper analysis of algorithmic bias but still shows awareness of structural disparities.

Hence, the control group acknowledges barriers to AI access but views them as fixed disadvantages rather than problems that can be addressed, while the experimental group sees AI as a tool that can bridge educational gaps, particularly for students struggling with English or self-study.

4.2 Learning Outcome Perceptions

Students' perceptions of their learning outcomes varied significantly between those with and without AI training:

The control group largely saw AI as a tool that hindered independent learning and encouraged over-reliance, reducing students' ability to critically engage with course material. One student expressed in a focus group, *"Ever since AI has been there, I only go to ChatGPT, ask it that one question, and that's my only avenue of learning… I don't explore as many channels as I used to because there's an easy way out."* This highlight concerns that AI discourages students from conducting deeper research or engaging with multiple sources, ultimately affecting the depth of their understanding. Another student admitted during a focus group discussion, *"Nowadays, during finals, we do not have that much time to solve a question 10 times and get to the right answer. So instead of even trying to open the book, I just went to ChatGPT, wrote the names of the topics, made it summarise the topic, and asked a few questions, and that's that."* This suggests that while AI provides convenience, it also diminishes traditional study habits by encouraging students to seek immediate answers instead of working through complex material. The control group primarily viewed AI as a shortcut that weakens traditional learning methods, making students more dependent on technology rather than fostering self-sufficiency.

In contrast, the experimental group, having received AI training, took a more strategic approach to AI use, recognising it as a support tool rather than a replacement for learning. One student stated during a focus group, *"It has summarised very long readings, like, for example, [Management Science] readings. It…gave quick notes."* This reflects a perception of AI as a time-saving tool that helps manage overwhelming workloads without necessarily replacing critical thinking. Another participant emphasised AI's role in compensating for gaps in instruction, particularly in cases where traditional teaching methods were insufficient: "*My economics teacher is not that great, so I had to self-study everything. So it helped me with my readings from the economics book... ChatGPT ne hi mujhe pass kar wale [I only passed because of ChatGPT]."* Here, AI is seen to fill in gaps where human instruction falls short, demonstrating that students with training recognise its value in enhancing, rather than replacing, their learning experience. Another participant acknowledged the risks of over-reliance but emphasised the importance of controlled use, stating, *"I feel like it does hinder your learning because you do become really reliant on it. But to some extent, if you learn how to control that, I think it can be really helpful."* This suggests that AI training helps students develop a balanced approach, where AI is used for efficiency without replacing independent thought.

The control group also expressed scepticism about AI's ability to support deep learning, particularly in creative and analytical subjects. One student remarked, *"For creative writing as we get in Writing & Communication, if everyone's using ChatGPT overall, it would sound very generic and we would lose the creativity that we used to have in our writings because everyone's using it, it would all be generalised."* This reflects a perception that AI-generated content lacks originality, raising concerns that students may lose critical thinking and creativity when relying too much on AI. Another participant echoed this sentiment in a discussion about academic dishonesty, saying, *"When you're faced with an assignment, you're just not able to think, and your brain's default is just to go to AI and ask it to do the work for you."* This

illustrates how students in traditional learning settings see AI as a crutch that diminishes cognitive engagement, ultimately weakening their problem-solving skills over time.

Meanwhile, the experimental group recognised AI's potential to enhance analytical and research skills when used correctly. One participant highlighted AI's role in helping structure essays and brainstorming ideas, stating, *"It helps us brainstorm. It helps us structure our essay or tells us which points to focus more on."* Unlike the control group, which viewed AI-generated ideas as a threat to originality, students with AI training saw it as a tool for improving argumentation and organisation. Another student emphasised how AI can aid comprehension by simplifying difficult material: *"For instance, I was reading something in my economics book, and I just couldn't understand it, so I asked ChatGPT to explain it to me like I'm a fourth grader. It broke everything down into simple terms with examples."* This suggests that AI training enables students to use AI as an educational assistant rather than a replacement for deep learning, making complex concepts more accessible while still requiring students to engage with the material. Even students in semi-structured interviews expressed similar sentiments: "*Using AI for brainstorming and structuring is fine, but total reliance on AI would be a hindrance to critical thinking—it should make learning accessible, not replace it."* This statement shows an awareness of both AI's benefits and risks, reinforcing the idea that training leads to a more balanced, intentional approach to AI in learning.

Overarchingly, however, the rapid expansion of AI use in higher education has created both opportunities and challenges for students. Many recognise AI's efficiency in summarising readings and assisting with coursework, but they also acknowledge the risks of over-reliance. One student explained, *"In my political science course, readings that would take me four hours to go through can be done in 15–20 minutes using AI."* While this significantly accelerates study time, it may also reduce deep engagement with the material, weakening critical thinking and research skills. Another participant admitted, "Instead of even trying to open the book, I just went to ChatGPT, wrote the name of the topics, and made it summarise the topic." This illustrates how AI can lead to passive learning, where students absorb information without actively analysing or synthesising it.

Students also highlighted AI's varying reliability across different subjects, shaping their perceptions of its role in education. One interviewee noted, "*In calculus, if you have the answer key, ChatGPT can help you verify your answer*," indicating AI's usefulness in objective, quantitative fields where solutions can be cross-checked. However, concerns arose in writing-intensive disciplines, where AI's role in brainstorming and structuring arguments remains debated. Without clear institutional guidelines, some students questioned whether AI-assisted work compromises originality and academic integrity. While some instructors integrate AI responsibly, inconsistencies in educational policies create uncertainty about appropriate AI use in academic writing.

Training played a significant role in shaping AI use, with students in the experimental group demonstrating a more critical and strategic approach. One participant stated during a focus group, *"Cross-checking can make it reliable. For example, in the last quiz, my roommate used AI to write his research essay and asked GPT for citations, but the citations were incorrect."* This awareness helped trained students maximise AI's benefits while mitigating risks, demonstrating a more informed approach to AI verification and accuracy. Conversely, control group students, who lacked structured guidance, exhibited either scepticism or misuse of AI, limiting their ability to engage effectively with course material. The absence of AI training led some to over-rely on AI without understanding its limitations, while others avoided it entirely due to concerns about unverified information and ethical dilemmas.

Overall, the control group perceived AI as undermining traditional learning by encouraging shortcuts, reducing independent research, and weakening creativity. They largely saw AI as a hindrance to deep learning, promoting superficial engagement rather than fostering meaningful intellectual growth. In contrast, the experimental group, having received AI training, saw AI as a tool for efficiency and comprehension, particularly in summarising content, structuring arguments, and clarifying complex topics. However, they also acknowledged its risks, recognising that AI should be used as an aid rather than a replacement for critical thinking. This suggests that AI training helps students develop a more nuanced and strategic approach, allowing them to integrate AI into their learning process without becoming entirely dependent on it.

4.3 Ethical Behaviour

Students' ethical awareness regarding AI use varied significantly depending on whether they had received formal AI training. The control group primarily framed AI ethics as a matter of personal responsibility, focusing on individual integrity and whether students were using AI in a way that could be considered cheating. One student stated, *"I feel like completely writing assignments is definitely considered cheating. But even if we talk about idea generation, the liberty to create our own ideas is sort of the purpose of some assignments. If we hand it out to AI, I feel like that is also a form of cheating."* This perspective suggests a rigid understanding of academic integrity, where even using AI for brainstorming is viewed as ethically questionable. Another participant highlighted a pragmatic approach to AI use, stating, "*There's always a workaround to AI detection tools, like paraphrasing AI-generated essays."* Rather than discussing ethical

guidelines, this student focused on how AI restrictions could be bypassed, reflecting a more individualistic and reactive approach to AI ethics, where the focus is on avoiding detection rather than considering AI's broader ethical implications.

In contrast, the experimental group placed greater emphasis on institutional responsibility and structured AI policies, recognising that ethical AI use is not just a personal decision but something that should be defined and regulated within educational institutions. One student stated in a focus group, *"If the student still chooses to use AI after being given clear guidelines, then the blame comes on to the student. But the institution has to ensure that AI use is defined properly."* This highlights a shared responsibility model, where students are expected to act ethically, but universities and schools must first establish clear rules and educate students on responsible AI use. Another participant went further, discussing how AI bias in grading must be addressed, saying, *"Teachers should second-check AI-graded assignments to ensure there's no bias, rather than just relying on AI blindly."* This reflects a broader awareness of ethical AI concerns beyond plagiarism, demonstrating that AI-trained students think critically about AI's fairness in academic assessment, not just about whether using it constitutes cheating.

The control group also viewed AI ethics in a binary way, where AI use was either acceptable or dishonest, with little discussion of nuance. One participant remarked, *"I think AI-generated posters are unfair because it wasn't our creativity that was being used; it was AI."* This statement reflects a concern that AI use diminishes original student work, reinforcing the idea that AI fundamentally undermines fairness in academic settings. Another student shared a similar sentiment during the focus group, arguing that *"Taking inspiration from ChatGPT is fine, but copying completely would be unethical."* While this acknowledges a spectrum of AI use, the focus remains on avoiding academic dishonesty, rather than considering ethical AI integration at a systemic level.

By contrast, the experimental group proposed concrete solutions for ensuring ethical AI use, suggesting structural approaches rather than relying on individual integrity. One participant recommended tracking students' work progression to prevent excessive AI dependence: "*We could track version history on Google Docs. If you're copying, the entire paragraph will show as pasted. That might help limit unethical AI use."* This reflects a proactive approach to academic integrity, where AI ethics is monitored through technological means rather than simply trusting students to self-regulate. Another student noted how educators could implement safeguards, saying, *"One of the instructors was asking students to write a daily log of their research process. This way, they couldn't just use AI to generate everything at once."* Unlike the control group, which focused on punishing AI misuse, the experimental group suggested institutional reforms that could encourage responsible AI use while still allowing students to benefit from its capabilities.

Students' understanding of ethical AI use and accountability also differed between the two groups. The control group primarily framed AI ethics as a matter of individual responsibility, emphasising personal integrity and self-discipline when using AI tools. One student stated in a focus group, *"I feel like completely writing assignments is definitely considered cheating. But even if we talk about idea generation, the liberty to create our own ideas is the purpose of some assignments. If we hand it out to AI, I feel like that is also a form of cheating."* This perspective suggests that the ethical burden of AI use falls solely on the student, without consideration of how institutional policies or systemic disparities might shape students' engagement with AI. Another student expressed during a focus group, *"There's always a workaround to AI detection tools, like paraphrasing AI-generated essays,"* which indicates a focus on evading detection rather than on ethical engagement with AI as a learning tool. This perspective reflects a reactive approach to AI ethics, where students navigate rules individually rather than advocating for broader structural accountability.

Conversely, the experimental group placed greater emphasis on institutional responsibility, seeing AI ethics as something that should be clearly defined and regulated by educational institutions. One student stated during a focus group discussion, *"If the student still chooses to use AI after being given clear guidelines, then the blame comes on to the student. But the institution has to ensure that AI use is defined properly."* This quote illustrates a belief that while students must act responsibly, universities and schools must first provide ethical AI training and clear expectations for its use. Another student suggested, *"Teachers should second-check AI-graded assignments to ensure there's no bias, rather than just relying on AI blindly."* This demonstrates an awareness of how AI might reinforce grading inequities and the need for institutional safeguards against potential biases in AI-generated evaluations. Additionally, a participant proposed using *"version history on Google Docs. If you're copying, the entire paragraph will show as pasted. That might help limit unethical AI use."* Similar ideas came up in semi-structured interviews: "*LUMS should have clearer guidelines—not just a one-time session in O-Week. We need structured policies on AI use throughout the semester."* This respondent sees AI ethics as an institutional issue requiring ongoing regulation and support rather than just detecting individual cheating. Unlike the control group, which focused on individual accountability, students with AI training suggested practical, structural solutions for ensuring ethical AI use while maintaining fairness.

Hence, the control group understood AI ethics at the level of personal integrity, seeing AI-related academic dishonesty as a matter of student choice without acknowledging systemic factors that shape AI access and use. Meanwhile, the experimental group advocated for institutional guidelines and monitoring mechanisms, recognising that ethical AI use cannot solely rely on individual integrity but must be reinforced by clear policies that prevent AI from exacerbating educational inequities.

Overall, the control group understood AI ethics primarily in terms of academic dishonesty, emphasising individual responsibility and cheating prevention, with little discussion of broader ethical considerations. Their approach to AI ethics was reactive, focusing on workarounds, detection methods, and moral dilemmas related to plagiarism. In contrast, the experimental group viewed AI ethics as a structural issue requiring institutional oversight, advocating for clear policies, monitoring tools, and ethical training to guide responsible AI use. While both groups recognised AI's ethical challenges, those with AI training were more engaged with developing solutions, suggesting that formal AI education enhances students' ability to think critically about AI ethics beyond issues of plagiarism.

5. CONCLUSION

The study explored student perceptions of learning outcomes, equity, and ethical considerations in AI-integrated higher education, comparing students with and without formal AI training. The findings reveal that AI literacy significantly influences students' understanding of algorithmic bias, responsible AI use, and learning strategies. AI-trained students emphasized the importance of institutional guidelines, while untrained students primarily viewed AI ethics as a matter of personal responsibility. Furthermore, ethical considerations varied between groups: control group students framed AI ethics as an individual concern, whereas the experimental group stressed institutional accountability and the necessity of structured policies. These findings underscore the importance of AI training in creating an equitable learning environment and mitigating the risks of over-reliance and bias.

Moreover, the study underlines the need for well-defined institutional policies to guide ethical AI use and ensure equity in higher education. Universities and policymakers must develop comprehensive frameworks that integrate AI responsibly into academic settings. Establishing clear guidelines and ensuring fair access to AI tools can create a more inclusive and ethical learning environment, equipping students with the skills needed for an AI-driven academic and professional landscape. To achieve this, institutions should consider the following measures:

1. Assess AI literacy levels to ensure students engage with AI on an equitable footing and receive appropriate support.
2. Integrate AI into curricula to cultivate responsible and ethical AI use across disciplines.
3. Provide faculty training on incorporating AI into pedagogy to enhance both teaching methods and student learning experiences.
4. Ensure equitable access to AI tools by offering institution-wide access to premium AI platforms, preventing disparities in academic resources.
5. Embed AI ethics training into coursework to develop students' critical awareness of algorithmic bias and ethical AI use.
6. Conduct AI literacy workshops as part of student orientation to establish foundational knowledge, particularly for those with limited prior exposure.
7. Develop institutional AI policies that clearly define acceptable use, ethical considerations, and academic integrity guidelines.
8. Encourage interdisciplinary AI research to explore innovative applications of AI across various academic fields.
9. Establish AI mentorship programs where AI-literate students assist peers in developing AI-related skills, fostering a collaborative learning environment.

By implementing these measures, universities can better prepare students for an AI-driven future, strengthening both their ethical awareness and technological proficiency.

6. CONFLICT OF INTEREST DECLARATION

There is no conflict of interest to declare by the corresponding author.

7. ROLE OF THE FUNDING SOURCE

Lahore University of Management Sciences (LUMS) Mushtaq Gurmani School of Humanities and Social Sciences (MGSHSS) Proposals & Grants(P&G) Committee funded the Research Assistants' salaries. The P&G committee had no role in the research process of the article.

8. ETHICS APPROVAL

Ethics approval was obtained from LUMS Institution Review Board (IRB) before commencing the research- Protocol no: IRB-0334

9. AI DISCLOSURE

AI has only been used to improve readability and language of the work. The author is ultimately responsible and accountable for the contents of the work.

10. ACKNOWLEDGEMENT

The author acknowledges Research Assistants Malaika Arbab Gohar and Rana Muhammad Ahmad.

11. REFERENCES

[1] Exponential growth — the science of machine learning & AI. (n.d.). The Science of Machine Learning & AI. https://www.ml-science.com/exponential-growth
[2]*Transforming communication through AI innovation*. (n.d.). https://www.grammarly.com/ai
[3] Roumeliotis, Konstantinos I.; Tselikas, Nikolaos D. (2023). "ChatGPT and Open-AI Models: A Preliminary Review". Future Internet. 15 (6): 192. doi:10.3390/fi15060192.
[4] *ChatGPT and other AI Tools: Implications for teaching and learning*. (2024, March 25). Center for Teaching and Learning. https://ctl.prod.acquia.cws.oregonstate.edu/chatgpt-and-other-ai-tools-implications-teaching-and-learning
[5] *Without universal AI literacy, AI will fail us*. (2022, March 17). World Economic Forum. https://www.weforum.org/stories/2022/03/without-universal-ai-literacy-ai-will-fail-us/
[6] Roshanaei, M., Olivares, H., & Lopez, R. R. (2023). Harnessing AI to Foster equity in Education: Opportunities, challenges, and emerging strategies. *Journal of Intelligent Learning Systems and Applications*, *15*(04), 123–143. https://doi.org/10.4236/jilsa.2023.154009
[7] Chai, F., Ma, J., Wang, Y., Zhu, J., & Han, T. (2024). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Frontiers in Psychology*, *15*. https://doi.org/10.3389/fpsyg.2024.1221177
[8] Shah, M. (2024, September 19). The deplorable state of Pakistan's digital divide. *Brecorder*. https://www.brecorder.com/news/40322876
[9] School, S. L. (2024, June 29). *How will AI Impact Racial Disparities in Education? | Stanford Law School*. Stanford Law School. https://law.stanford.edu/2024/06/29/how-will-ai-impact-racial-disparities-in-education/
[10] Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. https://doi.org/10.1111/bjet.13334
[11] Alotaibi, N. S., & Alshehri, A. H. (2023). Prospers and Obstacles in Using Artificial Intelligence in Saudi Arabia Higher Education Institutions—The Potential of AI-Based Learning Outcomes. *Sustainability*, *15*(13), 10723. https://doi.org/10.3390/su151310723
[12] Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2021). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, *31*(9), 5650–5664. https://doi.org/10.1080/10494820.2021.2015693
[13] Schiff, D. (2021). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2
[14] Borenstein, J., & Howard, A. (2020). Emerging challenges in AI and the need for AI ethics education. *AI And Ethics*, *1*(1), 61–65. https://doi.org/10.1007/s43681-020-00002-7
[15] Borg, J. S., Sinnott-Armstrong, W., & Conitzer, V. (2024). *Moral AI*. Penguin Press.
[16] Tang, L., & Su, Y. (2024, March 1). *Ethical Implications and Principles of Using Artificial intelligence models in the Classroom: A systematic literature review*. https://reunir.unir.net/handle/123456789/16210?show=full
[17] Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI And Ethics*, *2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7