

Beyond the Numbers: A Qualitative Exploration of Female University Students' Experiences with Generative AI in Personalized Learning at a Pakistani University

Muhammad Sikandar Ali ^{1*} Rimsha Arshad ² Fatima Akram ³

¹ Alumna of the M.Phil. in Education, University of Education, Lahore, Punjab, Pakistan.

Email: sikandarali231098@gmail.com

² PhD Scholar, Department of Education, Lahore College for Women University (LCWU), Lahore, Punjab, Pakistan.

Email: RimshaArshad.lcwu@gmail.com

³ Alumna of the M.Phil. in Educational Leadership and Policy Studies, University of Education, Lahore, Punjab, Pakistan.

Email: Fatimaakram382@gmail.com

ABSTRACT: Generative Artificial Intelligence (AI) has become an important part of academic practice in everyday academic life; however, there is a lack of understanding about how female students in the higher education system in Pakistan interact with it. The context in which the use of AI in learning occurs is marked by English-medium teaching, collectivist academic culture, and gender differences in terms of technology adoption. This qualitative descriptive study is an attempt to fill this research gap based on semi-structured interviews with twelve female students: nine studying at the Department of Education and three at the Department of Computer Science at a women only public university in Lahore. Interviews continued until theoretical saturation of the data was reached and they were analyzed using the six-phase reflexive thematic analysis described by Braun and Clarke (2006) within the framework of the Extended Mind Thesis (Clark & Chalmers, 1998). Four themes emerged in response to the research questions that guided this study: Generative AI as a Tool of Scaffolding Personalized Learning, Effectiveness and Ambivalence Towards Generative AI at the Institutional Level, AI as a Psychological and Academic Enabler, and The Over-Reliance Paradox. Instead of naive and uncritical users, the data presents the participants negotiating their dependence on AI and in some ways being more aware of this fact than their institution. The paper provides a context-specific contribution to the literature which is mainly dominated by survey-based studies that do not consider gender differences.

KEYWORDS: Generative AI, Personalized Learning, Gender, Higher Education, Pakistan

Introduction

There has been an incredible growth and evolution in generative AI technologies such as ChatGPT, Gemini, and GitHub Copilot in higher education to the point where these tools offer on-demand scholarly help to students in a conversational form which not a single teacher can provide to them (Kasneci et al., 2023; Lo, 2023). It is important because what this change in the landscape of generative AI has done to personalization is that personalized learning used to depend on provisioning, but now it has come down to the demands that the students put forward by

Pages: 14 – 31

Volume: 5

Issue: 5 (May 2026)

Corresponding Author

Muhammad Sikandar Ali

✉ sikandarali231098@gmail.com

Cite this Article: Ali, M. S., Arshad, R., & Akram, F. (2026). Beyond the Numbers: A Qualitative Exploration of Female University Students' Experiences with Generative AI in Personalized Learning at a Pakistani University. *The Regional Tribune*, 5(5), 14-31. <https://doi.org/10.55737/trt/v-v.264>

themselves privately and on their own devices. It is based on evidence that there is considerable focus on well-endowed educational systems of the US, UK, and similar countries (Chan & Hu, 2023).

The higher education sector in Pakistan has some unique aspects that relate to this research study. There are three important conditions that should be considered. First, the English language is used as the main language of instruction in university education despite the fact that it is the second or even third language for many students. It means that the Generative AI tool needs to be understood and prompted by non-native language speakers in order for it to be effective. Second, the continuous around-the-clock faculty support that is common in many studies related to the application of AI in the educational process is not typical of the conditions in Pakistan. Third, all-female universities that account for a considerable number of female students in the country have not been considered in the qualitative aspect of students' experience of Generative AI.

The difference is not limited to location; rather, it pertains to a lack in the kind of knowledge that is created. The current body of literature on Pakistan and other countries mainly involves the measurement of indicators like the presence of Generative AI use by students, level of satisfaction, and correlation between use and performance (Barra et al., 2024; Zaman & Malik, 2026). The questions may be valid and pertinent to ask; however, they have nothing to do with the main topic of this paper, which revolves around how students perceive, evaluate, and justify their use of AI technology. The statistics about the adoption of technology show the presence of Generative AI use by women students in Pakistan; however, it does not explain how the students understand the technology being used.

The analysis acquires another, specifically gender-related layer in the process of development. As Barra et al. (2024) note, the technology adoption of Pakistani women is affected not only by technical skills but by social expectations equally, making gender the factor actively affecting technology adoption in this setting and not just a controlling one. As this kind of shaping is performed through the act of interpretation, the interpretation of confidence, hesitation, or the justification of using a certain technology, the data obtained in the process of such an act cannot be measured through adoption intention questionnaires dominating the existing literature. That is why the present study requires this specific approach to be applied.

Thus, the present research deals not with the questions of whether Pakistani university female students use Generative AI for personalized learning and how they estimate its effects, but how they do this and perceive the technology in terms of personal academic confidence and performance. In other words, this research is significant twice as well. From the academic point of view, it adds new information to the corpus of Generative AI literature, which is characterized by the dominance of large-sample surveys (Kasneci et al., 2023). On the practical side, it helps solve the problem of developing policy on the technology by providing evidence on how students in Pakistan use it, because at present there is little direct evidence available to Pakistani universities.

To address this problem, the study conducted semi-structured, in-depth interviews with female students at Lahore College for Women University (LCWU), guided by three research questions:

RQ1: How do female university students describe and make sense of their use of Generative AI for personalized learning?

RQ2: What are students' perceptions of the effectiveness of Generative AI, and what concerns do they associate with its use?

RQ3: How do students perceive the influence of Generative AI on their academic performance, engagement, and confidence?

The paper consists of five sections. In the Literature Review section, the current research is positioned in relation to the generative AI research in general and in Pakistan, related to the existing literature on gender and technology adoption, and linked to the theoretical framework consisting of three components. In the Methods section, the qualitative descriptive approach to the research design, participant selection, data collection, and analysis, along with approaches used to ensure the trustworthiness of the research, are described. In the Findings section, three main themes, one divergent case, and one cross-cutting theme are presented, accompanied by quotes from participants. In the Discussion section, the findings are interpreted based on the previous literature review.

Literature Review

Generative AI in Higher Education

The claim that Generative AI is able to personalize academic assistance on a scale impossible for any single educator has become a well-substantiated one (Crompton & Burke, 2023). The authors Baidoo-Anu and Ansah (2023) describe applications like ChatGPT as being a harbinger of actual change in the essence of academic communication through the generation of personalized feedback, instructions, and solutions to problems posed by students. There are, however, some considerations to be made when approaching this statement with too much enthusiasm. According to Holmes et al. (2019), technologies are equally capable of perpetuating the same inequalities as addressing them, unless educational institutions deliberately seek out equity in the development of technology. This is precisely what the current research tries to tackle; instead of taking the benefits or drawbacks of Generative AI for granted, it asks the students themselves to express their opinions on the matter.

Generative AI in Pakistani Higher Education

There are two such studies carried out recently in Pakistan. The first is by Majeed et al. (2025). The researchers conducted interviews of sixteen students attending both public and private universities in Lahore, and it transpired that the use of Generative AI was considered to improve research efficiency and provide individualized assistance, although there was apprehension over possible problems with misinformation, bias, and the collectivist approach in academia making it difficult to rely on the AI. The second study was conducted by Zaman and Malik (2026), who sampled students of three Lahore universities through purposive sampling, and concluded that Generative AI usage was linked to the perception of an easy and useful tool for research while worrying about dependency on such tools and decreased creativity. However, in these studies, gender was treated as another demographic characteristic, and no gender-specific analysis was done in either paper. In addition, the authors did not explore whether the subject of the discipline a student studied could play a role in her experience of using these tools.

Generative AI and Personalized Learning

The common theme in the existing literature is that the genuine educational value of Generative AI resides in its adaptability, which refers to the ability to interact with a learner based on their current level, as opposed to giving the same explanation to all students (Dai et al., 2023). The research by Tlili et al. (2023) shows that such adaptability does not come for free: deliberate, well-crafted prompts produce significantly greater academic benefit than the generic or passive ones, which means that how the student interacts with the tool matters just as much as what the tool itself is. As stated by Lo (2023), who conducted a systematic review of the field, artificial intelligence-based personalization brings its value when used in an environment that encourages higher-order thinking as opposed to the passive repetition of the answer provided. All of this implies that prompting becomes a skill that should be studied systematically rather than just being a technical procedure.

Gender, Technology Adoption, and Higher Education in Pakistan

In addition to literature on AI adoption specifically, an additional stream of research that deals with gender and technology adoption in Pakistan can be considered an important source of information about the topic of the current study. According to Barra et al. (2024), technology adoption among men in Pakistan is most likely to depend on the perception of one's technical skills, while women are most influenced by social approval when adopting technology. The authors illustrate the phenomenon with a "multitasking whirlpool," where technology adoption by women is constantly affected by various academic and household needs. Gender should not be seen in this case as an external factor that has no influence on the process of technology adoption but rather as something that affects the time, manner, and justification of technology adoption.

Women-oriented institutions, like LCWU, have a distinctive place in relation to the stated trend. In the classroom environment, there are no differences in the educational experiences of the learners based on their gender since all the students attend the same institution for education. In other environments, however, the students continue living in the society described by Barra et al. (2024), where women's adoption of technology differs widely depending on individual characteristics. An important question arises here, which cannot be answered using adoption-intent surveys: not only do female students of a Pakistani university use Generative AI, but also how do they talk about this phenomenon, what do they justify or claim, and whether being in a women-only environment changes anything.

Student Perceptions of AI Effectiveness

Students across different countries perceive Generative AI as an effective method of learning new things, researching, and getting ready for exams, and such perceptions sometimes lead to the official introduction of the technology into curricula (Chan & Hu, 2023; Wardat et al., 2023). Gerlich (2025), however, conducted a massive mixed-methods research study that sheds light on how Generative AI is perceived and used by students and suggests that the higher the use of the technology becomes, the poorer the results in critical thinking become. This effect happens due to cognitive offloading – delegation of mental efforts to the technology. Interviews in the study demonstrated that cognitive offloading could happen while students still felt positive about the tool being used – thus, there was a disconnect between the affective evaluation of the tool and its actual effectiveness in terms of developing one's own reasoning abilities. This is one of the reasons why the current study focuses on interviews rather than scaling the phenomenon.

Generative AI, Academic Engagement, and Second-Language Learners

Various researchers have found that the usage of Generative AI increases engagement among students (Zhai et al., 2021); nevertheless, the process behind the relationship still lacks clarity. According to Kohnke et al. (2023), it is not the increased availability of information, but the decrease in academic anxiety that causes the increased engagement. This theory seems especially relevant when it comes to students who need to submit academic assignments in a second language.

Academic Integrity, Critical Thinking, and Over-Reliance

One of the most controversial areas of concern within this area of research is academic integrity. According to Cotton et al. (2024), there was a sudden need for universities to counteract AI-driven cheating, given the advent of tools such as ChatGPT, as the response was mostly reactive (detection and punishment of offenders rather than redesigning assessments) rather than proactive. In addition, Perkins (2023) argues that the majority of breaches of

integrity concerning the use of AI are not the result of deceit but the lack of clearly defined rules for the appropriate use of such technologies in various educational establishments. Kasneci et al. (2023) take this discussion one step further by identifying the link between the use of AI and critical thinking as an unresolved issue, as the technology which helps in performing better at a certain task can at the same time undermine the level of critical thinking used in order to accomplish this task. The constant interplay between visible advantages and hidden disadvantages becomes the major theme of concern in this study.

Qualitative Approaches to Studying AI Adoption

It is important to note that almost all the cited research carried out in Pakistan as well as elsewhere uses structured and semi-structured interviews based on existing instruments used for measuring technology acceptance (Barra et al., 2024; Wardat et al., 2023). Although this methodology allows comparison between various samples, it forces the interviewees to fit their experiences into already known frames, which is inappropriate for studying the technology which is still being constructed by the respondents. Reflexive thematic analysis by Braun and Clarke (2019), which regards the researcher's interpretation as an asset rather than a threat to research quality provided that it is done transparently, seems to be ideally suitable for such a topic. The emerging qualitative literature on the use of GenAI in Pakistan (Majeed et al., 2025; Zaman & Malik, 2026) already shows the emergence of patterns of advantages and disadvantages which are similar to those reported in surveys, but provides more flexibility for showing how students themselves resolve the conflict between them. However, something that has not been done yet and is planned to be done in this research is to ask this question specifically about the female students in the women-only educational institution and in different disciplines.

Theoretical Framework: The Extended Mind Thesis

The analysis of the data in this study is conducted via the framework of the Extended Mind Thesis of Clark and Chalmers (1998), which argues that cognitive activity could take place outside the brain, in terms of access to some external tools or resources that are readily available, reliable, and trustworthy enough to be used in an automatic way. Clark and Chalmers refer to the above criteria as the parity principle. The chosen theoretical framework was preferred to the one based on a more general technology acceptance theory because the focus of the research is to find out how the learner incorporates an external tool into her reasoning process instead of predicting her actions regarding adopting it. According to Dai et al. (2023), generative AI could act as an external cognitive tool for a learner, but only if she interacts with it in a way that helps to couple her reasoning to the output of the tool. In this study, the process of designing a prompt is not considered a secondary technical skill but is viewed as the means to establish the coupling, following the claims made by Ng et al. (2021) that AI literacy is the meta-cognitive competence that does not develop spontaneously. From this theoretical framework, this study received sensitizing concepts to interpret the participants' stories about generative AI without imposing predetermined content.

Theoretical Framework: Self-Regulated Learning

Another complementary theory comes from Zimmerman's (2002) self-regulated learning (SRL) theory, where learning is seen as a cyclic process involving foresight, performance, and self-reflection stages. Within such an approach, the learner is the one who is responsible for self-monitoring and regulating her strategies during all three stages of the process rather than receiving instructions passively. If the Extended Mind Theory explains how, using Generative AI technology, a student becomes an external resource, then the SRL theory gives the language for another goal of the research, namely, when and why a student needs to use scaffolding, when to stop, and how to assess her performance after the process. Such a theory is particularly relevant for studies of Generative AI because,

unlike a tutoring system, conversational AI gives the responsibility for the need for assistance and the decision to stop it to the learner herself.

Theoretical Framework: Cognitive Offloading

A third and more modern point of view is derived from the theory of cognitive offloading (Risko & Gilbert, 2016). Cognitive offloading is understood as the tendency of delegating the mental task to an external tool or aid in order to lower the cognitive efforts required to execute it – for instance, looking up information on the phone instead of trying to remember it, or increasingly, instructing the AI system to solve the problem instead of solving it on one's own. According to Risko & Gilbert (2016), such choices to offload are driven by individuals' metacognitive evaluations of their own skills, which could be either correct or poorly calibrated.

This model was designed to explain the phenomenon of the Over-Reliance Paradox, whereby the Extended Mind Thesis explains the benefit of using AI in terms of cognitive extension, while SRL by Zimmerman (2002) explains learners' monitoring of the scaffolding, and cognitive offloading explains the cost of it as substitution. This model is consistent with a line of modern research (e.g., Gerlich, 2025) where such concerns are associated with deterioration of critical thinking skills. Overall, all three models allow considering the benefit, its self-monitoring by the learner, and risk as one and the same process.

Method

Research Design

This study used a qualitative descriptive design (Creswell & Creswell, 2018). The decision to use this design was purposeful because of the need to create an interpretation that is close to the language of the participants without necessarily testing hypotheses or developing a theoretical framework from the data. Qualitative description offers enough rigor to answer a specified number of research questions, but it allows for flexibility by creating space for participants to bring up issues that the researcher may not have considered, especially since most of what will follow emerged voluntarily from the participants. The reason why semi-structured and in-depth interviews were used to generate data is based on the same rationale.

Setting and Participants

The study took place at the Lahore College for Women University (LCWU), which is a publicly funded and women-only university situated in Lahore, Pakistan. A purposive sampling method was used in recruiting participants because the purpose was to get information-rich cases that could answer research questions, not to have representative statistics (Creswell & Creswell, 2018). Criteria for recruitment included being enrolled in either the BS or MS program and using Generative AI tools regularly for academic purposes. Twelve students met the criteria and agreed to be part of the study: nine were from the Department of Education, and three belonged to the Department of Computer Science. Selecting students belonging to both these departments, since one was a humanities department and another a technical one, was an attempt to test whether discipline had any impact on the narrative shared by the participant. A small sample of the participants was chosen to make individual transcript analyses possible, instead of only aggregate interpretation of transcripts. Data were collected until saturation was achieved: the code log revealed 34 first-order codes in Interviews 1-8, and one more code related to culturally localized AI content came up in Interview 9, with no further codes coming up in Interviews 10-12. The latter interviews were kept because it was necessary to check if the pattern observed continued to hold. Participant details are presented in Table 1.

Table 1*Participant Profiles (N = 12)*

Code	Programme	Department	Semester	Interview Duration	AI Tools Used
P1	MS Education	Education	3rd	27 min	ChatGPT, Gemini
P2	BS Education	Education	7th	24 min	ChatGPT, MS Copilot
P3	BS Computer Science	Computer Science	5th	29 min	GitHub Copilot, ChatGPT, Gemini
P4	MS Education	Education	1st	22 min	Claude, Gemini
P5	BS Education	Education	5th	26 min	ChatGPT, Grammarly, Gemini
P6	BS Education	Education	3rd	21 min	ChatGPT, Perplexity AI
P7	MS Education	Education	3rd	28 min	ChatGPT, Perplexity AI
P8	BS Computer Science	Computer Science	7th	25 min	ChatGPT, GitHub Copilot
P9	BS Computer Science	Computer Science	5th	27 min	ChatGPT, GitHub Copilot
P10	MS Education	Education	Final	23 min	ChatGPT
P11	BS Education	Education	3rd	26 min	ChatGPT, Grammarly
P12	MS Education	Education	1st	30 min	ChatGPT

Note: All participants were female students at LCWU reporting regular Generative AI use in their academic work. Interview durations reflect audio-recorded portions only.

Data Collection

The data were collected from the respondents using the process of individual, semi-structured interviews through the medium of video conferencing, which lasted from 21 to 30 minutes. The questionnaire for the interviews (Appendix A) was prepared in the light of the research questions, and then it was pilot tested with one respondent not included in the actual data gathering. The questions in the interviews were open-ended, and they were followed by prompts from the researcher to obtain detailed responses instead of facts only. All interviews were carried out in English language, Urdu language or both according to the choice of the participants and were tape recorded after taking their permission. If Urdu was spoken, the researcher translated the particular parts into English and also a bilingual person translated randomly selected 20 percent of these parts independently in order to check the accuracy of the translation.

Data Analysis

The transcripts were analyzed following the six-stage reflexive thematic analysis approach of Braun & Clarke (2006): familiarization through multiple readings and listening; identification of initial codes from throughout the dataset; organization of the codes into possible themes; review of the themes in relation to both the coded material and the entire dataset; definition and labeling of the themes; and finally, construction of the analytical narrative detailed in this report. When the theme contained two or three unique internal patterns rather than a single concept, subthemes were identified below the theme, but this only happened rarely and only when the differentiation was inherent in the data rather than artificial in nature, in accordance with the cautionary note of Braun and Clarke (2019) against forced breakdown of themes. Manual coding took place and was documented using a structured spreadsheet,

providing an auditable chain of reasoning from transcript to theme. The three research questions provided a broad deductive structure for theme identification, although the actual coding was an inductive process, based on the participants' own language use. This allowed identification of the Over-Reliance Paradox, an issue which had not been covered in any interview question.

Trustworthiness

Rigor was assessed in terms of four criteria developed by Lincoln and Guba (1985): credibility, transferability, dependability, and confirmability, rather than validity and reliability, which would be more applicable to quantitative studies. The issue of credibility was approached from three directions: member checking, where the researcher shared her preliminary thematic findings with four chosen participants who were selected to represent both departments as well as different levels of enthusiasm regarding artificial intelligence; using participants' actual language in presenting the results; and looking for inconsistencies within the main pattern rather than for examples that would support it. Transferability was assured by providing a detailed description of both the setting and the participants to allow readers to evaluate the relevance of the findings in other situations. Dependability required documentation of the process of analyzing the data and decision-making during the coding process. Confirmability included keeping a reflective journal throughout the data collection and analysis stage, which always asked for identifying and describing non-fitting cases as explained in greater detail below.

The researcher also used triangulation in assessing credibility, following the classic distinction made by Denzin (1978) between data, investigator, and methodological triangulation. Data triangulation came from sampling two structurally different departments so that a finding such as the Over-Reliance Paradox could be identified only when found in both Education and Computer Science participants' accounts and not just in one department. Investigator triangulation was achieved in the process of member-checking described above, as participants were checking the interpretation of data by the researcher. Methodological triangulation was provided through the use of participants' verbatim transcripts and the researcher's analytical memos as well as independently back-translated segments of Urdu speech. Taken together, these sources allowed ruling out any potential for bias towards the departmental, interview or researcher-driven artifact.

Ethical Considerations

Ethical approval for the study was obtained from the relevant body in LCWU before conducting fieldwork. All the participants gave informed consent once they were briefed on the goals of the study, their right to withdraw at any time without being penalized, and the methods of ensuring confidentiality. The confidentiality of the data was ensured by giving code numbers (P1-P12). No participant's name or any other identification appears in this report. The recordings and transcripts have been stored in an encrypted device, which only the researcher could access.

Findings

Themes identified from the twelve interviews included three main themes relating to each research question, a case that diverges from the pattern established, and a final theme that arose from all three research questions rather than answering a specific one. The following table highlights these themes before proceeding to a full discussion, including illustrative quotes, after which point.

Table 2*Overview of Themes*

Theme	Research Question	Brief Description	Participants Evidencing Theme
Theme 1: Generative AI as a Scaffolding Tool for Personalized Learning	RQ1	Students actively adjust prompts (level, discipline, local examples) to draw richer, more tailored explanations out of AI tools.	12 of 12
Theme 2: Perceived Effectiveness and Institutional Ambivalence	RQ2	AI is seen as highly effective, but the absence of clear university policy leaves students uncertain how far its use is officially sanctioned.	12 of 12
Theme 3: AI as a Psychological and Academic Enabler	RQ3	AI lowers the emotional threshold for starting and persisting with difficult academic tasks.	11 of 12
Divergent Case: Selective and Cautious Engagement	RQ2 / RQ3	One participant's credibility-damaging experience led to durable, restricted, low-trust AI use.	1 of 12 (P4)
Emergent Theme: The Over-Reliance Paradox	Cross-cutting (RQ1-RQ3)	The students most positive about AI's benefits are also the most articulate about its risks to independent thinking.	7 of 12

Note: "Participants evidencing theme" reflects the number of transcripts in which the theme was clearly present during coding, not the number of participants directly asked about it.

The narrative account below presents each theme in turn, illustrated with participants' own words and followed by the researcher's interpretation of what that excerpt contributes to the theme, consistent with best practice in reflexive thematic analysis (Braun & Clarke, 2006).

Theme 1: Generative AI as a Scaffolding Tool for Personalized Learning

Subtheme 1a: Discipline-Specific Patterns of Tool Use

The interviewees, regardless of which departments they were from, were all engaging in continuously adapting their prompts to Generative AI to fit their comprehension level, disciplinary context, and level of detail needed. This adaptation was a constant negotiation process similar to Clark and Chalmers's (1998) description of a learner's coupling of her thinking process with an external trusted source. The Computer Science interviewees usually use various tools in one workflow process, often using GitHub Copilot and ChatGPT together. In comparison, the Education interviewees use conversational AI mostly for literature synthesis and lesson planning.

Subtheme 1b: Prompting as an Emerging Academic Literacy

A number of participants in both departments reported a specific moment of understanding only in cases where the AI's explanation was set in a locally familiar context, such as one in a Pakistani classroom setting, rather than a textbook-level example.

"Well, so when I say, like 'explain this at MS student level,' the answer is just... totally different. More theoretical, more complex, you know?" (P1, MS Education)

"When Constructivism was explained step-by-step with a Pakistani classroom example, the concept just made sense." (P1, MS Education)

"It took me like three days to get a hang of Data Structures honestly. However, ChatGPT provided three different ways of solving them with time complexity and... that really helped a lot." (P3, BS Computer Science)

Overall, all the quotes above reflect a process which is not as passive as the use of the tool. Specifically, P1's two quotes highlight the same participant who recognized two specific levers of engagement that provided a different kind of scaffolding from the same tool. Therefore, this pattern suggests that Generative AI served not as a pre-existing source of knowledge, but as a responsive conversation partner whose effectiveness depended on how efficiently and precisely it was used. Under this lens, the act of prompting becomes an emerging academic literacy in its own right.

Theme 2: Perceived Effectiveness and Institutional Ambivalence

Subtheme 2a: Effectiveness Without Embarrassment

Each participant felt that the Generative AI was truly useful in their academic research, with one participant giving a definition that is most appropriate to explain the reason behind this view:

"It's just like a 24/7 personal tutor with no shame."

Embarrassment is an important keyword here, as the value of AI technology is not the fact that it knows the facts but in the reduction of the need for a human presence while being uncertain. This feeling has been expressed by participants in the field of Computer Science using disciplinary language:

"It is really good, like pair programming with a senior developer around always." (P3, BS Computer Science)

Subtheme 2b: The Governance Gap

The effectiveness was, however, accompanied by evident worries about the lack of any institutional direction. All twelve respondents, on their own, mentioned the need for an official AI policy at the university. Although details of their suggestions differed depending on departments due to different functions AI performs in the context of computer science classes versus education courses, they all highlighted the importance of having a well-defined policy. As one of the respondents put it:

"It's important that LCWU should develop an AI use policy rather than a ban for each assignment type." (P1, MS Education)

"Integrate AI officially into CS labs with clear policies; banning is not the solution." (P3, BS Computer Science)

Subtheme 2c: Cultural and Contextual Relevance

The second general issue is the issue of cultural disparity in AI applicability, the perception that AI applications were geared towards contexts other than the one of this educational institution. *"AI must make use of culturally relevant data. Western-oriented examples are not applicable to the context of Pakistani education."* (P8, BS Computer Science)

"AI must consider the current education policy in Pakistan, like NEP 2017 and SNC, to be applicable to us." (P6, BS Education)

All in all, this theme shows that the participants' lack of trust in AI had nothing to do with the reliability of the tool itself, since all participants had some level of trust in the technology; the lack of guidance regarding how far this trust could go was the core issue.

Theme 3: AI as a Psychological and Academic Enabler

All eleven out of the twelve respondents experienced concrete benefits in their studies due to the utilization of AI. Students enrolled in education courses were more likely to mention improvements in their literature reviews, whereas those studying computer science mentioned improvements in laboratory work. Nevertheless, regardless of which department students belonged to, the most prevalent benefit that students mentioned was not content-related but rather affective in nature. This benefit was about the experience of starting a difficult task.

Subtheme 3a: Overcoming the Fear of Starting

"I used to just... avoid starting assignments, because I didn't know where to begin, you know? Now I ask AI for a starting point and the fear kind of just disappears." (P2, BS Education)

Subtheme 3b: Sustaining Persistence Through Difficulty

The psychological point of intervention in P2's account is at the hinge point of initiation, which is where the intervention of AI happens. On the other hand, in P3's account, the role of AI is different but in line with P2 in that it affects persistence after engagement with the task. In both cases, what the participants gained academically through the use of AI was not so much the information gained from it but the influence it had on their ability to initiate and persist in hard work.

A Divergent Case: Selective and Cautious Engagement

However, not all participants followed the generally positive trend discussed above, and one such participant stands out entirely and requires description for the reason of the complicating impact she has made on the research findings. Specifically, P4, who is an MS Education student in her first semester, expressed a significantly more cautious attitude toward Generative AI than did other participants.

"Last semester I used AI to help summarize a research article for my literature review, and it just... completely misrepresented the author's argument. I didn't catch it until my supervisor pointed it out in front of the whole class. That was embarrassing, and honestly, I don't fully trust it with anything conceptual anymore." (P4, MS Education)

After this incident, P4 indicated that she would limit her use of AI to only those instances involving low-stakes tasks such as grammar checks and rephrasing of sentences but would never use it for anything requiring analysis or concept. In contrast to the other eleven participants who supported institutional integration of AI, P4 suggested that in any institution policy regarding the use of AI, there should be "probably a warning to the students about how wrong it [AI] can sound." This illustrates that the positive inclination seen in Themes 1-3 is not universal, and one credibility-damaging incident can permanently affect the student's perception of the tool regardless of discipline.

Emergent Theme: The Over-Reliance Paradox

One pattern surfaced across seven of the twelve interviews, spanning both departments, without ever being asked about directly. The participants who described the greatest academic benefit from AI were, without exception, also the ones who spoke most critically and precisely about its risks, a pairing this paper terms the Over-Reliance Paradox: enthusiasm and wariness held at the same time, by the same people, rather than trading off against each other.

"I do worry that over-reliance might weaken critical thinking. Like, if AI always gives the answer, then... when do students actually develop their own analytical skills?" (P7, MS Education)

"Students submit AI-generated code without understanding the underlying logic. That defeats the purpose entirely." (P3, BS Computer Science)

What stands out about these narratives is not a lack of clarity but a sense of coherence. The participants P7 and P3 do not seem to be struggling with two opposing views but rather keeping up both viewpoints and using this understanding in order to devise practical ways to cope with it. The practical ways of coping with it are trying to solve the problem on one's own before turning to AI or asking for an explanation of why AI chose a certain answer. This situation means that the institution policy in such circumstances needs to rely on an existing self-regulating attitude of many students.

Discussion

The purpose of this research is to explain how the process of Generative AI use takes place among female university students in Pakistan, not just measure usage rates. The three research question-derived themes and a divergent case, together with a cross-cutting emergent theme, allow for a richer interpretation of the phenomenon at hand compared to the survey-based literature that has primarily been used for the research of the area thus far.

Theme 1 contributes to the understanding of AI as a generative extended mind, that is, a cognitive system which is an extension of the individual outside of the human body (Dai et al., 2023). By explicating the way the scaffolding relationship actually works, it shows that the students do not receive help; they generate it using particular prompting techniques. In that sense, the present findings agree with the conclusions made by Tlili et al. (2023) that carefully developed prompts result in more valuable outputs; however, they also contribute to this idea a certain level of specificity that their research does not possess: these participants understood that in order for the example to be relevant and clarify the concept properly, it had to be situated in a Pakistani setting.

Theme 2, secondly, matches with the findings of Chan and Hu (2023) in that students tend to prefer formal integration of AI within the curriculum and those of Perkins (2023), in which Perkins states that the majority of the issues related to integrity are caused by confusion about the policy rather than any malintent. This study takes those results further to demonstrate both the level of unity among students with regard to that demand for a policy and its specific nature in relation to the individual departments: all 12 students spontaneously brought up that issue and made their particular suggestions in accordance with their departmental affiliations.

Thirdly, Theme 3 states a mechanism behind the idea of Kohnke et al. (2023) that anxiety relief, rather than mere accessibility of information, is what drives engagement gains enabled by AI. In the current case, this mechanism can be identified very specifically right at the very start of the task, before any information was generated, making it a more precise statement compared to what is currently known about engagement gains in general. Yet, there is a limitation here, which is that one negative experience with AI, such as that provided by P4 regarding a factual mistake that makes AI less credible, could overwhelm any anxiety-relief effects and drive students to narrow their AI usage down considerably.

Moreover, and finally, the Over-Reliance Paradox identifies an open question from the existing literature as one that the participants themselves have settled for themselves. The use of artificial intelligence in regard to critical thinking is identified by Kasneci et al. (2023) as a problem needing further investigation; however, these participants themselves have already engaged with a similar question during their own academic journeys, and in a manner more nuanced than can be encapsulated in the term "open question." This contrast between the open question designation of the literature versus the prior conclusion reached by the participants suggests a particular pragmatic action: the incorporation of reflective questions, or metacognitive exercises, into the AI-based assignment would formalize and

foster a previously held skill among the students, rather than viewing them as a demographic that needs to learn to think before acting.

Fifth, these results do not support but complicate the existing body of knowledge on gender and technology adoption. Barra et al. (2024) show that the use of technology by Pakistani women is driven mostly by social pressure and not by technical expertise; yet, the subjects of this research, all students of a women-only college, do not describe their AI adoption in this manner. On the contrary, the subjects see their AI skills as something self-made and used for academic goals, not something adopted due to the need to earn social approval. This shows that the institutional setting of being in a single-sex learning environment plays an important role in mediating the effects of gender on technology adoption, in shaping an environment in which a student's interaction with technology happens in a way defined by her academic identity and not gendered social pressure seen in mixed-gender studies. An experiment performed in one institution alone cannot prove this statement; instead, it puts forward a hypothesis for further testing.

Inevitably, the method that is used in this research is tightly tied to the results that are obtained. Research through surveys of the introduction of GenAI in Pakistan (Majeed et al., 2025; Zaman & Malik, 2026) has found very similar categories of advantages and disadvantages. However, the difference in the current case is that a reflexive thematic analysis was applied (Braun & Clarke, 2019), which allowed an understanding of the contradiction as data and not as noise, which resulted in statements like P4 or the phenomenon known as the Over-Reliance Paradox.

Theoretical Implications

As evidenced by the results, two additional implications of the Extended Mind Thesis can be put forth based on this study. First of all, the research has shown that the role of an external cognitive resource, which is worthy of trust, is not something intrinsic about the tool itself but something that is actively constructed by the learner through her practices of making prompts. Namely, the participants who made more specific prompts by specifying their academic level, subject and asking for locally situated examples seemed to receive more sophisticated scaffolding as a result. Thus, this limitation refines Dai et al.'s (2023) conceptualization of AI as a cognitive partner by putting some of the responsibility for the reliability of the coupling onto the learner's prompting skills.

The second implication of the Over-Reliance Paradox lies in the fact that the standard parity principle does not capture one of the aspects of the learner's behavior: namely, she seems to maintain the coupling of two systems at once. The first one is an academic system which utilizes the help of AI in order to enhance her content knowledge, while the second one is a self-regulatory system that is constantly evaluating whether the current situation still meets the requirements of trust and reliability defined by the extended-mind thesis. The notion of dual-coupling has not been explicitly stated before in extended-mind discussions of the use of AI in education (Dai et al., 2023; Ng et al., 2021), yet it seems to be useful for the discussion of how an AI-saturated environment helps students form independent judgment of the external resource's trustworthiness. Furthermore, when related to the stages of the self-regulated learning (SRL) cycle by Zimmerman (2002), the second self-regulatory layer corresponds to the self-reflection stage quite well.

A further extension arises where the phenomenon of extended mind dynamics is brought into relation with the existing research on the topic of gender and technology uptake (Barra et al., 2024). If trust in the reliability of the external cognitive tool is not the only factor determining the degree to which a learner is ready to use it as the part of an extended mind process, but it is also dependent on the social permission and confidence she brings into the situation, then the all-women setting observed in this case might operate as the socially extended coupling, i.e. a situation where the gender-related hesitations seen in the other Pakistani technology studies are diminished, making

the learners feel comfortable working with AI as the confident authors of their cognitive extension, rather than hesitant participants constrained by societal norms. This would be an inductive claim, drawn from the description of agency given by the participants themselves, that extends the initially individually-based framework of Clark and Chalmers (1998).

The fourth extension involves viewing the Over-Reliance Paradox in light of the theory of cognitive offloading (Risko & Gilbert, 2016), instead of just through the lens of the Extended Mind Thesis. According to Risko and Gilbert (2016), the decision about offloading a mental task is based on a metacognitive evaluation of one's abilities, and the evaluation can either be accurate or inaccurate. Using this framework, the seven individuals who expressed the Over-Reliance Paradox exhibited signs of well-calibrated offloading judgment, relying on AI in situations when it could do more good, and being attentive to situations when the use of AI could replace their thought processes. On the other hand, P4's exceptional case does not seem to be a display of caution for its own sake but rather a one-time mistake in the process of metacognitive judgment: one-time offloading of a conceptual task to the artificial intelligence was costly for her, leading to a more cautious offloading approach in future.

Limitations

Similar to qualitative research in general, the current results are meant to be transferable and contextually rich but not statistically generalizable (Creswell & Creswell, 2018). The study was carried out at one all-women institution, which may affect the ability to transfer its results to co-educational institutions or male students. Even though the sample size was sufficient to reach thematic saturation, the study included only two departments, which means that it may have left out experiences of students from other departments and may need to be supplemented by such information. In addition, because the sample included nine students from Education and three students from Computer Science, observations related to the latter discipline were based on relatively weak evidence. Finally, even though a colleague who was bilingual verified the quality of some translations from Urdu into English, she verified just 20% of all the translations, and the remaining ones were done by the researcher herself.

Future Research

First, there are three paths for future investigation that can be clearly drawn from the limitations mentioned above. To begin with, a comparison between qualitative research carried out at both the co-educational and the single-sex universities will reveal whether the agentic, self-regulating approach found here is particular to the latter university or rather common to all Pakistani educational institutions. Secondly, by using a longitudinal design where the same subjects are followed throughout several semesters, we can find out whether the Over-Reliance Paradox is a fixed tendency or whether it changes with the development of policies regarding artificial intelligence in the university. Thirdly, P4's case requires further investigation on the level of the general phenomenon.

Conclusion

The present study examines how female university students in Pakistan personally perceive Generative AI in the context of personalized learning, and it reveals a group of users that exhibits self-awareness higher than expected according to either the optimistic or the cautionary branch of prior scholarship. Participants saw AI as a scaffold tool whose worthiness depends on purposeful and culturally nuanced prompting but, simultaneously, without any prompting, articulated a critical evaluation of the danger that this tool poses to their capacity for independent thought. The combination represents the unique contribution of the present study, called the Over-Reliance

Paradox: the existence of a cohort of Pakistani university students who use Generative AI proactively in ways that current institutional policies did not account for.

In terms of policy implications, this means that the guidance emphasizing prohibition attempts to answer questions that these students had already passed. Pedagogically speaking, the conclusion from the present research implies that the self-regulatory mindset that is inherent to many students could be cultivated through the course design without being introduced anew. As far as future research is concerned, it suggests that further progress does not mean carrying out another survey on the matter, but a more focused investigation into the extent to which this trend persists or develops over time and in other settings.

References

- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Barra, C., Grimaldi, M., Muazzam, A., Troisi, O., & Visvizi, A. (2024). Digital divide, gender gap, and entrepreneurial orientation: How to foster technology adoption among Pakistani higher education students? *Socio-Economic Planning Sciences*, 93, Article 101904. <https://doi.org/10.1016/j.seps.2024.101904>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589-597. <https://doi.org/10.1080/2159676x.2019.1628806>
- Chan, C. K., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00411-8>
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19. <https://doi.org/10.1093/analys/58.1.7>
- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00392-8>
- Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP*, 119, 84-90. <https://doi.org/10.1016/j.procir.2023.05.002>
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. McGraw-Hill.
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELJ Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>

- Majeed, D., Khan, D. Y., Munir, M., Shan, R. U., Khalid, H., Khan, K., Aurangzeb, T., & Atique, E. (2025). Opportunities and challenges of generative AI in Pakistani higher education: A qualitative study on student perspectives in learning, integrity, and innovation. *Nature-Nurture Journal of Psychology*, 4(2), 30-52. <https://doi.org/10.53107/nnjp.v4i2.89>
- Ng, D. T., Leung, J. K., Chu, S. K., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676-688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1). <https://doi.org/10.1186/s40561-023-00237-x>
- Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286. <https://doi.org/10.29333/ejmste/13272>
- Zaman, S., & Malik, S. (2026). Understanding the perceptions of higher education students about personalized learning through generative artificial intelligence. *Gomal University Journal of Research*, 42(1), 124-141. <https://doi.org/10.51380/gujr-42-01-10>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S., Istenic, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021(1). <https://doi.org/10.1155/2021/8812542>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Appendix A

Semi-Structured Interview Guide

Opening / Rapport-Building

1. Can you tell me a bit about your program of study and how far along you are?
2. Which Generative AI tools do you currently use, and how often would you say you use them for academic work?

RQ1: Describing and Making Sense of AI Use

3. Can you walk me through a recent time you used AI for a university assignment?
4. When you're using AI for academic work, how do you decide what to ask it, or how to phrase your request?
5. Has the way you use AI changed since you first started, and if so, how?

RQ2: Perceived Effectiveness and Concerns

6. How effective would you say AI tools have been for your academic work? Can you give an example?
7. What concerns, if any, do you have about using AI for your studies?
8. If your university introduced an official policy on AI use, what would you want it to say?

RQ3: Academic Performance, Engagement, and Confidence

9. How, if at all, has using AI changed the way you feel about difficult academic tasks?
10. Can you describe a time AI helped you get "unstuck" academically? What did that feel like?
11. Do you ever worry about relying on AI too much? Can you tell me more about that?

Closing

12. Is there anything about your experience with AI in your studies that we haven't talked about, but that you think is important?

Note: Questions were followed by spontaneous probes (e.g., "Can you tell me more about that?", "What did that feel like?") to elicit depth. The guide was piloted with one participant outside the final sample prior to use.