INTERNATIONAL JOURNAL OF SOCIAL SCIENCES BULLETIN
ISSN: (E) 3007-1917 (P) 3007-1909
Volume 4, Issue 6, 2026

ARTIFICIAL INTELLIGENCE LITERACY AND ITS ROLE IN DEVELOPING CRITICAL THINKING SKILLS AMONG 4TH SEMESTER STUDENTS OF UNIVERSITY OF GUJRAT

Dr. Nazish Andleeb[*1], Dr. Nishat Zafer[2], Dr. Azmat Farooq[3]

[*1]*Assistant Professor, Department of Education, University of Gujrat, Pakistan*
[2]*Associate Lecturer, Department of Education, University of Gujrat, Pakistan*
[3]*Assistant Professor, Khwaja Fareed University of Engineering and Information Technology (KFUEIT), Pakistan*

[*1]nazish.andleeb@uog.edu.pk

Corresponding Author: *
Dr. Nazish Andleeb
DOI: https://doi.org/10.5281/zenodo.20843528

Received	Accepted	Published
25 April 2026	04 June 2026	21 June 2026

ABSTRACT

Artificial Intelligence (AI) has become a permanent part of academic and professional life and has brought along a whole new set of skills AI literacy that educators are just starting to seriously consider. This paper is a study of how AI literacy affects critical thinking skills among 4th semester students of University of Gujrat (UoG), Pakistan. With a mixed-methods sequential explanatory design to collect and analyse data first a quantitative phase was undertaken. Quantitative data were collected from 150 students using a validated AI Literacy Scale and a critical thinking instrument based on the Watson-Glaser Critical Thinking Appraisal and the California Critical Thinking Skills Test. Later, semi-structured interviews were conducted with 15 purposively selected participants. ... AI literacy and critical thinking scores Quite a bit and positively correlated, as the Pearson correlation test ($r = 0.54$, $p<0.001$). With regression analysis, it was established that AI literacy was a significant predictor of critical thinking and could explain 38% of the variance. The qualitative themes revealed the three ways of mechanisms: questioning AI outputs, recognizing AI limitations, and developing intellectual autonomy. According to the study, AI literacy is considered a large intellectual capability rather than merely a technical skill, and it also has a significant impact on the students' reasoning, evaluation, and engagement with knowledge. The paper calls for increased initiatives in curriculum revision, teacher training, and strengthening college policy.

1. INTRODUCTION

Something has silently changed in university lectures and it is even more profound than one single new technology. Students who a couple of years ago would write their first draft by hand and check the dictionary when revising, are now dealing with a reality where an AI program can generate a whole paragraph in less than a second, provide a literature review without having read a single document, or answer a test-type question with the fluency of a student who has really prepared. The main concern of teachers should not be if this is really happening because it obviously is but if students are equipped with the right skills to deal with it ethically.

Being literate when using Artificial Intelligence, or AI literacy, is the shorthand for the set of knowledge and skills that give a person the capability to grasp the operating principles of AI systems, identify their strengths and limitations capabilities, and critically, rather than passively, interact with their results (Long and Magerko, 2020). It is relevant to stress what it is not: AI literacy is not the same as programming nor is it

equivalent to computer skills in general. For example, a student proficient in digital skills - that is, knowledgeable and comfortable with smartphones, social media, and online research - may, at the same time, be completely AI-illiterate in the sense relevant to this discussion: he or she does not sufficiently understand how to assess if the output of an AI is accurate biased incomplete or misleading within its context.

Critical thinking is actually one of the most long-lasting aims of higher education. Educators and philosophers, from Dewey (1910) to Facione (1990) and Paul and Elder (2001), have constantly pointed out that the ability to engage in reflective, evidence-based reasoning characterizes the educated mind. The Higher Education Commission (HEC) of Pakistan has officially included critical thinking as a necessary graduate attribute, and universities in the country, at least in theory, are devoted to facilitating it. But research continues to show that a large number of students in Pakistani universities find difficulty in performing higher-order cognitive tasks that critical thinking involves, and also that traditional teaching methods, which typically emphasize memorization and examination performance, have very little success in solving this problem.

This research explores the AI literacy level of the students, the relation of their AI literacy to critical thinking, and the influence of various factors on that relation. It also adds to the literature on AI literacy in higher education internationally and offers findings directly relevant to curriculum and policy making at UoG and Pakistani higher education generally.

The study starts with outlining the research objectives and questions, making a case for the importance of the topic, and describing the problem. Next, the literature review discusses AI literacy frameworks, critical thinking in higher education, and the emerging empirical evidence on the overlap between the two. The methods section covers research design participants instruments, and methods of analysis. The results are shown and discussed in relation to the previous literature. Finally, main points and suggestions are presented.

2. Research Objectives

This research had these three objectives in total;

Objective 1: To assess the degree of AI literacy among 4th semester undergraduate students of University of Gujrat from various fields of study and gender.

Objective 2: To examine the extent of the relationship between AI literacy and critical thinking skills in the student body.

Objective 3: To identify individual and contextual factors (for example prior experiences with AI tools discipline gender and self-efficacy) that influence or explain the association between AI literacy and critical thinking development.

3. Research Questions

The study is guided by the following three research questions:

RQ1: What is the current level of AI literacy among 4th semester students at the University of Gujrat, and how does it vary across academic disciplines and gender groups?

RQ2: Is there a statistically significant relationship between AI literacy levels and critical thinking competency among this student population?

RQ3: Which individual and contextual factors moderate or mediate the relationship between AI literacy and critical thinking among 4th semester students at UoG?

4. Significance of the Study

The significance of this study is not only at one level. but, it is the institutional level, the national level, and the international level. So, it is only right to be clear about each of these. Firstly, at the institutional level, the University of Gujrat is the one that is currently the most served by students using AI for their academic work. But, at the same time, the university curriculum today does not comprise any formal provision for AI literacy education. This research, Because of this, is setting up the first empirically data of AI literacy levels among UoG undergraduates, and such data is a must for any curricular decision-making to be based on information.

Secondly, at the national level, the research is a contribution to a topic that Pakistan's higher education is in dire need of discussing. In fact, the

HEC has promised commitment to digital transformation and technology integration in universities. But again, technology integration without critical literacy education might lead to students who are more and more dependent on AI, rather than less - they are capable of using the tools efficiently but cannot assess them meaningfully. In other words, this research demonstrates that AI literacy is linked with critical thinking and that means that spending on AI literacy education is also spending on the type of thinking that Pakistani universities have made a formal commitment to inculcate.

Thirdly, at the international level, the research offers a developing-country perspective that is a worthwhile addition to the literature mostly dominated by research from North America, Europe, and East Asia. The Pakistani higher education environment - with its unique combination of resource limitations, pedagogical methods, socioeconomic diversity, and fast-growing technological culture among the youth - is truly unique. In fact, features of such a context may be useful not only for others...

Lastly, the paper matter for its message about the wider connection between technology and education. The issue of how new information technologies influence human thinking is not a new one-it was the concern of the printing press, television and internet. But, the question is being asked with special urgency about AI, since AI's capability to replace human reasoning is much more direct than any previous technology. Finding out if being literate in AI can act as a protective factor - leading to critical thinking rather than dependency - is a real educational issue.

5. Problem Statement

Even though AI tools have been rapidly and widely embraced by Pakistani university students, no empirical study has investigated the relationship between AI literacy and critical thinking in this population, so far. It is a considerable void and, in fact, it has real consequences. Universities that are unaware of the cognitive aspects of their students' use of AI are the least capable of responding to it, either by developing it constructively or by dealing with its risks.

The situation is exacerbated by what could be called the illusion of competence. Students who rely on AI tools regularly might raise a false sense of AI fluency (a feeling that they understand the technology) that is not in correspondence with their conceptual knowledge over the long run. Laupichler et al. (2022) identified this discrepancy between perceived and actual AI literacy in European university students; Because of this, it is very likely that there is a similar if not bigger gap in Pakistan where formal AI education is hardly available and students' knowledge about AI systems mainly comes from their own, unreflective self-education.

Besides, this has an immediate impact on critical thinking. Rapidly and highly, students at Pakistani universities have really gone for AI tools but as yet, no research paper has done the job of exploring how critical thinking changes with AI literacy in the population of university students in Pakistan. It is a big gap in knowledge and, in fact, these educational institutions are most at risk of being unproductive and stale if they do not realize the cognitive side of the students' use of AI, both for the positive development of the tool and for limiting the hazards.

The situation is made worse by what one might call the illusion of competence. Students who habitually use AI tools might develop an unfounded confidence in their AI skills (a feeling that they comprehend the technology) which does not align with their actual understanding over time. Laupichler et al. (2022) found a significant misalignment between perceived and actual AI literacy among European university students; Because of this, it is very likely that the situation is similar if not worse in Pakistan where there is hardly any formal AI education and students' knowledge about AI systems is mainly derived from their own unreflective self-education.

And, this has a direct impact on critical thinking. Students will almost Of course be unable to critically analyze information provided by AI if they cannot accurately judge the reliability of an AI output, for example, if they do not know what training data is, what a knowledge cutoff means,

or how an algorithm can be biased. In a way, they are thinking with a tool they do not understand. The consequences can be as superficial as a paper with made up references or as serious as making life-changing decisions.

This study addresses the specific problem of a lack of baseline data on AI literacy at UoG, absence of empirical evidence on the AI literacycritical thinking relationship in Pakistani higher education, and As a result the inability of curriculum designers and policymakers to make evidence-based decisions about integrating AI literacy in undergraduate education. The study is aimed at handling all three aspects of the problem.

6. Literature Review

6.1 Defining and Conceptualizing AI Literacy

AI literacy as a separate topic first appeared in academic publications mostly from 2019 to 2020, although its theoretical foundation can be traced back to studies on digital literacy, computational thinking, and media literacy. The detailed discussion of AI literacy so far has been done by Long and Magerko (2020) who carried out an in-depth review of the concept by studying various AI literacy setups and educational resources. They found 17 basic skills grouped under three main capacities: the ability to recognize what AI is, the understanding of the working principles of AI, and the ability to critically assess and engage with AI systems. Besides, their model Mainly highlights the point that AI literacy is a broad human skill rather than an exclusive technical skill for STEM students since AI systems influence daily life in all social sectors. Through the AI4K12 project, Touretzky et al. (2019) came up with five major concepts in AI they believed every literate person should know: perception (how computers detect the world), representation and reasoning (how AI depicts and processes information), learning (how AI systems get better from data), natural interaction (how humans and AI communicate), and societal impact (how AI changes society). Even though their model was meant for school education, their requirement for mixing technical knowledge with ethical and societal consciousness is valid at the university level as well.

Recent publications have been mostly aimed at studying AI literacy measurement among college students. Laupichler et al. (2022) created and tested a tool for measuring AI literacy of university students and after research came up with 4 areas: technical knowledge of AI, awareness of AI applications, ability to evaluate AI outputs, and moral reasoning. Their German validation research showed good psychometric qualities and also a regular pattern of results in different student samples: self-rating of AI literacy was always higher than the assessed literacy level, implying that students often overestimate their own skills when there is no formal teaching.

Research on AI literacy is an emerging field in Pakistan and South Asia. Alam (2021) carried out a study on digital literacy in Pakistan at the university level. He revealed the existence of great disciplinary differences with social science and humanities students reporting very low digital self-efficacy compared to engineering and computer science students. These disciplinary differences will most probably show up in AI literacy rates and So impact directly the ways AI literacy initiatives are planned and implemented.

6.2 Critical Thinking in Higher Education

Critical thinking is by no means the last word here; after all, it is one of the most broadly researched constructs in educational psychology and also one of the most debated. The definition proposed by Facione (1990) in the landmark Delphi Report – purposeful, self-regulatory judgment that involves interpretation analysis evaluation inference explanation, and self-regulation – remains the most cited in the research literature and serves as the conceptual basis for the most frequently used assessment instruments.

Still, Paul and Elder (2001) developed the scheme further by the inclusion of a role of intellectual virtues such as humility perseverance integrity, and empathy, in the endorsement of critical thinking over a long period of time. Their theory has been a significant point of reference when talking about higher education mainly because it sees critical thinking not as a single skill but as a disposition, something one develops through sustained

practice and reflective engagement rather than something one acquires through a single course.

Developing critical thinking in universities in Pakistan is becoming a hot topic among researchers, partly because the Higher Education Commission (HEC) has made it mandatory that all bachelor's programs promote critical thinking as one of the qualities in graduates. Documentation of the persistent gap between this formal pledge and the real teaching practice has been done by Naz and Murad (2017) and Shaheen (2010). They have highlighted that the main reasons for these gaps are teaching geared towards examinations, large number of students in a class, teachers' limited training in active learning methods, and cultural norms which limit students' questioning and debating their teachers (who are viewed as authority figures). These major barriers imply that even if one intends to foster critical thinking, one still encounters a lot of difficulties- which is why the new methods, like AI literacy education, may Actually be helpful as they present a totally different way of engagement.

6.3 Digital Literacy, AI Literacy, and Critical Thinking: Theoretical Connections

The logical underpinning for a relationship between digital/AI literacy and critical thinking is quite solid. For instance, Eshet-Alkalai (2004) mentioned that digital literacy is a critical competence by nature because it incessantly entails evaluating sources, identifying affected content, and assembling the meaning out of a piece of fragmented and often contradictory information. Same thing, Buckingham (2007) suggested that media literacy, i.e. the ability to read analyze evaluate and create media, is a type of critical literacy that calls for the interrogation of representations, questioning of authorship and recognition of ideological systems.

AI literacy is a continuation of these lines of thought, yet it includes a few distinctive characteristics. AI-generated works rather than the ones created by older digital technologies do not in any way indicate that they are constructed - They act like simple answers. A Wikipedia article can be evaluated due to its visible construction (citations, history of edits, talk page); a ChatGPT response is totally without such a layer of transparency. This fact On top of making AI literacy sharply dependent on critical thinking makes the user who is required to apply a critical structure on the content that does not plainly invite it. Then again, it indicates that learning how to assess AI outputs might be a really effective means of nurturing one's critical thinking abilities, just because the difficulty posed is greater and less apparent.

6.4 Empirical Evidence on AI Literacy and Critical Thinking

There is still not much empirical research checking the direct link between AI literacy and critical thinking, but several studies of the last few years can work as useful references. For example, Ng et al. (2021) did a study to find out what AI literacy competencies secondary and post-secondary students in Hong Kong have. They found that students who understood AI concepts better were generally able to come up with more sophisticated reasoning about whether information is credible and the reliability of sources. What is more, the authors found that this influence existed irrespective of the academic disciplines, which is why they excluded it being an effect of STEM education.

Seo and his colleagues (2021) studied the relation between AI education and computational thinking among students in South Korea and their findings revealed that those who received formal AI instruction made advances in logical reasoning and problem decomposition that were Much greater than the control group. Besides being highly statistically significant, the effect sizes were also large enough to be of practical importance, This way providing solid evidence that AI literacy education on a formal level can be used as a critical thinking strategy.

Riedl (2019) brought up a solid point that should really determine how we take these results: simply exposing kids to AI tools doesn't guarantee that they will automatically get AI literacy or critical thinking improvements. In fact, it is through a combination of teaching and reflection that really lead to the right conditions for learning. Students who use AI tools but do not have knowledge of these i.e. the majority, as most studies do not

exhibit the evaluative skills that AI-literate students are capable of demonstrating. This insight So leads naturally to direct consequences for the current study situation: heavy AI tool usage by UoG students by itself is not a marker of either high AI literacy or strong critical thinking.

Kim and Kim (2022) contributed another angle to this discussion by showing that engineering and computer science students performed better on the technical elements of AI literacy, but social science and humanities students actually showed a bigger relative increase in learning from AI literacy instruction. So, what this points to is a scenario of floor-and-ceiling which has practical importance for the design of programs: it may well be that it is the non-STEM students who get the most out of an AI literacy program.

6.5 Gaps in Literature

There are a few reasons why this study is necessary based on the literature. Firstly, almost all research on AI literacy has been done in high-income countries like the United States Germany South Korea, and Hong Kong. Still, little has been written about developing countries whose constraints in resources, cultural attitudes towards teaching and learning, and technology access patterns are very different. Secondly, no research has so far focused on the AI literacycritical thinking link in a public university of Pakistan. Third, most research work studies only one method, either quantitative or qualitative; the present study's mixed-methods design makes it possible to understand not only the strength but also the ways in which the relationship works. These gaps together justify the research and give it significance beyond its immediate institutional context.

7. Research Methodology

7.1 Research Design

This study uses a mixed methods sequential explanatory design (Creswell & Plano Clark, 2018) wherein quantitative data is collected and analyzed in the first phase and then followed by qualitative data in the second phase to explain or elaborate the quantitative results. This design was chosen for two reasons. First of all, the main research question, "Is there a relationship between AI literacy and critical thinking, is best suited to quantitative data and statistical analysis. Third, the relationship, and the reasons for and how it relates, would be best understood with the kind of rich, contextual data qualitative interviews are likely to generate. The sequential design allows for the qualitative phase to be guided by the quantitative findings, as well as to be responsive to it, and it does not take place in parallel without meaningful integration into each other.

7.2 Research Paradigm

The study is based on a pragmatist epistemological model, which views the worth of a research method in its ability to address the research question rather than sticking to a particular philosophical tradition. Pragmatism advocates the combination of quantitative and qualitative methods as it does not agree with the assumption that these structures are mutually exclusive. On the contrary, it acknowledges that they can offer different perspectives on the same thing (Tashakkori and Teddlie, 2010).

7.3 Population and Sample

The population for this research project includes all 4th semester undergraduate students at the University of Gujrat in Spring 2026 semester. Monitoring critical thinking development of students is an excellent option at 4th semester since this is the time when students move towards a higher level of independent academic engagement. They would have finished the foundation courses by this time.

A stratified random sampling of 150 students was conducted. Stratification was done by academic faculty (Social Sciences, Natural Sciences Commerce Computer Science and IT) and gender. Stratification ensured proportional representation of these groups and also made the subgroup comparisons more meaningful. This sample size is larger than the minimum one recommended by Tabachnick and Fidell (2013) for the multiple regression analysis used in the study ($N > 50 + 8m$, where m is the number of predictors; with four predictors, $N > 82$; actual N

= 150), Because of this giving the study a strong statistical power.

For the qualitative part, a purposive sample of 15 students was drawn from the quantitative sample. These students were selected to represent diversity in AI literacy scores (low medium high), academic faculty, and gender. Purposive sampling in qualitative research is not concerned with statistical representativeness but rather with gathering the broadest range of experiences and perspectives.

7.4 Research Instruments

There were two established instruments that were translated and modified for the quantitative phase of the study. AI literacy was gauged by the AI Literacy Scale which was created by Laupichler et al. (2022) and Later modified for the Pakistani cultural and academic setting. The process of modification included these steps: the instrument was first evaluated by three faculty members who are experts in educational technology and AI; next, two Pakistani teachers checked a culturally appropriate version; and finally, Urdu versions of unclear English items were prepared. The modified instrument is divided into four subscales: general knowledge of how AI works (8 items), knowledge of AI applications (7 items), judging AI outputs (8 items), and ethical issues for AI (7 items). Altogether, there are 30 items on the scale which are answered on a 5-point Likert scale. As part of a composite instrument, critical thinking was assessed partially by the Watson-Glaser Critical Thinking Appraisal (Watson and Glaser, 1980) and the California Critical Thinking Skills Test (Facione, 1990). Firstly, the instrument had items measuring inference (ability to draw reasonable conclusions from evidence), evaluation of arguments (ability to consider the strength and relevance of arguments), and interpretation (ability to judge whether the conclusions follow the stated premises). There were 25 items consisting of multiple-choice and short-answer formats. Preliminary research shows that the hybrid instrument is reliable (Cronbach's alpha = 0.81).

For qualitative data collection, a semi-structured interview schedule was created which covered three content areas: students' reported use of AI and literacy perceptions; their description of AI influence on their thinking and reasoning; and their experiences in checking AI-generated materials for accuracy, bias, and trustworthiness. The schedule had eight open-ended questions with probes and before administration was also checked for clarity and non-leading wording by two qualitative researchers.

8. Data Collection

8.1 Overview of Data Collection Procedures

Data collection was undertaken at two separate stages, in line with mixed-methods sequential explanatory design which is the system for this study. The qualitative phase came after the quantitative phase, and the decision on which qualitative participants to select was based on quantitative results. Table 1 summarizes the whole data collection process from a different viewpoint and also lists the instruments used, participant groups, timelines, and administration formats.

Table 1

Overview of Data Collection: Phases, Instruments, Participants, and Administration

Phase	Data Type	Instrument / Tool	Participants	Period	Format
Phase 1 (Quantitative)	AI Literacy	AI Literacy Scale (Laupichler et al., 2022) — adapted, 30 items	150 undergraduates (stratified)	March 2026 (3 weeks)	Paper survey, in-class

Phase	Data Type	Instrument / Tool	Participants	Period	Format
Phase 1 (Quantitative)	Critical Thinking	Watson-Glaser & CCTST composite — 25 items (MCQ + short answer)	150 undergraduates (stratified)	March 2026 (3 weeks)	Paper survey, in-class
Phase 1 (Quantitative)	Background Variables	Demographic questionnaire — faculty, gender, AI use frequency	150 undergraduates	March 2026 (3 weeks)	Paper survey, in-class
Phase 2 (Qualitative)	In-depth Accounts	Semi-structured interview guide — 8 open-ended questions	15 purposively selected students	April 2026	Individual interview, audio-recorded
Pilot Testing	Instrument Refinement	Both instruments (pre-main study)	20 4th-semester students (not in main sample)	February 2026	Paper survey, convenience sample

Note. All participants gave their written informed consent before the data was collected. The interviews were done in English, Urdu, or a combination of both as what each participant wanted. AI Literacy Scale was modified from Laupichler et al. (2022); Critical Thinking tool was taken from Watson and Glaser (1980) and Facione (1990). MCQ = multiple-choice question; CCTST = California Critical Thinking Skills Test.

8.2 Quantitative Data Collection

Quantitative data was gathered over a period of three weeks, in March 2026. Before the data collection, permission was secured from the university's research ethics committee, department heads concerned, and instructors of individual courses. During the class sessions, few pen-paper based surveys were conducted to assure the highest number of responses and students who don't have the internet connection at home will still be reachable. Students received a piece of paper stating the purpose of the research, informing that participation is voluntary, anonymous and that they can leave at any time without any penalty. Out of 150 students who were approached, all of them agreed to participate in the study, So the response rate was 100%.

After the original order was changed in different classrooms to avoid the effect of the order, each student took the AI literacy scale and the critical thinking instrument back-to-back. The total time needed for the survey was 30 minutes. Academic faculty gender year of study (verified to be 4th semester), and one question measuring frequency of AI tool use by a five-point scale ranging from never to daily were background data collected.

8.3 Qualitative Data Collection

A sample of 15 students was selected for semi-structured interviews, five from each faculty from the low, medium and high AI literacy tertiles and

stratified by gender. 15 all agreed to participate. Interviews were held in a private room on campus in a 30- to 45-minute period, alone. All interviews were recorded on audio recorder, and with participants' informed consent. Interviews took place in the language most comfortable for the interviewees — English, Urdu or both — to reduce the likelihood of language barriers impeding the full representation of interviewees.

All audio recordings were transcribed immediately after each interview within 48 hours. Interviews in Urdu or mixed language were translated into English by the research team. Transcripts were provided for member-checking (to review or clarify or retract statements) and a few minor corrections were made before the analysis.

8.4 Pilot Testing

Before the main data gathering, both tools were tested on a pilot basis together with 20 fourth semester students who were a convenience sample and apart from the main study sample.The pilot helped in three ways: finding confusing or unclear items, measuring the time for completion, and presenting preliminary estimates of reliability. After pilot feedback, four survey items were changed to be clear and one was eliminated due to poor item-total correlation ($r < 0.20$). The interview guide was also improved as pilot participant remarks about question clarity and relevance. Final pilot results of reliability were good enough for both instruments, This way justifying their use in the main study.

9. Reliability and Validity

9.1 Quantitative Reliability and Validity

Internal consistency reliability of both tables was measured by Cronbach's alpha. The AI Literacy Scale showed excellent internal consistency in the main study sample (alpha = 0.83), and the critical thinking instrument showed a good level of reliability (alpha = 0.81). Both numbers are higher than the widely recognized 0.70 cut-off point for research instruments (Nunnally, 1978), which supports the reliability of the scales.

To confirm construct validity, CFA analysis was done with the help of AMOS software. Four-factor model of AI Literacy Scale (technical understanding, application awareness, output evaluation, ethical reasoning) was an acceptable fit for the data (CFI = 0.91, RMSEA = 0.06), meaning that the adapted scale maintained the intended factor structure well in the Pakistani student sample. Content validity was reaffirmed during the adaptation process by the expert review outlined in Section 7.4.

To check convergent validity, correlations between constructs that are theoretically related were explored. AI literacy was positively correlated with the self-reported frequency of reflective AI use ($r = 0.43$, $p < 0.001$) and with students' academic performance in courses involving information evaluation ($r = 0.38$, $p < 0.001$), This way supplying more proof that AI literacy scale is In fact measuring the intended construct.

9.2 Qualitative Trustworthiness

Qualitative trustworthiness was ensured through four strategies corresponding to Lincoln and Guba's (1985) criteria of credibility transferability dependability, and confirmability. Credibility was enhanced through member checking, prolonged engagement with participants, and triangulation across quantitative and qualitative data sources. Transferability was taken care of by thick description of the study context, participants, and procedures, which gave readers the freedom to see if findings can be applied to other situations. Dependability was illustrated by the use of a very detailed audit trail of all the analytical decisions puzzled out. Confirmability was a matter of the lead researchers journaling reflections and doubts about one's assumptions and possible biases which were explicitly documented throughout the research process.

Two researchers independently coded the same set of six transcripts for classification of qualitative data to check the consistency of coding between different researchers (inter-rater reliability). Afterwards, they compared their coding schemes and calculated Cohen's kappa, a statistic that describes the level of agreement between two raters in evaluating qualitative items, which yielded a value of 0.78, indicating a good level of agreement for qualitative research (Landis and Koch, 1977). The discrepancies were resolved through

discussion, and the final coding setup was applied to the transcripts that had not yet been coded by the lead researcher.

10. Data Analysis

10.1 Quantitative Analysis

Quantitative data were entered into SPSS 26 and were thoroughly checked and cleaned before analysis. Checks for missing data did not find any significant patterns of missingness (< 2% across all items), and missing values were filled by series mean imputation. For all variables, summary statistics like means, standard deviations, and ranges were computed. Normality of the distributions was tested with the Shapiro-Wilk test and by examining the histograms and Q-Q plots; both AI literacy and critical thinking scores were nearly normally distributed, which means we could use parametric tests.

To start, Pearson's correlation was used to find the relationships between AI literacy and critical thinking, as well as between all the other variables in the study. We also calculated subgroup correlations to check if the connection between AI literacy and critical thinking changed according to the academic faculty or gender of the participants. Then, using multiple regression analysis, we took critical thinking as the dependent variable and considered AI literacy, academic faculty, gender, and frequency of AI tool use as predictors. We ran all the standard regression assumption tests: residuals' normality (Shapiro-Wilk), homoscedasticity (Breusch-Pagan test), errors' independence (Durbin-Watson statistic = 1.94, which is an acceptable value), and multicollinearity absence (VIF values all below 3.0), respectively. All tests confirmed that the assumptions were satisfied.

10.2 Qualitative Analysis

The qualitative data analysis method utilized was thematic analysis that was guided by Braun and Clarke's (2006) six steps: (1) becoming acquainted with the data by repeatedly reading the transcripts and identifying first thoughts; (2) producing initial codes by systematically tagging the aspects of the data that are relevant to the research questions; (3) determining themes by assembling codes into possible theme groups; (4) examining themes about the coded data and the entire data set to confirm that they are consistent, different, and well-supported; (5) characterizing and naming themes with clearly defined margins; and (6) writing up the analysis by choosing illustrative quotations and developing an analytical narrative. Our analytical method was inductive in that the themes were derived from the data rather than being dictated by an existing theoretical system. At the same time, the analysis was theoretically informed by the AI literacy and critical thinking concepts that were reviewed in the literature. This combination of induction and theoretical sensitivity is a feature of reflexive thematic analysis as explained by Braun and Clarke (2019).

10.3 Integration of Quantitative and Qualitative Findings

Integration of the two data strands occurred at the interpretation stage, following the standard practice for sequential explanatory designs. Qualitative themes were examined for the extent to which they explained, elaborated, or challenged the quantitative findings. Where quantitative and qualitative findings converged, this was taken as evidence of stronger validity. Where they diverged, this was taken as an invitation for more nuanced interpretation, with the qualitative data used to contextualize and complicate the statistical picture.

11. Findings

11.1 AI Literacy Levels Among 4th Semester Students

The overall mean AI literacy score of all 150 participants was 58.4 (SD = 12.7), which signifies a moderate level of AI literacy. This indicates that there is a great amount of change across the four subscales. Students' familiarity with AI-powered tools like search engines, recommendation systems, and chatbots was highest on the application awareness subscale (M = 67.2/100). They did worst on the output evaluation subscale (M = 44.8/100) that measured their critical evaluation of the accuracy, bias, and reliability of the information they receive from AI — the aspect

of critical thinking most closely related to the use of AI.

Significant disciplinary differences were found ($F(3, 146) = 18.42$, $p < 0.001$). Students in Computer Science and IT scored highest ($M = 72.3$, $SD = 10.1$), followed by Natural Sciences ($M = 61.4$, $SD = 11.8$), Commerce ($M = 54.7$, $SD = 12.4$), and Social Sciences and Humanities ($M = 49.1$, $SD = 13.2$). Post-hoc Tukey tests revealed significant differences between the Computer Science group and all other groups, between the Social Sciences group and the Computer Science and Natural Sciences groups. However, gender differences were not statistically significant when academic faculty was controlled for ($F(1, 148) = 1.87$, $p = 0.17$). Descriptive statistics for AI literacy by subscale and academic faculty are shown in table 2.

Table 2

Descriptive Statistics for AI Literacy Scores by Subscale and Academic Faculty (N = 150)

Subscale / Group	M (Score/100)	SD	Min	Max	n
Overall AI Literacy	58.4	12.7	29	91	150
Application Awareness	67.2	11.4	38	96	150
Technical Understanding	60.1	13.5	27	93	150
Ethical Reasoning	58.8	13.1	25	90	150
Output Evaluation	44.8	14.2	18	82	150
By Faculty — Computer Science & IT	72.3	10.1	47	91	38
By Faculty — Natural Sciences	61.4	11.8	35	88	38
By Faculty — Commerce	54.7	12.4	29	83	37
By Faculty — Social Sciences & Humanities	49.1	13.2	24	79	37

Note. Scores are expressed on a 0–100 scale. Faculty groups comprise n = 37–38 participants each. $F(3, 146) = 18.42$, $p < .001$ for between-faculty differences; gender differences were non-significant after controlling for faculty ($p = .17$).

11.2 Critical Thinking Scores

On average, critical thinking scores was about 61.2 in 100 ($SD = 14.3$). Following are the subscale means: inference, $M = 65.4$ ($SD = 13.1$); interpretation, $M = 62.6$ ($SD = 14.8$); and evaluation of arguments, $M = 55.8$ ($SD = 15.2$). The relatively low score on evaluation of arguments is consistent with previous studies on Pakistani university students and is indicative of the complexity of the task of evaluating the logical structure of arguments, unlike just deriving conclusions from the given information. Table 3 shows the descriptive statistics for critical thinking subscales.

Table 3
Descriptive Statistics for Critical Thinking Subscale Scores (N = 150)

Critical Thinking Subscale	M (Score/100)	SD	Min	Max
Overall Critical Thinking	61.2	14.3	28	94
Inference	65.4	13.1	31	96
Interpretation	62.6	14.8	27	95
Evaluation of Arguments	55.8	15.2	22	90

Note. Scores are expressed on a 0–100 scale. Instrument adapted from Watson and Glaser (1980) and Facione (1990); Cronbach's alpha = .81.

11.3 Relationship Between AI Literacy and Critical Thinking

Pearson correlation analysis showed a statistically significant positive correlation between overall AI literacy scores and critical thinking scores ($r = 0.54$, $p<0.001$). By the conventional measure of "effect size" (Cohen, 1988) this is a moderate to strong effect size, and suggests a meaningful increase in students' critical thinking competency with increased AI literacy. The relationship was noteworthy for each academic faculty studied, although it was highest for the social sciences students ($r = 0.61$) and slightly lower for the computer science students ($r = 0.44$).

The subscale correlations revealed that output evaluation of the AI, as one of the dimensions, had the strongest association with critical thinking ($r = 0.59$, $p < 0.001$). Second was ethical reasoning ($r = 0.51$, $p < 0.001$), followed by technical understanding ($r = 0.43$, $p < 0.001$) and finally application awareness ($r = 0.31$, $p < 0.001$). The more evaluative and reflective features of AI literacy are closely linked to critical thinking, while the dimension with the highest score, which is general awareness of the tools, has a relatively weak relationship with critical thinking.

When running a multiple regression analysis it was found that, regardless of the academic faculty, gender, and AI usage frequency, AI-literacy was something important in predicting critical thinking (beta = 0.41, $t = 5.73$, $p < 0.001$). The entire model explained 38% variance in critical thinking scores ($R = 0.38$, adjusted $R = 0.36$, $F(4, 145) = 17.64$, $p < 0.001$). The full regression model is presented in Table 4.

Table 4
Multiple Regression Analysis: Predictors of Critical Thinking Scores (N = 150)

Predictor	B	SE B	β	t	p
AI Literacy (overall)	0.46	0.08	0.41	5.73	< .001
Academic Faculty	2.14	0.81	0.22	2.64	.009
Gender	0.88	1.12	0.06	0.79	.433
Frequency of AI Tool Use	0.72	0.94	0.07	0.77	.446
$R^2 = 0.38$ | Adjusted $R^2 = 0.36$ | $F(4,145) = 17.64$, $p < .001$					

Note. Dependent variable: Critical Thinking total score (0–100). Predictors entered simultaneously. B = unstandardized coefficient; SE B = standard error; β = standardized coefficient. All VIF values < 3.0; Durbin-Watson = 1.94.

11.4 Qualitative Findings: Themes

Thematic analysis of the 15 interviews produces three major themes, described below.

Theme 1 — Questioning the Machine.

Questioning the Machine. Students who scored higher in AI literacy illustrated frequently questioning AI results instead of blindly accepting them Actually. They talked about verifying AI-generated data with academic publications, figuring areas where an AI could get its data, and spotting AI responses which appeared overconfident or totally complete in a suspicious way. Awareness that language models are trained on data up to a certain point was the basis of one student's description of how that made her change the way of reading AI-generated content - she now automatically wonders if the information may be outdated or coming from unreliable sources. That habit of questioning is parallel to the inference and interpretation aspects of critical thinking.

Theme 2 — Recognizing Limitation.

Students with a high level of literacy identified AI limitations like biases in training data, hallucination, and knowledge cutoffs, which students with low literacy were not aware of. Their knowledge of the limitations of AI was not a mere abstraction as they could cite specific instances of AI errors and provide a general explanation of why such errors occur. Many of them also mentioned how finding an AI error - receiving a confident statement that turned out to be false - was a major change in the way they perceive AI. This understanding of limitation is directly linked to the critical thinking dimension of evaluation of arguments and involves the identification of assumptions that are not explicitly stated as well as the assessment of the quality of evidence.

Theme 3 — Intellectual Autonomy and Agency.

A third theme focused on the idea that participants made what they call thinking for themselves. Higher-literacy students, for example, considered AI as a tool that can provide not the final answer but only the first idea they would either form their own opinion before checking the AI or treat the AI output as a hypothesis to be tested rather than a conclusion to be accepted. On the contrary, students with lower literacy skills described AI more often as a time-saving tool that they used to get answers rather than to help with their own reasoning. Being ready to engage in intellectual autonomy is a way the self-regulatory dimension of Facione's (1990) critical thinking corresponds; it is also the way Paul and Elder's (2001) notion of intellectual independence is understood.

12. Conclusion

One of the main goals of this research paper was to explore if AI literacy is a determinant of critical thinking among 4th semester BS students of the University of Gujrat, and the result - both statistically and qualitatively - is a definite yes. Students with high level of AI literacy are found to be more thoughtful, inquisitive and autonomous in their interaction with information. Such a linkage is not simply a correlation to few statistical figures; in fact, qualitative research shows cognitive ways through which becoming literate in AI leads to strengthening and activating critical thinking dispositions.

At one point Still the research results in a big and discomforting discovery, to some extent. The aspect of AI literacy that is most strongly linked to critical thinking-the skill to judge AI is also the one on which students rate themselves the lowest. Students know about AI tools well; yet, majority of them do not critically evaluate them. They only get the exposure without any structure. Working on this gap is the main challenge the research study has identified.

Besides that, it also confirms a matter which those thinking of critical thinking development as a desirable goal could be quite pleased by, that the association between AI literacy and critical thinking is highest among students who before

learning about AI literacy were the poorest critical thinkers. Social Sciences students had the strongest correlations between AI literacy and critical thinking gains, indicating that students without formal instruction in systematic reasoning will benefit to a larger extent in overall critical thinking by first learning to think carefully about AI.

Being a cross-sectional study, cause-effect can hardly be definitively established from correlation data alone. Yet, the combination of quantitative and qualitative evidence of two different data sources, plus the theoretical unity of the identified mechanisms, strongly supports the argument that AI literacy and critical thinking are not only correlated but also developmentally connected. Further longitudinal and experimental research will be necessary to verify this connection and to determine which interventions will be most effective in producing it.

13. Recommendations

The Following the results of this research, different groups of people involved in Pakistani higher education are given these proposals.

To the University of Gujrat: The first and foremost task for the university is to design and offer an AI literacy course as a part of its general education program that is accessible to students from all faculties. The contents of this course should not be limited to the technical aspects of AI such as machine learning, recommendation algorithms, and training data but should also address the criteria and ethical aspects of AI: verification of AI results for correctness and bias, recognition of AI limitations, and use of AI as a cognitive support rather than a cognitive substitute. Since the disciplinary differences in this study demand, special efforts should be made to ensure that students of Social Sciences and Humanities have equal opportunities for AI literacy.

To the Faculty and Instructors: Regardless of the discipline, AI literacy should be integrated as a component of the curriculum and made explicit in the teaching. The level of AI literacy achieved is to be adjusted by instructors to the specific AI tools and challenges of their field. It may seem that instructors need to be changed into AI specialists without understanding the main requirement which is that they should demonstrate a critical attitude towards AI. So, they should raise questions about AI's outputs in the class, use AI-produced content as a basis for testing tasks, and motivate students to research rather than simply using AI tools. The instructors should be supported with professional development programs aimed at fostering the skills and building the confidence assisting them to be more efficient.

To the Curriculum Designers: Treating AI literacy as a technical subject unrelated to other competencies during curriculum reform is a mistake. Scenario-based activities that involve students in evaluating AI-generated materials, identifying misinformation, and making decisions are good opportunities to incorporate AI literacy in the existing courses without having to introduce entirely new courses. Assessment techniques that appreciate a critical approach towards AI rather than simply penalizing the use of AI (would) change the incentive mechanisms in a way beneficial to the overall scheme.

Higher Education Commission of Pakistan: Incorporating formal AI literacy standards as part of the graduate attributes structure is one idea that the HEC can take into account alongside the existing critical thinking standards. Such a set of standards would not only guide institutions in Pakistan but also highlight the national significance of AI literacy as a key educational focus. Besides, HEC should finance the studies which follow the changes in AI literacy among undergraduate students for a long time, so that the standards can be changed based on actual data over time.

For Future Researchers: In reality this paper is based on the cross-sectional way means that the possibility for determining causality is limited. Researchers in the future should be able to follow the same students for a period of time (longitudinal study) to observe changes in AI literacy and critical thinking, and use of experimental studies to test specific AI literacy measures. Besides, the studies should also find out

whether AI literacy education shows different effects based on factors like gender, socioeconomic background, and prior educational experience, to design equitable programs.

References

Alam, A. (2021). Digital literacy and higher education in Pakistan: Exploring competencies, challenges, and policy implications. Journal of Education and Educational Development, 8(1), 45–67.

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.

Buckingham, D. (2007). Beyond technology: Children's learning in the age of digital culture. Polity Press.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

Dewey, J. (1910). How we think. D. C. Heath.

Eshet-Alkalai, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. Journal of Educational Multimedia and Hypermedia, 13(1), 93–106.

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

Higher Education Commission of Pakistan. (2020). Graduate attributes framework for Pakistani universities. HEC.

Kim, S., & Kim, J. (2022). Disciplinary differences in AI literacy among university students: Implications for curriculum design. Computers & Education, 181, 104468.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. Computers and Education: Artificial Intelligence, 3, 100101.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE Publications.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–16). ACM.

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). Faithfulness and factuality in generative abstractive summarization. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 1906–1919).

Naz, S., & Murad, H. S. (2017). Innovative teaching has a positive impact on the performance of diverse students. SAGE Open, 7(4), 1–8.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041.

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

Paul, R., & Elder, L. (2001). Critical thinking: Tools for taking charge of your learning and your life. Prentice Hall.

Riedl, M. (2019). Human-centered artificial intelligence and machine learning. Human Behavior and Emerging Technologies, 1(1), 33–36.

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner-instructor interaction in online learning. International Journal of Educational Technology in Higher Education, 18(1), 54.

Shaheen, R. (2010). Creativity and education. Creative Education, 1(3), 166–169.

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). Pearson.

Tashakkori, A., & Teddlie, C. (2010). SAGE handbook of mixed methods in social and behavioral research (2nd ed.). SAGE Publications.

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (pp. 9795–9799). AAAI Press.

Watson, G., & Glaser, E. M. (1980). Watson-Glaser critical thinking appraisal. Psychological Corporation.