ISSN (Print): 2788-8320 ISSN (Online): 2788-8339

JOURNAL OF SOCIAL RESEARCH DEVELOPMENT
www.jsrd.org.pk editor@jsd.org.pk

GENERATIVE AI DEPENDENCY AND ACADEMIC PERFORMANCE OF UNIVERSITY STUDENTS: THE MEDIATING ROLE OF CRITICAL THINKING

Ahmad Ali[1], Naeem Shahzad [2], Hira Shahbaz[3], Adeen Khan [4]

[1]PhD Scholar, Institute of Education and Research, University of the Punjab, Lahore, Pakistan
Email: ahmadfarzand34@gmail.com
[2]Monitoring Officer, Literacy and Non-formal Education Email: ali911ahmad 4@gmail.com
[3]PhD Scholar, Institute of Education and Research, University of the Punjab, Lahore, Pakistan
Email: hieerbaryar @gmail.com
[4]Instructor, Virtual University, Lahore, Pakistan Email: adeenkhanofficial05@gmail.com

KEYWORDS

Generative AI, ChatGPT; AI dependency; Critical thinking; Academic performance, Lahore, Pakistan.

ARTICLE HISTORY

Date of Submission: 04-11-2025
Date of Acceptance: 01-12-2025
Date of Publication: 02-12-2025

ABSTRACT

Generative artificial intelligence (AI) tools such as ChatGPT have spread rapidly through Pakistani universities. Many students now use them every day, which raises the question of whether dependence on these tools affects how well students think and perform. This study examined the effect of generative AI dependency on the academic performance of university students in District Lahore, and whether critical thinking mediates this effect. A quantitative, survey research design was used. Through multistage stratified random sampling, 384 undergraduate students were selected from four public and four private universities. Data were collected with a 12-item Generative AI Dependency Scale adapted from Morales-García et al. (2024), a 14-item Critical Thinking Scale adapted from the Critical Thinking Disposition Scale of Sosu (2013), and self-reported cumulative grade point average (CGPA). Both scales showed high reliability (α = .920 and .916). Most students (60.9%) showed moderate AI dependency and 21.1% showed high dependency. Students with high dependency had significantly lower critical thinking (M = 2.57) and CGPA (M = 2.87) than students with low dependency (M = 3.40 and 3.19). AI dependency negatively predicted CGPA (B = −0.138, p < .001) and critical thinking (B = −0.378, p < .001), while critical thinking positively predicted CGPA (B = 0.283, p < .001). Critical thinking fully mediated the effect of AI dependency on CGPA (indirect effect = −0.107, 95% bootstrap CI [−0.145, −0.075]), explaining 77.6% of the total effect. Dependency increased with daily hours of AI use, but no significant differences were found by gender or university sector. The study concludes that AI dependency appears to lower academic performance mainly by weakening critical thinking, and recommends AI literacy training and assessment that requires students' own reasoning.

2025 Journal of Social Research Development

Correspondence: Ahmad Ali
Email: ahmadfarzand34@gmail.com

DOI: https://doi.org/10.59075/jsrd.v6i4.666
Volume/Issue: 6(4), 2025, 60-90

Introduction

Background of the Study

In November 2022, ChatGPT was released, causing a paradigm shift in the world of education. By a few months later, millions of students were demonstrating the use of generative AI tools to explain complex ideas, to summarize reading, to write essays, to write code, to translate texts, and to solve problems. Unlike previous digital tools, which primarily were used to locate information, generative AI creates new content given a simple request. A student is now able to get an entire essay outline, a worked solution and a fine paragraph in a matter of seconds. These tools have definite advantages for learning. They can provide immediate feedback, explain ideas in simple terms, offer practice questions and help students who don't have the means for private tutors. Kasneci et al. (2023) suggest that LLM technologies could enhance learning personalization and accessibility, notably for students who face less support in the classroom. These options are especially appealing in nations where classrooms are big and teachers do not have sufficient time to provide one-on-one attention.

The very convenience also poses a risk. If a tool can give the answer in a flash, then students might not do the thinking they need to learn. This is called "cognitive offloading" which is defined as "a person relying on an external aid to reduce the load on their own memory and thinking" (Risko & Gilbert, 2016). Individuals have come up with some methods of delegating responsibilities, such as writing shopping lists, using calculators, etc. The worry about generative AI is that students can now delegate the very skills that they are supposed to learn, like analyzing an argument, evaluating evidence, and even creating a solution. Zhai et al. (2024) also found that systematic reliance on the AI dialogue system might diminish the students' decision-making capabilities, critical thinking and analytical reasoning, in their systematic review. In a similar study, Gerlich (2025) also discovered a negative correlation between the use of AI tools and critical thinking.

Pakistan is not an exception in this regard. The population of Pakistan is one of the youngest in the world and the Higher Education sector in Pakistan has grown a lot with the Higher Education Commission (HEC). University students can use AI tools widely through their smart phones and mobile internet and they can use free versions of ChatGPT, Gemini, Copilot, Grammarly and QuillBot. The evidence of use gathered from surveys illustrates the rapid increase in use. According to a study conducted at a female university of Bahawalpur, 354 students were asked about the frequency of their ChatGPT usage, with approximately 51% stating that they use it regularly (Zafar et al., 2024). A total of 82% of the 300 undergraduates from six public universities in Punjab had used at least one AI tool, of which 64% used ChatGPT (Shahzad & Farooq, 2025). The results of the national survey of 400 students across all four provinces showed that most students found ChatGPT to be helpful, while also being aware of its issues of accuracy, academic integrity and reliance on AI (Alghazo et al., 2025).

Lahore features prominently in this image. District Lahore is the capital of Punjab and is one of the most prominent educational hubs of Pakistan, its high education centers and universities draw students from all over the country. The universities of Lahore, being an important district for

educational purposes, have a varied range of resources, class sizes, teaching methods and assessment practices, providing a helpful context for comparing students between different sectors. There are studies taking place in Lahore that already have noticed the uses of ChatGPT by students. In the study by Mehran and Farooq (2025), the undergraduates at Government College University Lahore reported that ChatGPT contributed positively to their academic learning and motivation; however, the users of ChatGPT also reported that the daily use of ChatGPT negatively impacted on the usage of traditional resources like books and libraries. The study compared the correlation of self-efficacy with student engagement among frequent and infrequent users of ChatGPT for students in four universities of Lahore with 297 participants who were undergraduate students (Mujtaba & Ali, 2025).

Meanwhile, there has been an ongoing concern regarding the lack of critical thinking in education system in Pakistan. Pakistan education system has been termed as examination-oriented and rote memorization based. Jamil and Muhammad (2019) have reported that secondary school Science teachers in Punjab wanted to focus primarily on securing good examination grades and one teacher in their study estimated that only 10-15% of the test content was geared towards conceptual understanding. In the university level Bibi and Akhter (2020) administered the Watson–Glaser test on 160 prospective teachers of three universities of Lahore, and found that around three-quarter of the students had scores below 50%, especially in the inference, deduction and analysis of arguments items.

Combined, these two factors present a big problem. Students with low levels of independent thinking may be particularly drawn to generative AI, as it can be a way to get around tasks that seem challenging. They may fail to gain the opportunities to learn to think critically which the university education offers if they rely on AI for analysis and writing. This weaker critical thinking could then impact their studies, particularly in exams and other evaluations where AI is not permitted. It is crucial that teachers, university leaders and policymakers in Pakistan who have to make decisions about the use of AI in higher education understand this chain of effects.

Statement of the Problem

Generative AI has entered Pakistani universities faster than policies, teaching methods and assessment practices have been able to adapt. AI adoption among students in Punjab is high, with many reporting increased reliance on these tools over time, and many already concerned about depending on it (Shahzad & Farooq, 2025; Zafar et al., 2024). However, there is no clear evidence in universities about the positive and negative impacts of this dependency on learning. This uncertainty has real-world implications. Some teachers prohibit the use of AI, some teachers accept the work created by AI without a doubt, and students receive conflicting messages about the concept of responsible use.

The current evidence in Pakistan is conflicting. In some cases, students feel that AI enhances their learning and productivity (Qureshi et al., 2024; Shahzad & Farooq, 2025). Others claim that the use of ChatGPT is linked to procrastination, memory loss, and decreased academic achievement (Abbas

et al., 2024). This variation could be due to the fact that many studies have only measured the frequency of student's use of AI, not how reliant they have become. It's not the same to use AI occasionally for the purpose of learning as it is when students become dependent on AI to start or finish a task. There are not many studies in Pakistan that have estimated dependency directly.

Another challenge is many of the studies focus on direct impacts, like the correlation between AI usage and grades. They have no idea why it is that there's such a link. Results from international studies indicate that using AI can negatively impact critical thinking (Gerlich, 2025; Zhai et al., 2024), which is considered a key academic competency for university success (Halpern, 1998). It is particularly relevant in Pakistan to explore if the performance of the university students is affected by AI dependency or not due to its impact on CT, because CT is already a concern among the university students in Pakistan (Bibi & Akhter, 2020).

Moreover, there is limited knowledge regarding whether there is a difference in the dependency on AI between different groups of students in Lahore, including students in public universities, female students, and students with varying hours spent with AI per day. Private universities may have a higher availability of technology, and public universities may have larger classes, which may affect students' behavior with AI. The problem addressed by this study is therefore the lack of evidence on how generative AI dependency affects the academic performance of university students in District Lahore, whether critical thinking explains this effect, and whether students with different levels of AI dependency, and from different groups, differ in critical thinking and academic performance.

Research Objectives

The research objectives of this study were given as follows:

1. To determine the level of generative AI dependency among university students in District Lahore.
2. To examine the effect of generative AI dependency on students' academic performance.
3. To examine the effect of generative AI dependency on students' critical thinking.
4. To examine the effect of critical thinking on students' academic performance.
5. To examine the mediating role of critical thinking in the relationship between generative AI dependency and academic performance.
6. To compare critical thinking and academic performance of students with low, moderate and high levels of AI dependency, and to compare the study variables by gender, university sector and daily AI use.

Research Questions

The research study was consisted of the following research questions:

1. What is the level of generative AI dependency among university students in District Lahore?
2. What is the effect of generative AI dependency on students' academic performance?

3. What is the effect of generative AI dependency on students' critical thinking?
4. What is the effect of critical thinking on students' academic performance?
5. Does critical thinking mediate the relationship between generative AI dependency and academic performance?
6. Is there a significant difference in critical thinking and academic performance among students with low, moderate and high AI dependency, and do the study variables differ by gender, university sector and daily AI use?

Significance of the Study

This study will be significant for students, as it will help them understand how excessive dependency on AI tools may affect their critical thinking skills and academic performance. It will inspire students to leverage AI tools to support their own learning processes, not replace them. The study will also assist in understanding the level of dependency of university students on the use of AI and the impact of this on their critical thinking, which can benefit university teachers. The results can guide teachers to design learning, teaching and assessment strategies, which foster students' independent learning, reasoning and analytical abilities. Additionally, the research will contribute to the creation of effective policies related to the use of AI tools, academic integrity, and responsible integration of technology in higher education by university administrators. The research will also be relevant for the Higher Education Commission (HEC) and policymakers as the results will serve as local evidence to guide the decisions regarding AI literacy in Pakistani universities, educational policies, and assessment practices. Furthermore, teacher training institutions can leverage the results to encourage the responsible use of AI and enhance the critical thinking abilities of future educators. The research will also assist parents in understanding how over-reliance on AI could impact their students' ability to learn independently and grow in their studies. Moreover, the findings of this study could be valuable for future research in critical thinking as the variables between AI dependence and academic performance. It can also serve as a baseline measure for research on other factors like motivation, self-efficacy, and AI literacy.

Review of Related Literature

Generative AI in Higher Education

Generative AI refers to systems that generates new text, images, audio or code based on a prompt. The most common examples in educational use are large language models, like ChatGPT, Gemini and Copilot. Students use them to explain concepts, produce ideas, summarise a reading, to check grammar, write and debug code and prepare for examinations. Teachers use them in designing lessons, creating quizzes and giving feedback.

Kasneci et al. (2023) conducted a review on opportunities and challenges of Large Language Models (LLM) in education. Defining benefits included personalised learning, writing and research support, and student support for students with disabilities. They also noted issues like "reliance," "lack of

creativity," "bias in content created" and "academic dishonesty. They claimed that it can only be achieved if teachers and students acquire the skills necessary to critically leverage AI.

Empirical research has demonstrated that the impact of AI is highly dependent on its application. In a study of students in a technical university in Romania, Dogaru et al. (2025) showed that students who considered ChatGPT as helpful reported higher levels of motivation and competence. But strong dependence was linked with dependence and with weak recognition of the use of AI. In the U.S., Nasr et al. (2025) investigated student interactions with university academic libraries and found two types of interaction: passive, AI-directed interaction and collaborative, AI-supported interaction. The authors found that critical thinking was facilitated by only one collaborative pattern, and that the quality of the human–AI interaction is more important than the quality of the AI tool itself.

Lee et al. (2025) conducted a survey among knowledge workers on their working experiences with generative AI. The workers with higher confidence in AI indicated lower levels of critical thinking in their work, while the workers with higher confidence in their own abilities indicated higher levels of critical thinking in their work. In this study, students were not involved, but it proposes a general mechanism: people trust AI and they decrease their cognitive effort.

Generative AI Use in Pakistani Higher Education

In Pakistan, the study of generative AI has increased significantly since 2023. The majority of studies are survey-based with university students, with high usage and mixed findings. Zafar et al. (2024) conducted a study on 354 Master's and PhD students of Government Sadiq College Women University, Bahawalpur. Approximately half (51%) used ChatGPT on a regular basis. Students appreciated it for its speed of completing work, for customised instruction and for helping them with work at any time of day, but also feared it would stifle investigation, critical thinking and honesty in learning.

Qureshi et al. (2024) did a study on the BS students from the faculties of education in three public universities of southern Punjab. Although students rated their training in AI tools as low as 3.03 on a five-point scale, they considered ChatGPT to be a helpful tool to enhance their learning ($M = 3.96$). There were no significant differences between male and female students. The study indicates a high level of adoption of AI among students but with limited oversight.

Shahzad and Farooq (2025) conducted a study on 300 students who were selected from each of the six public universities in Punjab using stratified random sampling. A total of 82% reported using at least one AI tool, the most popular of which were Grammarly (75%), ChatGPT (64%) and QuillBot (58%). The use of AI was positively correlated with time management ($r = .47$), task completion ($r = .41$) and self-reported academic performance ($r = .39$). Despite this, the authors pointed out their worry with relying too heavily on AI-generated content.

Alghazo et al. (2025) conducted a survey and measured physical activity among 400 students of all four provinces of Pakistan. The majority of students found ChatGPT beneficial, with more perceived

benefits among private university students than among government university students. Accuracy, academic integrity and dependence on AI were the main concerns. The majority of respondents said they checked ChatGPT's results with other sources, but this could be an exaggeration of actual checking.

There have been a few studies carried out in Lahore. Mehran and Farooq (2025) conducted a survey with 226 students studying under graduation at Government College University, Lahore. The overall rating of ChatGPT for academic learning and motivation was high and low for accuracy and reliability, respectively. Those who used a resource daily reported the greatest decrease in the resources they used and were least inclined to consider academic dishonesty as a risk. Mujtaba and Ali (2025) have conducted a study on 297 students of four universities in Lahore. Users who used ChatGPT more often showed higher self-efficacy and engagement in the online environment, and lower engagement with peers and with the classroom environment, than infrequent users. Only frequent users endorsed a relationship between self-efficacy and engagement.

Direct concerns have also been pointed out in other Pakistani studies. In a study conducted on the university students of Pakistan, Abbas et al. (2024) determined that heavy academic workload and time pressure acted as good catalysts for the use of ChatGPT, while ChatGPT usage was linked with procrastination, memory loss, and lower academic performance. Shah et al. (2024) conducted a survey of students in Sindh to produce the argument that the over-dependence on ChatGPT might lead to a decline in deep learning, but the conclusions of this survey were made with a small sample size.

The Pakistani studies combine into three patterns. The first is that AI usage is prevalent within university students. Secondly, students have a positive idea of how AI can be useful. Third, there is a concern from both students and researchers that too much reliance is being made. However, frequency of use or perceived benefits are measured in most studies, not dependency, and few studies look at the mechanisms underlying the relationships between AI use and academic outcomes.

The Concept of Generative AI Dependency

Generative AI dependency refers to a pattern where students use generative AI to do tasks, they are able to do without it and feel unable or unwilling to do them without it. A dependency is not the same as a use. Using AI to edit grammar after having written an essay yourself is using a tool. If a student can't start an assignment without asking AI for an answer, they are being dependent.

One of the first university student dependences on AI scales has been developed by Morales-García et al (2024). They trained 528 students in Peru and created a five item, single factor scale that had good reliability ($\alpha = .87$). Its entries reflect a sense of being vulnerable without AI, a sense of being behind if using AI, an effort to keep up with AI, seeking AI validation for feeling good, and a fear that AI will take the place of one's abilities. Their research highlights the psychological as well as the behavioral aspects of relying on AI.

The present study builds on this work and takes the view that dependency has three characteristics. The first is behavioral reliance, which involves over-reliance on AI and the tendency to overuse it. The first is behavioral reliance -- over-dependence on AI, and overuse. The second is less effort on the part of students, as they take AI-generated work at face value and do not verify or edit it. The third one is psychological attachment, like when AI isn't around, there's anxiety or helplessness. Risko and Gilbert (2016) argue that the more readily available and reliable external tools are, the more likely offloading will happen. Generative AI is both and dependency can easily be formed.

The Concept of Critical Thinking

There are many definitions of critical thinking. A consensus of experts has identified it as purposeful, self-regulatory judgement that leads to interpretation, analysis, evaluation and inference, along with explanation of the evidence and reasoning (Facione, 1990). Critical thinking can be learned and adapted to different contexts, as students practice it purposefully, according to Halpern (1998).

Critical thinking has both dispositions and skills. Skills include analyzing arguments, judging the credibility of sources and drawing sound conclusions. Dispositions are open-mindedness, curiosity and questioning of one's own assumption. Sosu (2013) created the Critical Thinking Disposition Scale, which comprises of two dimensions: critical openness (i.e. openness to new ideas and evidence) and reflective skepticism (i.e. questioning and evaluation of information). Practices are the means by which skills and dispositions are developed.

Critical Thinking in the Pakistani Context

Critical thinking has long been a concern in Pakistani education. Jamil and Muhammad (2019) conducted interviews with science teachers of four public secondary schools of Faisalabad and discovered that teachers were employing question and group work techniques, but their primary goal was to secure better grades on the exams rather than to foster critical thinking. The focus of examinations was on memorizing more than understanding, and teachers had little motivation to vary from that.

Bibi and Akhter (2020) evaluated 160 prospective teachers using the Watson-Glaser Critical Thinking Appraisal at three universities in Lahore at university level. Approximately three-quarters of students made less than 50% and students' performance was poorest in making inferences, identifying assumptions, deduction and evaluation of arguments. The authors concluded that teacher education programmes were not effectively developing critical thinking. Manzoor et al. (2019) conducted a study with 712 prospective teachers in two public universities in Lahore and concluded that there were significant positive relationships between critical thinking disposition, problem-solving skills and problem-solving self-efficacy.

The findings from these studies indicate that many students in Pakistan have limited experience of independent thinking when they enter into and go through the university. This background is relevant to the current study. If critical thinking is already fragile, then a tool that allows students to

avoid thinking tasks could have a stronger effect than in systems where critical thinking is already well developed.

Academic Performance

Academic performance refers to the extent to which students achieve the learning goals of their courses. It is commonly measured through grades, examination scores or CGPA. In Pakistani universities, CGPA is the main indicator of academic standing and is used for scholarships, admission to further study and employment. CGPA summarises performance across many courses and assessors, although it is affected by differences in grading practices across institutions and disciplines.

Generative AI Dependency and Academic Performance

The relationship between AI and academic performance is not simple, and the Pakistani evidence reflects this. Shahzad and Farooq (2025) reported a positive correlation between the use of AI and self-reported academic performance, Abbas et al. (2024) reported negative correlation between using Chat-GPT and academic performance. It can be argued that the distinction between use and dependency can help resolve the differences in the findings. Purposeful use could assist the student in their approach to time and task management. However, dependency can result in reduced performance in assessment activities where AI is not available like invigilated exam, which still make up a significant part of assessment in Pakistani universities. Dependent students can also not immerse themselves fully in the content of the course even if their work is good.

Generative AI Dependency and Critical Thinking

If critical thinking develops through practice, then reducing practice should slow its development. Students avoid developing their own skill by asking AI to analyze a text or consider an argument. Gerlich (2025) identified a strong negative relationship between frequent use of AI and critical thinking, with cognitive offloading as a mediator, and the relationship was stronger for younger participants. In their systematic review, Zhai et al. (2024) similarly concluded that there is no evidence to support this. Lee et al. (2025) reported that individuals who were more confident in AI, were less likely to use critical thinking effort. In collaborative use, Nasr et al. (2025) demonstrated that final, resolution, and thinking stage was supported, while final and resolution stage in the passive and AI-directed use were not supported. Similarly, Pakistani students have pointed out the impact of relying on critical thinking excessively (Zafar et al., 2024).

Critical Thinking and Academic Performance

University assessment typically rewards analysis, argument and evaluation. Students with stronger critical thinking should therefore perform better. Halpern (1998) states that critical thinking is a key goal of higher education: it lets students use their knowledge to solve new problems. In Pakistan, a positive relationship exists between critical thinking disposition and problem-solving skills, self-efficacy, which positively influences academic success (Manzoor et al., 2019).

Critical Thinking as a Mediator

A mediator variable clarifies the relationship between an independent variable and a dependent variable, whether it is the way in which they are related or the reason behind their relationship (Baron & Kenny, 1986). As AI reliance decreases critical thinking, and critical thinking increases academic achievement, critical thinking could be the mechanism through which AI dependency impacts academic achievement. Gerlich (2025) showed that this type of mechanism can be tested, by demonstrating that cognitive offloading mediated the link between AI use and critical thinking. The present study builds upon this by analyzing the role of critical thinking in turn as a mediator between AI dependency and academic performance.

Theoretical Framework

The study is based on three complementary theories, which together explain why AI dependency might weaken critical thinking and, through it, academic performance.

Cognitive offloading theory. According to Risko and Gilbert (2016) cognitive offloading is the process by which a physical action or external tool is used to reduce the cognitive load associated with a task. An external aid which is available, reliable and less effortful than internal processing leads to people offloading. Offloading can help, as it will allow for mental resources to be freed for other tasks. It can be damaging, however, if the offloaded skill is the skill a person needs to work on. Generative AI is a very potent offloading machine and one that can replace reasoning, as well as memory or calculations, but writing. The theory therefore suggests that as easy access to AI increases, students will become more likely to outsource thinking and that these repeated acts of outsourcing will eventually turn into being overly dependent on AI.

Cognitive load theory. Sweller (1988) proposed working memory is a limited resource and that achieving the goal of learning is dependent on how this is used. Subsequent work identifies three types of load. Intrinsic load is due to the difficulty of the content. Extraneous load is due to an inadequate construction of instruction and does not support learning. Germane load is a measure of productive effort for developing understanding and enduring knowledge structures. Generative AI can be of service to remove unnecessary burden such as making confusing instructions clearer. If it also eliminates germane load, by performing analysis and reasoning for the student, it eliminates the effort that makes learning last. The theory is used to understand why a student might have successfully finished a task with AI but has little to no learning from it.

Sociocultural theory. Vygotsky (1978) proposed that learning takes place as a result of the interaction with a more knowledgeable other in a zone of proximal development, the difference between what a learner can do independently and with the help of another. From this perspective, AI might be seen as a better informed other, assisting learners (Nasr et al., 2025). The critical element is that the support should be slowly faded as the learner is able to perform the task—this is sometimes referred to as scaffolding. If the student is relying on AI for a lifetime, the scaffold is not phased out and the learner does not progress to independent functioning.

Integration of the theories. These theories put together imply a sequence. Students rely on AI, and this can be a form of dependency, due to the cognitive offloading theory. The cognitive offloading theory is a theory that is used to explain why students use AI and why it can be a form of dependency. The cognitive load theory can help to understand how dependency eliminates the germane efforts by which critical thinking can develop. According to the socio cultural theory, the support of AI is helpful if it is temporary and if the learners gradually take over the thinking themselves. The performance in university is dependent upon critical thinking, so weaker critical thinking is expected to result in weaker performance. Thus, the key expected impact of AI dependency on academic performance is through critical thinking, the mediation model examined in this study.

Conceptual Framework

Figure 1 presents the conceptual framework. Generative AI dependency is the independent variable, critical thinking is the mediating variable, and academic performance, measured by CGPA, is the dependent variable.

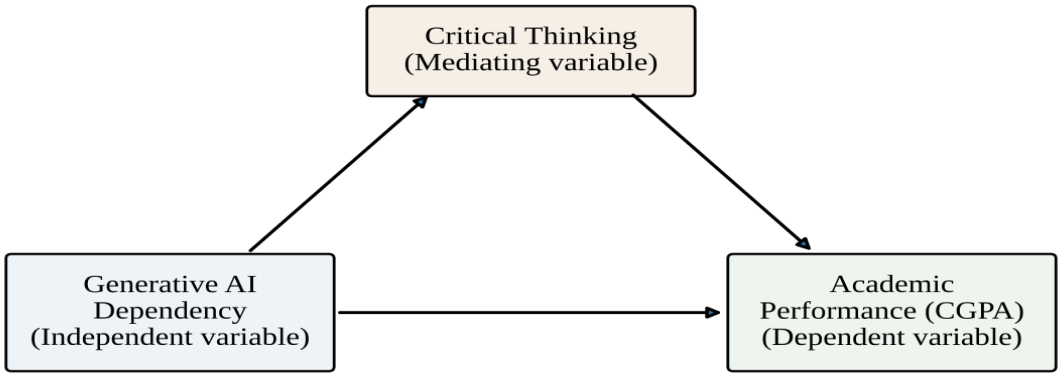


Figure 1. Conceptual framework of the study

The framework shows a direct path from AI dependency to academic performance, and an indirect path that runs from AI dependency to critical thinking and from critical thinking to academic performance. The mediation analysis tests whether the indirect path explains the effect of AI dependency on academic performance.

Research Methodology

Research Design

The study used a quantitative research approach with a survey research design. A quantitative approach was chosen because the study examines numerical relationships between measurable variables, namely generative AI dependency, critical thinking and academic performance. A survey

design was considered appropriate because data were collected from a large sample of university students through a structured questionnaire. Within this design, a correlational approach was adopted, since the study aimed to examine the effect of AI dependency on academic performance and to test whether critical thinking mediates this effect. Variables were measured as they naturally occurred, without manipulation by the researcher. Regression-based mediation analysis was used to examine the direct and indirect effects. In addition, students were compared by level of AI dependency, gender, university sector and daily AI use to identify group differences. Because the design is non-experimental, the findings indicate predictive relationships rather than definite cause-and-effect.

Population of the Study

The population consisted of all undergraduate students enrolled in HEC recognized public and private sector universities located in District Lahore, Punjab, Pakistan. The target population was further limited to undergraduate students who had used a generative AI tool for academic purposes at least once during the current semester, because students who had never used AI could not report dependency on it. Since the population of undergraduate students in Lahore is large (more than 100,000), the sample size was determined with the table of Krejcie and Morgan (1970), which recommends a sample of 384 for populations of this size at a 95% confidence level and a 5% margin of error.

Sample and Sampling Technique

A multistage stratified random sampling technique was used to select the sample, and Figure 2 presents the sampling framework. In the first stage, universities in District Lahore were divided into two strata according to sector: public sector and private sector. In the second stage, four universities were selected at random from each stratum, giving a total of eight universities; in this practice example, they are labelled A–D for the public sector and E–H for the private sector. In the third stage, departments were chosen from each selected university to represent four broad disciplines: sciences, social sciences, arts and humanities, and business. In the fourth and final stage, students were selected through simple random sampling from the class lists of the selected departments, in numbers roughly proportional to the size of each university's undergraduate enrolment. This process produced a final sample of 384 students, of whom 197 were from public universities and 187 were from private universities.

Table 1: Distribution of the sample by university and sector

Sector	University	Sample (n)	Percentage
Public	University A	50	13.0
	University B	49	12.8
	University C	49	12.8
	University D	49	12.8
	Public subtotal	**197**	**51.3**

Private	University E	47	12.2
	University F	47	12.2
	University G	47	12.2
	University H	46	12.0
	Private subtotal	**187**	**48.7**
Total		**384**	**100.0**

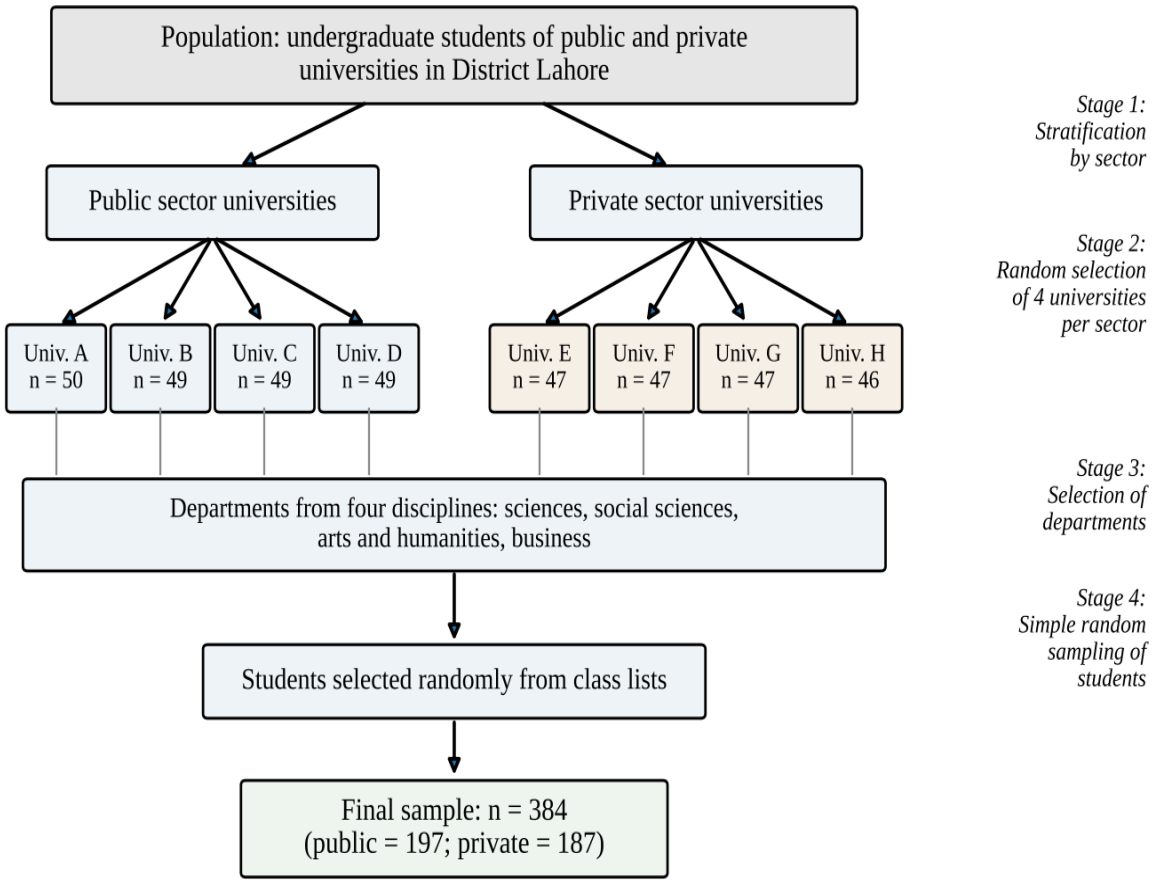


Figure 2. Sampling framework of the study

Instrumentation

Data were collected through a self-administered questionnaire in four parts. All attitude items used a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). The full questionnaire appears in the Appendix.

Part A: Demographic information. This part recorded gender, university sector, discipline, year of study and average daily time spent using generative AI tools for study.

Part B: Generative AI Dependency Scale (GAIDS). AI dependency was measured with a 12-item scale adapted from the Scale of Dependence on Artificial Intelligence developed by Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., and Morales-García, M. (2024), published in Frontiers in Education. The original scale contains five items in a single factor and was validated with 528 university students in Peru, with Cronbach's alpha and McDonald's omega both equal to .87. For the present study, the five original items were retained, slightly reworded for the Pakistani context, and seven items were added to cover behavioural reliance and reduced self-effort. The adapted scale therefore has three sub-dimensions: behavioural reliance (4 items), reduced self-effort (4 items) and psychological attachment (4 items). Higher scores indicate greater dependency.

Part C: Critical Thinking Scale (CTS). Critical thinking was measured with a 14-item scale adapted from the Critical Thinking Disposition Scale (CTDS) developed by Sosu, E. M. (2013), published in Thinking Skills and Creativity. The original CTDS was validated with undergraduate and graduate students in two studies (n = 467 and n = 371) and measures two dimensions, critical openness and reflective scepticism. For the present study, items were adapted around four aspects of critical thinking described by Facione (1990): analysis, evaluation, inference and reflective judgement. Higher scores indicate stronger critical thinking.

Part D: Academic performance. Academic performance was measured through students' latest CGPA on a four-point scale, as reported by the students.

Validity and Reliability of the Instruments

The adapted questionnaire would be reviewed by a panel of three experts in education, educational technology and research methods. The experts would judge each item for relevance, clarity and suitability for Pakistani university students, and items would be revised according to their suggestions. The questionnaire was piloted with 30 undergraduate students from another university other than the main one to ensure that the questions were easy to understand, clear, and acceptable to university students. Such students would not be in the final analysis. The item loadings, composite reliability (CR) and average variance extracted (AVE) were used to check construct validity. Cronbach's alpha was used to determine internal consistency. The results are given in Table 2.

Table 2: Validity and reliability of the instruments (N = 384)

Scale	Items	Loading range	Cronbach's α	CR	AVE	√AVE
Generative AI dependency	12	.698–.777	.920	.932	.532	.729
Critical thinking	14	.643–.733	.916	.928	.479	.692

Note. CR = composite reliability; AVE = average variance extracted. Correlation between the two scales: $r = -.398$.

Both scales exceeded the usual threshold of .70 for Cronbach's alpha and composite reliability (Nunnally, 1978), and all item loadings were above .60. The AVE for the AI dependency scale exceeded .50, which supports convergent validity. The AVE for critical thinking (.479) was slightly below .50, which Fornell and Larcker (1981) consider acceptable when composite reliability exceeds .60, as it did here. Discriminant validity was supported because the square root of each scale's AVE (.729 and .692) was greater than the correlation between the two scales ($r = -.398$). The instruments were therefore considered valid and reliable for the study.

Data Collection Procedure

Permission would be obtained from the heads of the selected departments, and the purpose of the study would be explained to students. Questionnaires would be administered in person during class time, with the researcher present to answer questions. Participation was voluntary and anonymous, and students could withdraw at any time. Completing the questionnaire took about 15 minutes. All 384 returned questionnaires were complete.

Data Analysis

Data were analysed with statistical software. Table 3 links each research question to its analysis.

Table 3: Data analysis plan

Research question	Purpose	Statistical technique
RQ1	Level of AI dependency	Mean, standard deviation, frequency and percentage
RQ2–RQ4	Effects among the three variables	Pearson correlation and linear regression
RQ5	Mediation by critical thinking	Mediation analysis with 5,000 bootstrap samples (Hayes PROCESS Model 4 approach) and the Sobel test
RQ6	Group differences	One-way ANOVA with Tukey's post-hoc test, and independent-samples t-test

Before the main analysis, the assumptions of normality (skewness and kurtosis), homogeneity of variance (Levene's test), independence of errors (Durbin–Watson) and absence of multicollinearity (variance inflation factor) were checked. Mediation was tested with the approach of Hayes (2018). The indirect effect, calculated as $a \times b$, was judged significant when its 95% percentile bootstrap confidence interval did not include zero (Preacher & Hayes, 2008).

Ethical Considerations

Informed consent would be obtained from all participants. The identities of participants would be kept confidential and no names or student numbers were recorded. Data would be securely stored

and only used for research purposes. Prior to data collection, ethical approval would be secured from an appropriate university research ethics committee.4. Data Analysis and Results

Demographic Profile of Respondents

Table 4: Demographic profile of respondents (N = 384)

Characteristic	Category	Frequency	Percentage
Gender	Female	215	56.0
	Male	169	44.0
University sector	Public	197	51.3
	Private	187	48.7
Year of study	1st year	98	25.5
	2nd year	101	26.3
	3rd year	92	24.0
	4th year	93	24.2
Discipline	Sciences	117	30.5
	Social sciences	110	28.6
	Arts and humanities	80	20.8
	Business	77	20.1
Daily AI use	Less than 1 hour	74	19.3
	1–3 hours	150	39.1
	3–5 hours	108	28.1
	More than 5 hours	52	13.5

The sample was fairly balanced by sector and year of study, with slightly more female students. Figure 3 shows the distribution by discipline and daily AI use. The largest group (39.1%) used AI tools for one to three hours a day, and 41.6% used them for more than three hours a day.

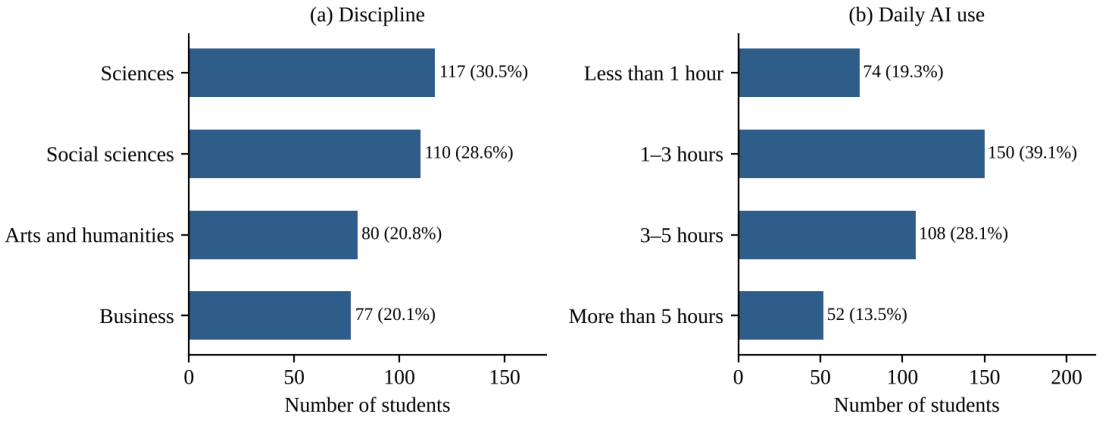


Figure 3. Respondents by (a) discipline and (b) daily hours of generative AI use

Descriptive Statistics and Normality

Table 5: Descriptive statistics of the study variables (N = 384)

Variable	Scale	Mean	SD	Min	Max	Skewness	Kurtosis
AI dependency	1–5	2.99	0.77	1.25	4.92	0.04	−0.53
Critical thinking	1–5	3.00	0.73	1.14	4.71	−0.02	−0.35
CGPA	0–4	3.02	0.42	1.87	4.00	−0.03	−0.33

Mean AI dependency (M = 2.99) and critical thinking (M = 3.00) lay at the midpoint of the scale, and the mean CGPA was 3.02. All skewness and kurtosis values were within ±1, which indicates approximately normal distributions (Figure 4).

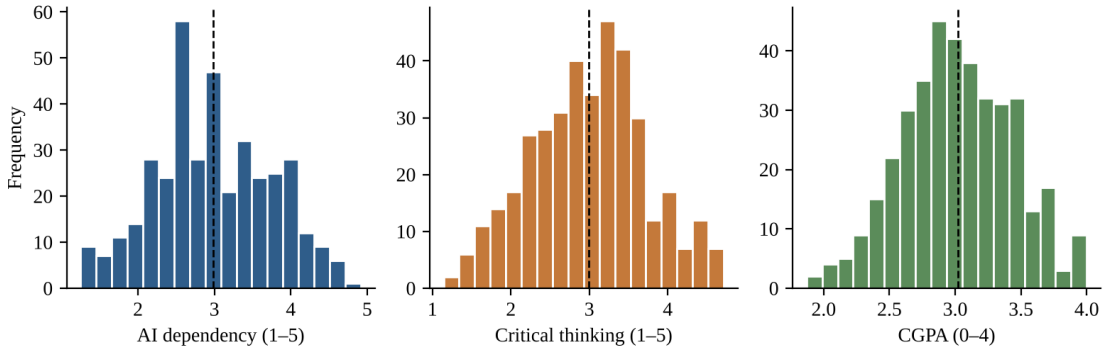


Figure 4. Distribution of AI dependency, critical thinking and CGPA (dashed lines show means)

Level of Generative AI Dependency (RQ1)

Table 6: Levels of generative AI dependency

Level	Score range	Frequency	Percentage
Low	1.00–2.33	69	18.0
Moderate	2.34–3.67	234	60.9
High	3.68–5.00	81	21.1

Most students (60.9%) showed moderate dependency, about one in five (21.1%) showed high dependency, and only 18.0% showed low dependency. Reliance on AI was therefore common, and a substantial minority relied on it heavily.

Correlation Analysis

Table 7: Pearson correlations among the study variables (N = 384)

Variable	1	2	3
1. AI dependency	1		
2. Critical thinking	−.398***	1	
3. CGPA	−.253***	.516***	1

Note. ***p < .001.

AI dependency had a moderate negative correlation with critical thinking ($r = -.398$) and a weak negative correlation with CGPA ($r = -.253$). Critical thinking had a strong positive correlation with CGPA ($r = .516$). All correlations were significant at $p < .001$ (Figure 5).

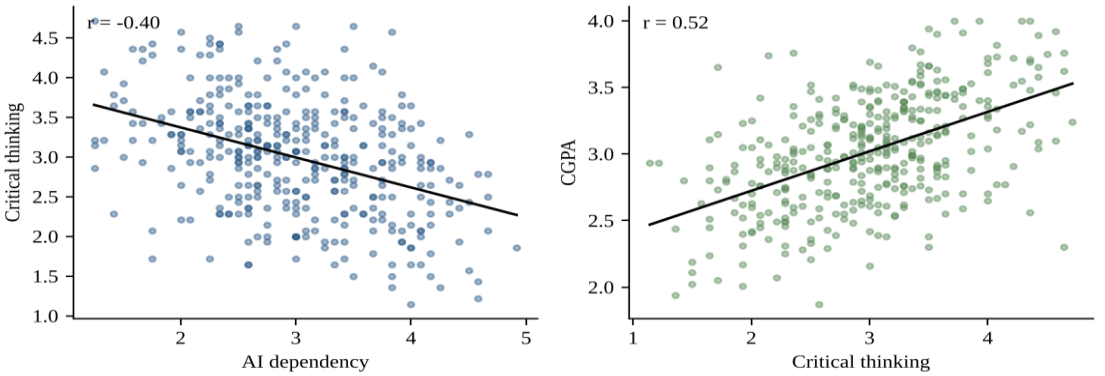


Figure 5. Scatterplots of AI dependency with critical thinking (left) and critical thinking with CGPA (right)

Effect of AI Dependency and Critical Thinking (RQ2–RQ4)

The variance inflation factor was 1.19 and the Durbin–Watson statistic was 1.88, so the assumptions for regression were met. Table 8 shows the regression models.

Table 8: Regression analyses of the direct effects (N = 384)

Model	Path	Predictor → Outcome	B	SE	β	t	p	R^2
1	c	AI dependency → CGPA	−0.138	0.027	−.253	−5.11	< .001	.064
2	a	AI dependency → Critical thinking	−0.378	0.045	−.398	−8.48	< .001	.158
3	b	Critical thinking → CGPA	0.283	0.027	.493	10.33	< .001	.269
	c′	AI dependency → CGPA	−0.031	0.026	−.057	−1.19	.236	

Note. Model 1: $F(1, 382) = 26.11$; Model 2: $F(1, 382) = 71.83$; Model 3: $F(2, 381) = 70.01$; all $p < .001$.

AI dependency significantly and negatively predicted CGPA ($B = -0.138$, $p < .001$), explaining 6.4% of its variance (RQ2). AI dependency also significantly and negatively predicted critical thinking ($B = -0.378$, $p < .001$), explaining 15.8% of its variance (RQ3). Controlling for AI dependency, critical thinking significantly and positively predicted CGPA ($B = 0.283$, $\beta = .493$, $p < .001$), and the two variables together explained 26.9% of the variance in CGPA (RQ4).

Mediating Role of Critical Thinking (RQ5)

Table 9: Total, direct and indirect effects of AI dependency on CGPA

Effect	Path	B	SE	95% CI	p	Result
Total effect	c	−0.138	0.027	[−0.191, −0.085]	< .001	Significant
Direct effect	c′	−0.031	0.026	[−0.082, 0.020]	.236	Not significant
Indirect effect	a × b	−0.107	0.018	[−0.145, −0.075]	—	Significant

Note. The CI for the indirect effect is a percentile bootstrap interval based on 5,000 resamples. Sobel test: $z = -6.55$, $p < .001$.

The indirect effect of AI dependency on CGPA through critical thinking was −0.107 (a × b = −0.378 × 0.283). Its 95% bootstrap confidence interval [−0.145, −0.075] did not include zero, so the indirect effect was significant. When critical thinking was included in the model, the direct effect fell from −0.138 to −0.031 and became non-significant ($p = .236$). This pattern shows full mediation, with the indirect effect accounting for 77.6% of the total effect. Figures 6 to 8 illustrate these results.

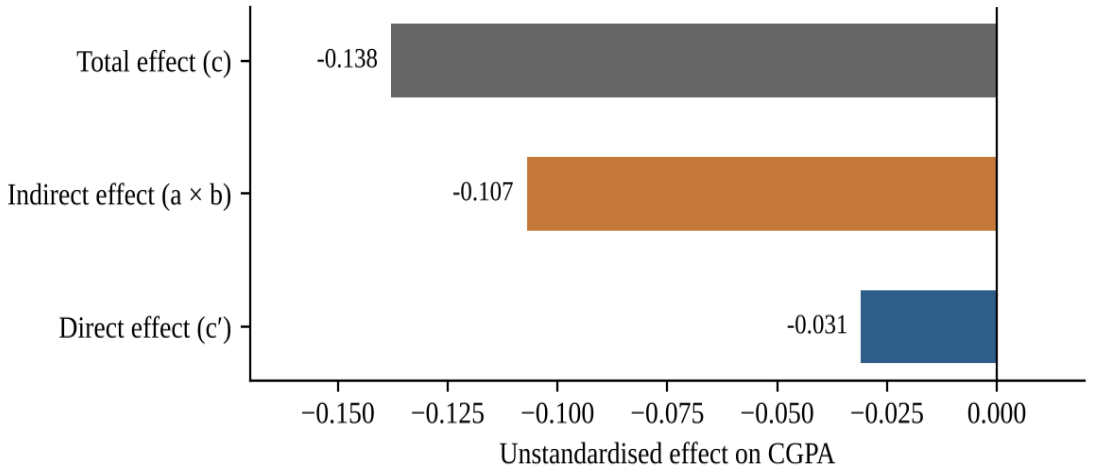


Figure 6. Decomposition of the total effect into indirect and direct effects

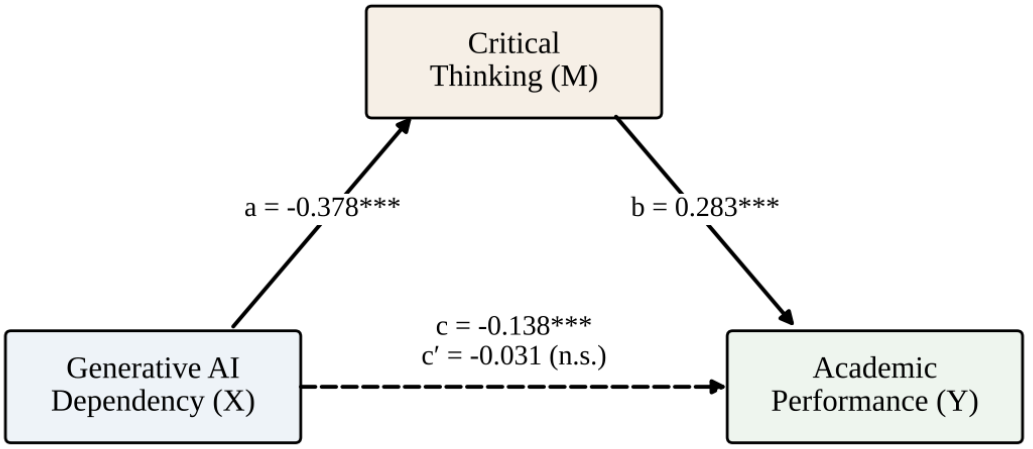


Figure 7. Mediation model with unstandardised path coefficients (***p < .001; dashed line = non-significant direct effec

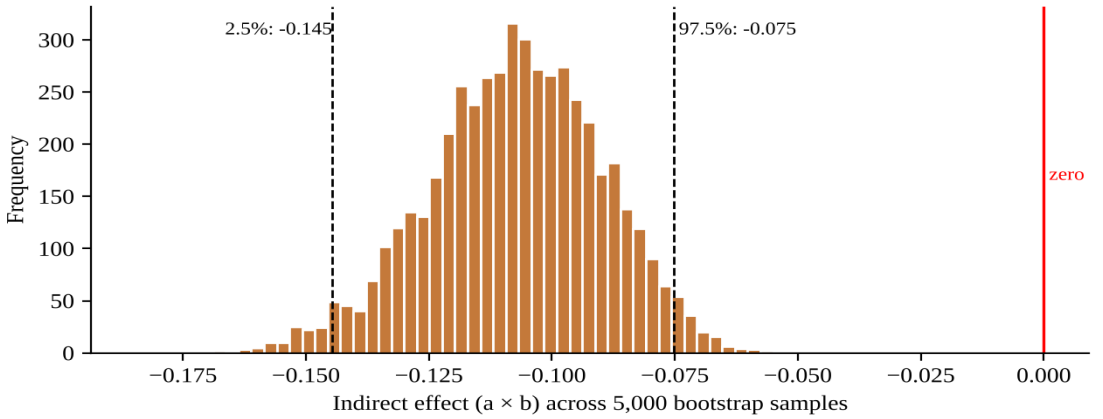


Figure 8. Bootstrap distribution of the indirect effect (5,000 resamples)

Comparison by Level of AI Dependency (RQ6)

Levene's test showed equal variances across the three dependency groups for critical thinking (p = .515) and CGPA (p = .661), so one-way ANOVA was used.

Table 10: Critical thinking and CGPA by level of AI dependency

Variable	Low (n = 69) M (SD)	Moderate (n = 234) M (SD)	High (n = 81) M (SD)	F(2, 381)	p	η^2
Critical thinking	3.40 (0.67)	3.03 (0.67)	2.57 (0.71)	28.41	< .001	.130
CGPA	3.19 (0.39)	3.02 (0.41)	2.87 (0.40)	11.60	< .001	.057

Table 11: Tukey post-hoc comparisons by level of AI dependency

Variable	Comparison	Mean difference	p	95% CI
Critical thinking	Low – Moderate	0.37	< .001	[0.15, 0.59]
	Low – High	0.83	< .001	[0.57, 1.09]
	Moderate – High	0.46	< .001	[0.25, 0.67]
CGPA	Low – Moderate	0.17	.009	[0.03, 0.30]
	Low – High	0.32	< .001	[0.16, 0.48]
	Moderate – High	0.16	.009	[0.03, 0.28]

Critical thinking differed significantly across the three groups, $F(2, 381) = 28.41$, $p < .001$, $\eta^2 = .130$, a large effect. CGPA also differed significantly, $F(2, 381) = 11.60$, $p < .001$, $\eta^2 = .057$, a medium effect. Tukey's test showed that all three groups differed from one another on both variables. As dependency increased from low to high, mean critical thinking fell from 3.40 to 2.57 and mean CGPA fell from 3.19 to 2.87 (Figure 9).

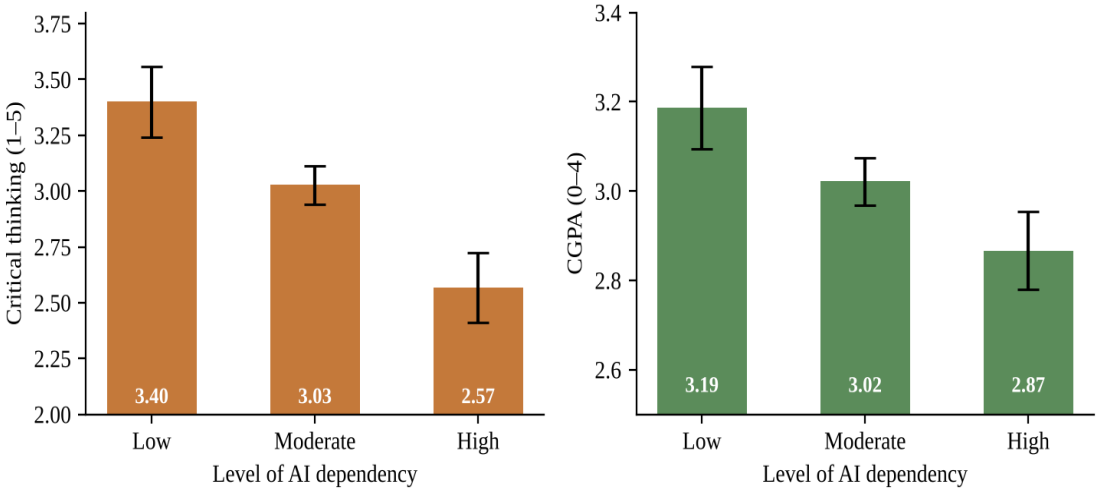


Figure 9. Mean critical thinking and CGPA by level of AI dependency (error bars show 95% confidence intervals)

Comparison by Gender and University Sector (RQ6)

Table 12: Independent-samples t-tests by gender and university sector

Variable	Male M (SD)	Female M (SD)	t (382)	p	Public M (SD)	Private M (SD)	t (382)	p
AI dependency	2.96 (0.76)	3.01 (0.78)	−0.62	.534	2.92 (0.78)	3.06 (0.75)	−1.86	.064
Critical thinking	3.02 (0.72)	2.98 (0.74)	0.61	.540	3.05 (0.71)	2.94 (0.74)	1.47	.142
CGPA	3.04 (0.42)	3.00 (0.42)	0.87	.385	3.03 (0.42)	3.01 (0.42)	0.63	.530

No significant differences were found between male and female students, or between public and private university students, on any variable (all $p > .05$). Private university students reported slightly higher AI dependency, but the difference did not reach significance ($p = .064$).

Comparison by Daily AI Use (RQ6)

Table 13: AI dependency, critical thinking and CGPA by daily AI use

Daily AI use	n	AI dependency M (SD)	Critical thinking M	CGPA M
Less than 1 hour	74	2.54 (0.64)	3.26	3.09
1–3 hours	150	2.90 (0.73)	2.95	3.01
3–5 hours	108	3.19 (0.69)	2.95	3.00
More than 5 hours	52	3.46 (0.81)	2.87	2.99

Note. ANOVA for AI dependency: $F(3, 380) = 20.55$, $p < .001$, $\eta^2 = .140$.

AI dependency differed significantly by daily use, $F(3, 380) = 20.55$, $p < .001$, $\eta^2 = .140$. Tukey's test showed that all pairs of groups differed except the 3–5 hours and more than 5 hours groups ($p = .113$). Dependency rose steadily from 2.54 among students using AI for less than an hour a day to 3.46 among those using it for more than five hours, while critical thinking fell from 3.26 to 2.87 (Figure 10).

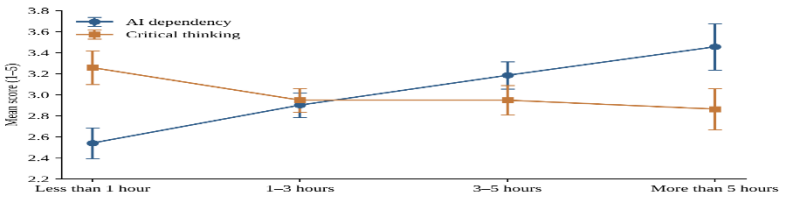


Figure 10. Mean AI dependency and critical thinking by daily AI use (error bars show 95% confidence intervals)

Findings

- Most students showed moderate generative AI dependency (60.9%), 21.1% showed high dependency and only 18.0% showed low dependency. The overall mean was 2.99 on a five-point scale.
- Daily AI use was high: 41.6% of students used AI tools for more than three hours a day.
- AI dependency had a significant negative effect on academic performance ($B = -0.138$, $\beta = -.253$, $p < .001$), explaining 6.4% of the variance in CGPA.
- AI dependency had a significant negative effect on critical thinking ($B = -0.378$, $\beta = -.398$, $p < .001$), explaining 15.8% of its variance.
- Critical thinking had a significant positive effect on academic performance ($B = 0.283$, $\beta = .493$, $p < .001$), and was the strongest predictor of CGPA.
- Critical thinking fully mediated the effect of AI dependency on academic performance. The indirect effect was −0.107 (95% CI [−0.145, −0.075]), and the direct effect became non-significant ($p = .236$). The indirect effect accounted for 77.6% of the total effect.
- Students with high AI dependency had significantly lower critical thinking ($M = 2.57$) than students with moderate ($M = 3.03$) and low dependency ($M = 3.40$), $F(2, 381) = 28.41$, $p < .001$.
- Students with high AI dependency also had significantly lower CGPA ($M = 2.87$) than students with moderate ($M = 3.02$) and low dependency ($M = 3.19$), $F(2, 381) = 11.60$, $p < .001$.
- AI dependency increased significantly with daily hours of AI use, $F(3, 380) = 20.55$, $p < .001$, $\eta^2 = .140$, while critical thinking decreased as use increased.
- No significant differences in AI dependency, critical thinking or CGPA were found between male and female students, or between public and private university students.

Discussion

This study examined the effect of generative AI dependency on the academic performance of university students in District Lahore, and the mediating role of critical thinking. Because the data in this practice article are simulated, the discussion shows how such findings would be interpreted in a real study.

The first finding was that most students showed a moderate level of AI dependency, and that about one in five showed a high level. This is in line with other Pakistani studies that found high levels of AI adoption. Zafar et al., (2024) reported that approximately 50% of the students used ChatGPT regularly, while Shahzad and Farooq (2025) reported that 82% of all the undergraduate students from Punjab used at least one AI tool. The present study provides an additional example of dependency to use in the minds of many students. Note that there is a particularly high correlation between hours of usage per day and dependency ($\eta^2 = .140$). Nearly one point higher on dependency was recorded by the students who used AI more than five hours a day compared to those who used it for less than an hour. This pattern is consistent with the cognitive offloading theory (Risko & Gilbert,

2016) which explains that our brains become more likely to turn to a tool when the tool is available and convenient, until the behaviour becomes a habit.

The researchers conclude that relying on AI was strongly negatively correlated with academic achievement. The findings are consistent with those of Abbas et al. (2024) who found that the usage of ChatGPT among the university students in Pakistan was negatively related to academic performance. It seems to contradict Shahzad and Farooq (2025), who suggested that AI usage positively impacted self-reported academic performance, and Qureshi et al. (2024), whose respondents felt that ChatGPT enhanced their learning. The difference is attributed to the different measurements taken. The studies cited in the previous section were on general usage and perceived benefits while the present study was on dependency. When used for tasks that students can do independently, it can become a dependency and potentially affect their ability to manage time and perform tasks. The adoption of AI coursework writing tools in Pakistani universities, where exams are frequently supervised, could pose difficulties for students who rely on the technology for their academic work when they are required to write exams without it. This may also vary from student to student as students may feel productive when they are not learning much, or vice versa.

The second significant negative effect of dependency in the model was on critical thinking, which was the most powerful effect of dependency in the model. This finding aligns with that of Gerlich (2025), who reported decreased critical thinking among frequent AI users, and with the systematic review by Zhai et al. (2024) that found that excessive dependence on AI dialogue systems can reduce analytical thinking. It also agrees with Lee et al. (2025), who reported that individuals who had more trust in AI were less critical in their work. Cognitive load theory (Sweller, 1988) provides a good explanation. Critical thinking is an emergent product of germane cognitive effort, which includes weighing evidence, comparison of arguments, drawing conclusions, etc. As AI does these steps, students are spared the work and thus the learning that is supposed to come in the process is not gained. This effect can be particularly worrisome in the Pakistani context. Approximately 75% of prospective teachers in Lahore were unable to score more than 50% in a critical thinking test (Bibi & Akhter, 2020) and exams are frequently prioritised in schools over thinking (Jamil & Muhammad, 2019). Students arriving at university with less experience working independently may be drawn to AI, and may fail to gain experience in independent thinking at university.

These were reinforced by the group comparisons. Students who were highly dependent, scored significantly lower than students with moderate and low dependency on critical thinking and CGPA and there was a difference between the groups and students with moderate dependency and those with low dependency. The trend of the pattern was that dependency increased from low to moderate to high and critical thinking and CGPA decreased in step. This dose-like pattern would be predicted if dependency influences thinking and performance, but, as a causal-comparative design, this study does not rule out any other interpretations. In some cases, the relationship between student's critical thinking and using AI can be bi-directional: If a student's critical thinking is weaker, then they may use AI more often because they feel tasks are tougher for them.

The best predictor of CGPA was critical thinking. This is consistent with Halpern (1998), who stated that critical thinking is one of the major objectives of higher education as well as with the findings reported by Manzoor et al. (2019) who reported a positive correlation of critical thinking disposition with problem solving skills and self-efficacy among prospective teachers in Lahore. The ability to interpret, analyze and argue is usually the type of thinking that is encouraged and valued in the university, and therefore students who have a good critical thinking are likely to fare well in all their subjects.

The most significant result was related to mediation. Pathway analysis showed that the association between AI dependency and academic performance was fully mediated. After accounting for critical thinking, AI dependency had little direct impact on CGPA, and the indirect impact was responsible for 77.6% of the total impact. The finding indicates that relying on AI is not a detrimental effect on grades, per se. Dependency does harm to students' thinking. This explanatory model is consistent with the overall theory of the study that integrates. This is because cognitive offloading is the reason students resort to AI, cognitive load theory is the reason that students are less dependent on it and thus less critical thinking is involved, and the latter is the reason for their lower performance. There is a positive take from the Socio-cultural theory. Vygotsky (1978) has referred to learning as a transition from assisted activity to self-directed activity. As students learn to do this themselves, AI can serve as an aid, as demonstrated by Nasr et al. (2025) with students working collaboratively. What is bad about dependency is that no support is ever withdrawn. The mediation outcome thus suggests that a ban on AI is not the answer; it is about supporting instead of replacing students' reasoning.

The mediation outcome is also in line with past research. In a study, Gerlich (2025) demonstrated that the use of AI correlates with a decrease in critical thinking, because of cognitive offloading. This current study contributes the next piece of the puzzle – demonstrating the relationship between lower levels of critical thinking and academic performance and AI dependency. A majority of the existing research conducted in Pakistan has focused on the direct relationship between AI usage and outcomes (Qureshi et al., 2024; Shahzad & Farooq, 2025). This study indicates that, in future, critical thinking should be included in the model of AI and learning in Pakistan, after testing a mechanism.

Gender and university sector didn't result in any significant differences. The gender result is consistent with that of Qureshi et al. (2024) who found no significant difference between gender with regards to students' perceptions of ChatGPT, and Alghazo et al. (2025) who found no significant difference between the genders with regards to the perceived benefits. It is surprising that the students in private universities reported higher benefits from the use of ChatGPT compared to those in government universities, which was revealed by Alghazo et al. (2025). The results of the current study indicated that there was a non-significant difference ($p = .064$) between the dependency of private university students. This indicates that reliance on AI is not confined to any sector or gender in Lahore's universities; it is a widespread phenomenon. Any answer must therefore be applicable to the institution.

There are a number of restrictions to take into consideration. Most importantly this article uses fake data as a learning tool so results of this article cannot be generalized with regard to real students. More restrictions would pertain to a real study of this design. Causal-comparative and cross-sectional design cannot claim to prove causality with assurance because there may be differences between the groups that are measured which are not captured in the study. All items were self-reported, which could lead to social desirability bias and common method variance in the reporting of dependency, critical thinking and CGPA. The adapted scales would require confirmatory factor analysis to validate the scales. As the study was conducted with under graduate students at district Lahore, the results may vary in other Cities, rural areas and post-graduate programmes. Longitudinal or experimental study design and the use of performance-based critical thinking tests like the Watson–Glaser appraisal, official academic records, and other variables like self-efficacy, motivation, and AI literacy could be the next steps for future research.

Conclusion

This study looked at whether depending on generative AI tools, such as ChatGPT, affects how well university students in Lahore do in their studies. It also asked whether critical thinking, meaning students' ability to analyse, question and reason for themselves, explains this effect. The results showed that AI dependency is common. Most students depended on AI to a moderate degree, and about one in five depended on it heavily. Students who spent more hours each day using AI were more dependent on it. Students who depended more on AI tended to have weaker critical thinking and lower grades. When students were divided into low, moderate and high dependency groups, critical thinking and CGPA fell step by step from the low group to the high group. Students with stronger critical thinking, on the other hand, achieved higher grades. The most important result was that critical thinking explained almost the whole link between AI dependency and grades. In simple words, depending on AI did not lower grades directly. It lowered grades because it was linked to weaker thinking skills. This suggests that the real danger is not AI itself, but letting AI do the thinking that students should do themselves. Male and female students, and students in public and private universities, did not differ significantly. This means AI dependency is a shared challenge for all universities in Lahore, not a problem for one group only.The main message is clear and practical. Universities do not need to ban AI. Instead, they should help students use AI in ways that keep their own minds active. When students think first and use AI only to check, support or extend their work, they can benefit from AI without losing the critical thinking that their studies, and their future careers, require.

Recommendations

Based on the findings and conclusions of the study, the following recommendations are made:

1. The study found that most university students in District Lahore had a moderate to high level of generative AI dependency. Universities should therefore run AI literacy workshops at the start of each degree programme, teaching students when to use AI, how to check its output and how to acknowledge its use.

2. Generative AI dependency had a negative effect on students' academic performance. Teachers should therefore include assessments that students must complete independently, such as in-class writing, oral presentations, viva examinations and staged drafts, so that grades reflect students' own learning.
3. Generative AI dependency had a negative effect on students' critical thinking. Teachers should design tasks in which students' critique, correct or improve AI-generated answers, so that AI is used to practice critical thinking rather than replace it.
4. Critical thinking had a strong positive effect on academic performance. Universities and the Higher Education Commission should include explicit critical thinking instruction across undergraduate curricula, rather than assuming that it will develop on its own.
5. Critical thinking fully mediated the relationship between generative AI dependency and academic performance. University AI policies should therefore focus on protecting students' independent reasoning, for example by encouraging a "think first, then check with AI" approach, rather than banning AI.
6. Students with high AI dependency had significantly lower critical thinking and CGPA than students with low and moderate dependency. Universities should identify these students early and provide academic support, mentoring and study-skills training.
7. AI dependency increased with daily hours of AI use. Student advisers should encourage students to monitor the time they spend using AI and should offer guidance to students who report very heavy daily use.
8. No significant differences were found by gender or university sector. AI guidelines and training should therefore be applied equally to all students in both public and private universities. Teacher education institutions should prepare prospective teachers to model responsible AI use and to teach critical thinking explicitly.

References

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. International Journal of Educational Technology in Higher Education, 21(1), Article 10. https://doi.org/10.1186/s41239-024-00444-7

Alghazo, R., Fatima, G., Malik, M., Abdelhamid, S. E., Jahanzaib, M., Nayab, D. e., & Raza, A. (2025). Exploring ChatGPT's role in higher education: Perspectives from Pakistani university students on academic integrity and ethical challenges. Education Sciences, 15(2), Article 158. https://doi.org/10.3390/educsci15020158

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bibi, S., & Akhter, M. (2020). A study to assess critical thinking skills among prospective teachers. Pakistan Social Sciences Review, 4(4), 134–143. https://doi.org/10.35484/pssr.2020(4-IV)09

Dogaru, M., Pisică, O., Popa, C.-Ş., Răgman, A.-A., & Tololoi, I.-R. (2025). The perceived impact of artificial intelligence on academic learning. Frontiers in Artificial Intelligence, 8, Article 1611183. https://doi.org/10.3389/frai.2025.1611183

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report). The California Academic Press.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), Article 6. https://doi.org/10.3390/soc15010006

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. American Psychologist, 53(4), 449–455. https://doi.org/10.1037/0003-066X.53.4.449

Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (2nd ed.). Guilford Press.

Jamil, M., & Muhammad, Y. (2019). Teaching science students to think critically: Understanding secondary school teachers' practices. Journal of Research and Reflections in Education, 13(2), 256–272.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, 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, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30(3), 607–610. https://doi.org/10.1177/001316447003000308

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. https://doi.org/10.1145/3706598.3713778

Manzoor, E., Hussain, T., & Hashmi, A. (2019). Prospective teachers' critical thinking disposition, problem solving skills and self-efficacy: A relationship study in Pakistan. Pakistan Social Sciences Review, 3(2), 687–696. https://doi.org/10.35484/pssr.2019(3-2)53

Mehran, M., & Farooq, A. (2025). Role of ChatGPT in undergraduate education: Benefits, challenges, and ethical considerations. Journal of Development and Social Sciences, 6(2), 1–12. https://doi.org/10.47205/jdss.2025(6-II)01

Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., & Morales-García, M. (2024). Development and validation of a scale for dependence on artificial intelligence in university

students. Frontiers in Education, 9, Article 1323898. https://doi.org/10.3389/feduc.2024.1323898

Mujtaba, S., & Ali, M. (2025). Moderating role of ChatGPT usage between self-efficacy and student engagement among students in Pakistan: Toward academic technoutopia. Applied Psychology Review, 4(2), 1–20. https://doi.org/10.32350/apr.42.01

Nasr, N. R., Tu, C.-H., Werner, J., Bauer, T., Yen, C.-J., & Sujo-Montes, L. (2025). Exploring the impact of generative AI ChatGPT on critical thinking in higher education: Passive AI-directed use or human–AI supported collaboration? Education Sciences, 15(9), Article 1198. https://doi.org/10.3390/educsci15091198

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). McGraw-Hill.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Qureshi, R., Noreen, S., & Ismaeel, G. (2024). The role of AI and ChatGPT in enhancing learning achievement among university students: A quantitative analysis. iRASD Journal of Educational Research, 5(1), 75–81. https://doi.org/10.52131/jer.2024.v5i1.2551

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

Shah, H. A., Khoso, I. A., & Nisar, S. (2024). The role of ChatGPT in shifting cognitive approaches to education in Sindh, Pakistan: Explaining a theoretical TAM framework. Journal of Development and Social Sciences, 5(4), 594–600. https://doi.org/10.47205/jdss.2024(5-IV)51

Shahzad, M., & Farooq, M. S. (2025). Ameliorating university students' academic productivity through artificial intelligence tools. Sir Syed Journal of Education & Social Research, 8(2), 28–35. https://doi.org/10.36902/sjesr-vol8-iss2-2025(28-35)

Sosu, E. M. (2013). The development and psychometric validation of a Critical Thinking Disposition Scale. Thinking Skills and Creativity, 9, 107–119. https://doi.org/10.1016/j.tsc.2012.09.002

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Zafar, S., Shaheen, F., & Rehan, J. (2024). Use of ChatGPT and generative AI in higher education: Opportunities, obstacles and impact on student performance. iRASD Journal of Educational Research, 5(1), 1–12. https://doi.org/10.52131/jer.2024.v5i1.2463

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11, Article 28. https://doi.org/10.1186/s40561-024-00316-7

Appendix: Research Questionnaire

Dear student, this questionnaire is part of a study on the use of generative AI tools (such as ChatGPT, Gemini, Copilot, Grammarly and QuillBot) and academic learning. Your answers are anonymous and will be used only for research. Please answer honestly; there are no right or wrong answers.

Part A: Demographic Information

Item	Response (tick one)
1. Gender	☐ Male ☐ Female
2. University sector	☐ Public ☐ Private
3. Discipline	☐ Sciences ☐ Social sciences ☐ Arts and humanities ☐ Business
4. Year of study	☐ 1st ☐ 2nd ☐ 3rd ☐ 4th
5. Average daily time spent using generative AI tools for study	☐ Less than 1 hour ☐ 1–3 hours ☐ 3–5 hours ☐ More than 5 hours
6. Latest CGPA (out of 4.00)	_____

Part B: Generative AI Dependency Scale

Adapted from Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., & Morales-García, M. (2024), Frontiers in Education, 9, Article 1323898. Items marked * are from the original five-item scale (slightly reworded); the other items were added for this practice study.

Response scale: SD = strongly disagree (1), D = disagree (2), N = neutral (3), A = agree (4), SA = strongly agree (5). Please circle one number for each statement.

No.	Statement	SD	D	N	A	SA
	Behavioural reliance					
1	I use AI tools for almost every assignment I complete.	1	2	3	4	5
2	My first step when I receive a task is to ask an AI tool.	1	2	3	4	5
3	I use AI to answer questions I could find in my notes or textbooks.	1	2	3	4	5
4	I use AI tools more often now than when I first started using them.	1	2	3	4	5
	Reduced self-effort					
5	I often submit AI-generated content with only minor changes.	1	2	3	4	5
6	I rarely check whether the information given by AI is correct.	1	2	3	4	5
7	I ask AI to explain readings instead of reading them myself.	1	2	3	4	5
8	I need validation or feedback from AI to feel confident in my academic decisions.	1	2	3	4	5
	Psychological attachment					
9	I feel unprotected when I do not have access to AI.	1	2	3	4	5

10	I am concerned about falling behind in my tasks or projects if I do not use AI.	1	2	3	4	5
11	I do everything possible to stay updated with AI to remain relevant in my field.	1	2	3	4	5
12	I fear that AI might replace my current skills or abilities.	1	2	3	4	5

Part C: Critical Thinking Scale

Adapted from the Critical Thinking Disposition Scale of Sosu, E. M. (2013), Thinking Skills and Creativity, 9, 107–119, and organised around the dimensions of critical thinking described by Facione (1990). The items below were written for this practice study; a real study should use the original items with permission.

No.	Statement	SD	D	N	A	SA
	Analysis					
1	I break complex problems into smaller parts before solving them.	1	2	3	4	5
2	I identify the main argument when I read an academic text.	1	2	3	4	5
3	I can recognise the assumptions behind an argument.	1	2	3	4	5
4	I compare different viewpoints before forming an opinion.	1	2	3	4	5
	Evaluation					
5	I check the credibility of the sources I use.	1	2	3	4	5
6	I can tell when an argument is supported by weak evidence.	1	2	3	4	5
7	I question information even when it comes from an authority.	1	2	3	4	5
8	I judge the strengths and weaknesses of my own work.	1	2	3	4	5
	Inference					
9	I draw logical conclusions from the evidence available.	1	2	3	4	5
10	I can predict the consequences of different solutions.	1	2	3	4	5
11	I support my conclusions with reasons.	1	2	3	4	5
	Reflective judgement					
12	I am willing to change my view when I meet strong evidence.	1	2	3	4	5
13	I think about how I reached a conclusion.	1	2	3	4	5
14	I look for alternative explanations before deciding.	1	2	3	4	5

Thank you for your participation.