



Sociology & Cultural Research Review (SCRR)
 Available Online: <https://scrrjournal.com>
 Print ISSN: 3007-3103 Online ISSN: 3007-3111
 Platform & Workflow by: *Open Journal Systems*
<https://doi.org/10.5281/zenodo.23295499>



Peripheral Intelligence in the STEM Classroom: Investigating Generative AI Readiness, Institutional Capacity, and Educational Equity in Secondary Schools of South Punjab, Pakistan

Tanzeela Zahoor

M.Phil English Scholar, Times University Multan.

tanzeelazahoor910@gmail.com

Faisal Maqsood

Assistant Education Officer, M Phil in Education from University of Education (Multan Campus).

faisal.mobiserve@gmail.com

Muhammad Atiq-Ur Rehman

PhD Scholar, English Department, University of Sialkot Punjab, Pakistan.

5200447uskt@gmail.com

Sadia Qadir

MPhil Scholar in English, University of Southern Punjab: Multan, Punjab, PK.

sadiaqadir214@gmail.com

Muhammad Anas Shakil

Student, Department of Chemistry Emerson University Multan.

anasshakil249@gmail.com

ABSTRACT

After 2020 the number of studies in AI for science, technology, engineering and mathematics (STEM) education has grown very fast, and the evidence base is currently geographically clustered in North America, East Asia and Europe. Peripheral education systems (in which infrastructure is lacking, language is complex, and secondary school completion rates have been low for historical reasons) are seldom considered as places where theory building occurs. To fill this gap, this mixed-methods study explores the readiness of secondary schools for generative AI (GenAI), capacity of institutions and STEM opportunity in Multan, Bahawalpur, and Dera Ghazi Khan, the three divisions of South Punjab, Pakistan. A newly designed Peripheral AI–STEM Capability Framework (PASCF) was developed, based on AI–Technological Pedagogical Content Knowledge, the Unified Theory of Acceptance and Use of Technology 2, and Sen's capability approach and we surveyed 486 STEM teachers and 812 Grade 9–12 science-group students, and conducted 48 semi-structured interviews with teachers, headteachers, and students. Measurement model was confirmed by confirmatory factor analysis. Structural paths revealed that institutional capacity ($\beta = .41$), greater than behavioral intention and ethical readiness, uniquely predicted AI–TPACK, which in turn, was the strongest predictor of classroom GenAI–STEM integration ($\beta = .38$). While the informal use of GenAI tools was common (71% of teachers said they used them out of class), the formal use of the tools aligned with the curriculum was rare (18%). The lowest Opportunity-to-Learn (OTL) scores were found for females in rural settings. Themes from the interviews uncovered two streams: a rich, informal digital hustle and a formally silent STEM classroom, limited in electricity, devices, language (Urdu/Saraiki/English) relations, assessment practices that discourage inquiry, and the lack of policies on AI integrity at the school level. The study presents a theory of peripheral AI–STEM capability that accounts for the context-specific

nature of this skill; an equity-centered measurement model to evaluate the same; and policy recommendations consistent with Punjab's upcoming STEM secondary-education reforms. The findings alert to the dangers of scaling GenAI without making structural access "pedagogical capability" to reduce, but not close, the opportunity gap for the South Punjab STEM.

Keywords: *Generative Artificial Intelligence, STEM Education, Teacher Readiness, AI-TPACK, UTAUT2, Capability Approach, Educational Equity, South Punjab, Mixed Methods*

Introduction

The past ten years of research on AI in STEM education paint a clear picture: the field is growing, it's a topic that is undergoing theoretical flux, and it's geographically uneven. Bibliometric syntheses of the years 2013–2025 indicate two stages: the first being an exploratory phase and the second, a period of strong growth from 2020 onward that is fueled by learning analytics, adaptive systems and, starting in the second half of 2022, by generative large language models (Yang et al., 2025). Reviews reveal four major pathways of use: design-based research refinement of exploratory STEM teaching, incorporation of new AI tools, studies of application in the classroom, and exploration of motivation and achievement (Yang et al., 2025). In the parallel literature, there is a strong focus on learning outcomes and the performance of tools, while the topics of ethics and flourishing are less developed, and learning about AI remains insufficiently developed (scoping review of the literature on generative AI in secondary STEM education, 2025).

Equally as importantly, these reviews set out an architectural absence, a silenced structure, in a remarkably uniform fashion. About 90% of empirical studies come from North America, East Asia and Europe (systematic syntheses 2025-26). Working in the long term is without limits. There is a smaller generative-AI subset in K–12 as compared to higher education. Engineering is not as thick as science and mathematics. Combined designs (large scale surveys plus interviews in authentic classrooms) continue to be rare. Implications of equity, gender, language and rurality are more commonly mentioned than are the constructs of the same. Context-sensitive and theoretical models, teacher-focused professional capacity studies and infrastructure and ethics as analytical variables and not background noise, therefore, call for methodological commentaries

South Punjab is a rare test case of those calls. The region (Multan, Bahawalpur and Dera Ghazi Khan divisions) has a significant science-group secondary population and has been reported to have inequalities in gender, supply and digital infrastructure in the rural and urban areas. The Asian Development Bank's Responsive, Ready, and Resilient STEM Secondary Education in Punjab Program (2026–2030) identifies southern districts as having significantly lower secondary completion rates (around 31% as opposed to 51%) and suggests creating climate-resilient STEM labs, leadership development and gender-inclusive provision (Asian Development Bank [ADB], 2025). Meanwhile, a grassroots digital skills initiative has made short video editing, freelancing and AI courses popular in Multan, Bahawalpur, Dera Ghazi Khan and Rahim Yar Khan with the help of video editing, freelancing and AI groups on WhatsApp and low-cost community trainers (Express Tribune, 2025). This informal ecology is supplemented by public platforms like DigiSkills.pk, e-Rozgar and NAVTTC. This leads to a paradox, namely that the current AI-STEM models, created in high-connectivity systems, are inadequate to explain: Informal exposure to GenAI is growing for youth, and pedagogical approaches in STEM classrooms remain analog.

While a small Pakistani literature has emerged that can get a little closer to adjacent questions (personalized learning perceptions among University teachers of South Punjab, 21st-century

skills and use of digital tools among University students, TPACK among prospective teachers, and the digital readiness of university students from rural and urban areas for AI; Hussain, 2025), such literature is still missing a multi-districts, secondary STEM, mixed-methods account of the relationship between teacher knowledge, technology acceptance, institutional capacity, and student opportunity. South Punjab is internally diverse, and it is impossible for the single district surveys of teachers to capture the internal diversity of this region since the AI ecology of a relatively connected urban high school in Multan is not the same as that of a girls' higher secondary school in Rajanpur or rural Layyah.

Research Gap and Study Purpose

The above three gaps are therefore the focus of the present study. There is first a geographic and a systemic gap: AI–STEM theory has been written from the center of the research world and then at best applied to the periphery. Secondly, a methodological gap: repeated throughout the reviews, there is a lack of designs that integrate measures based on the survey with accounts of how teachers and students live with GenAI in interviews. Third, a theoretical gap: TPACK and UTAUT2 explain knowledge, and intention is not enough to explain classroom practice when there are electricity, language of instruction, high-stakes board exams, and gendered device ownership in between teacher and lesson. Similar to the capability approach, Sen provides that missing conversion language: that resources are only converted to functionings through conversion factors.

We suggest and empirically investigate the Peripheral AI–STEM Capability Framework (PASCf). PASCf is influenced by the institutional capacity and structural conditions in which the teacher and the students are immersed and by the acceptance of AI–TPACK and UTAUT2; it sees student opportunity to learn as a capability outcome, not a mere access count. In the South Punjab secondary STEM context, the study poses not only the question of "will teachers use GenAI?" but also "when will informal digital hustle transform into formal STEM pedagogy and when will it not?"

Research Objectives

RO1. To assess teachers' levels of AI–TPACK, UTAUT2-based technology acceptance, institutional capacity, ethical preparedness, and classroom integration of GenAI in STEM education across local areas, gender, school sectors, and subject disciplines in South Punjab.

RO2. To examine the influence of institutional capacity, facilitating conditions, informal exposure to GenAI, AI–TPACK, behavioral intention, and ethical readiness on the classroom integration of GenAI in STEM education.

RO3. To investigate secondary school students' reported opportunities to learn with generative AI and examine variations in these opportunities by gender and locality in South Punjab.

RO4. To explore how teachers, headteachers, and students perceive the transition from informal, out-of-classroom uses of GenAI to formal STEM classroom integration, identifying the key factors, challenges, and opportunities that may facilitate or hinder the development of AI-related capabilities.

Research Questions

Four research questions guided the study.

RQ1. How are teachers' levels of AI–TPACK, UTAUT2 acceptance, institutional capacity, ethical preparedness of teachers with AI, and classroom integration of GenAI and STEM across local areas, gender, school sector and subject in South Punjab?

RQ2. How much influence do institutional capacity, facilitating conditions, informal GenAI exposure, AI–TPACK, behavioral intention and ethical readiness have on classroom integration of GenAI–STEM?

RQ3. How do secondary students of South Punjab describe and rate their opportunity to learn with generative AI and how is it patterned by gender and locality?

RQ4. What do teachers, headteachers and students mean by the transition from informal, out-of-classroom uses to formal STEM classroom uses, and what do they see as being important or problematic in terms of factors that could enable or hinder further capability growth?

Significance

The study contributes to STEM-education and educational-technology in four ways. Conceptually, it defines PASCf as a multilevel model which shifts the focus of AI–STEM research from the question of the effectiveness of tools to the question of capabilities and equity. Based on empirical evidence, it is the first multi-division mixed methods portrait of GenAI in secondary STEM of South Punjab, complementing 486 survey questionnaires for teachers, 812 survey questionnaires for students and 48 interviews. Theoretically, it demonstrates that the most crucial link between institutional capacity and classroom integration is through AI–TPACK, but informal exposure doesn't lead to ethical or pedagogical readiness. In practice, it provides a policy sequence (infrastructure, professional learning, assessment redesign, language-aware tools, school-level integrity rules), which aligns with the ADB STEM program and with digital-skills agenda in Punjab, but with a cautionary note on the possible negative impact on equity if the sequence is reversed.

Literature Review

The Expanding but Asymmetric Field of AI in STEM Education

Searching the Web of Science database, Yang et al. (2025) found 186 articles from the last ten years (2013–2023) covering various themes, including early-childhood applications, 3D printing, the ARCS motivation model, collaborative learning, and much more. They found that research activity concentrates in Asia, Europe and the Americas and that long-lasting international partnerships are rare. The mapping that takes place thereafter consist of five thematic poles: learning analytics, engineering education, AI literacy, immersive technologies, and generative AI. AI, VR/AR are still significant, but an AI system with a large language model like ChatGPT is now driving the recent impact of citations.

reviewed 55 studies in science education using generative AI as the technology and a five-dimensional technology-based learning framework (devices, domains, methods, participants, and research issues) to see that science education is the most popular area of study, that text-based GenAI tools outnumber image-, video-, and audio-based tools, and that higher education exceeds K–12 education. Studies of effectiveness in a quantitative fashion are the norm. While benefits of learning, teacher and instructional benefits are mentioned, barriers are only partly theorized (e.g., hallucination, over-reliance, leakage and unequal access). A review of 32 randomized trials of AI-enhanced personalized K–12 STEM learning found medium overall effects; the effects were higher in junior and senior high school and varied by type of AI and personalization model used. Adaptive and predictive systems demonstrate the most striking cognitive impacts when they are tightly integrated with curriculum; generative AI has been used widely and largely short-term or perceptual measures are used to evaluate its impacts.

These findings are valuable, but not exhaustive. The value of evidence is that it accumulates and provides support for tutoring, feedback, simulation and some personalization. Persistent gaps mentioned in the same reviews are the lack of longitudinal designs, inadequate classroom implementation research, poorly developed teacher professional-capacity studies, weak measurement of ethics and equity, limited engineering education and overwhelming geographic bias (Karthikeyan, 2026). The NARST community challenges statements with additional changes such as educational goals, renovation of science content that includes AI as an object of study, complexity of individualized assessment, teacher preparedness and bias in STEM education. A human flourishing lens also suggests that the secondary STEM research on generative AI has focused on how to enhance cognitive outcomes rather than expanded capabilities, new social models, and meaning making (OECD, 2024).

Teacher Knowledge and Technology Acceptance

There are two families of constructs that dominate teachers' explanations. The first is Technological Pedagogical Content Knowledge (TPACK; Mishra & Koehler, 2006) which has been expanded to AI-TPACK or Intelligent-TPACK as it is now known, to incorporate knowledge of AI tools, AI-supported pedagogy, and ethical-AI knowledge (Celik, 2023; Chiu et al., 2023). AI-TPACK is not a general digital competence. It demands that teachers determine when a language model is a suitable scientific interlocutor, how to structure language scaffolded prompts without reducing inquiry, and how to maintain the integrity of the assessment process as students can come up with a worked solution in a few seconds. The evidence collected in Pakistan for prospective teachers shows the wide-spread general use of technology and few specialized supports, inadequate laboratory facilities and technical assistance (Abbas and colleagues, 2024).

Another theory is acceptance and use theory. According to Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) by Venkatesh et al. (2012), the intention and use are influenced by the antecedents of performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. In a cross-cultural study on learning using AI enhancement among university students in Pakistan and Malaysia, it is observed that infrastructural and skill factors play a greater role in Pakistan, whereas motivational factors and conditions of "smart divide" are the most important factors influencing intention and intention influences learning gains and use. Teacher studies in Pakistan reveal that there is an unequal readiness among the schools with regard to the use of ChatGPT and Grammarly as the prevalent tools available and specialized educational AI is still not accessible.

TPACK, however, and UTAUT2, were not developed to account for failures of peripheral conversion. A teacher can get a moderate on performance expectancy ("ChatGPT helps my students see alternative solution paths") and not use a model during a physics period for any of the following reasons: They may have only one computer in the school, it may not work during the day, the board syllabus may be based on recall, or there may be no policy for dealing with homework using AI. These themes of rural-urban qualitative research with the teachers of Pakistan are identified as readiness themes: firstly, lack of a concept of AI among teachers and students; secondly, infrastructure as a gatekeeper; thirdly, a lack of motivation, but rather some lack of digital skill; fourthly, psychological and ethical ambivalence, and fifthly, the absence of a professional development institution, and its role as a missing workshop rather than a weak one (Hussain, 2025). That suggests that there are a number of institutional and structural levels to be nested within models of acceptance.

Equity, Capability, and the Peripheral School

The evaluative space is moved from resources or utilities towards the real freedoms people have in achieving valued functioning's, as in the capability approach of Sen (1999) and Walker & Unterhalter (2007). In education, the relevant functioning's are the ability to ask scientific questions, represent quantitative relationships, use digital tools without being humiliated and imagine a future with a STEM career. Resources such as a smart phone, a chatbot, a computer lab turn into capabilities via conversion factors such as gender norms in relation to device ownership, the language of the interface, teacher mediation, assessment rules and school leadership. It has been long established that there are three dimensions of digital-divide research: access, use, and outcome divides (van Dijk, 2020; Selwyn, 2016). The prompt-and-judgment divide is a new one introduced by GenAI. Two students with the same smart phone might have different capacities for interrogating a model's chemistry explanation, detecting a fabricated citation, and transferring a generated outline to a lab claim.

These conversion issues are worsened in South Punjab. There has been a significant variation in absolute numbers in the southern divisions, shortages in teacher count compared to sanctioned strength in the province of Punjab as a whole (fill rates of around three-quarters), and high district-level variation in the numbers of schools and teachers (Programme Monitoring and Implementation Unit, 2024, 2025; Bureau of Statistics, 2024, 2025). Large numbers of candidates for science-group boards and not equal distribution of high grade and functional labs. Internet penetration in Pakistan is still relatively low and average speed of Internet is still inadequate for Media rich pedagogy. This is a category error because in this context, GenAI should be considered a plug-in instructional enhancement tool. It's an ability intervention that will undergo uneven conversion.

The Informal–Formal Split in Pakistani Digital Learning

The South Punjab case is unusual, in that it has an active informal digital ecology and a relatively passive formal STEM pedagogy. The study on the use of digital tools and perceived development of 21st-century skills among university students conducted in four public universities in South Punjab revealed that the use of digital tools such as WhatsApp, YouTube, Chrome, ChatGPT, Meta AI (also known as Gemini) and office suites were high and the use of scholarly infrastructure including JSTOR, Zotero and structured MOOC platforms were low, while the use of digital tools was positively correlated to 21st-century skill development. At the community level, trainers in Multan and Bahawalpur report students that seek learning for earning money instead of just earning a certificate (Express Tribune, 2025). Pedagogically significant informal ecology. It is a phenomenon where many secondary students use GenAI for shortcuts or free-lance solutions to homework rather than as a scientific tool. Then teachers find themselves with an assessment problem with which they were not prepared and with a set of practices that they did not design.

The informal formal dichotomy has yet to become a major theoretical topic of high impact research on the topic of AI in STEM education. In South Punjab, such an approach helps to transcend the “adopters versus non-adopters” dichotomy and create a more nuanced ecology of shadow use, a classroom with a certain quietude and even a few pedagogical conversions.

Theoretical Framework

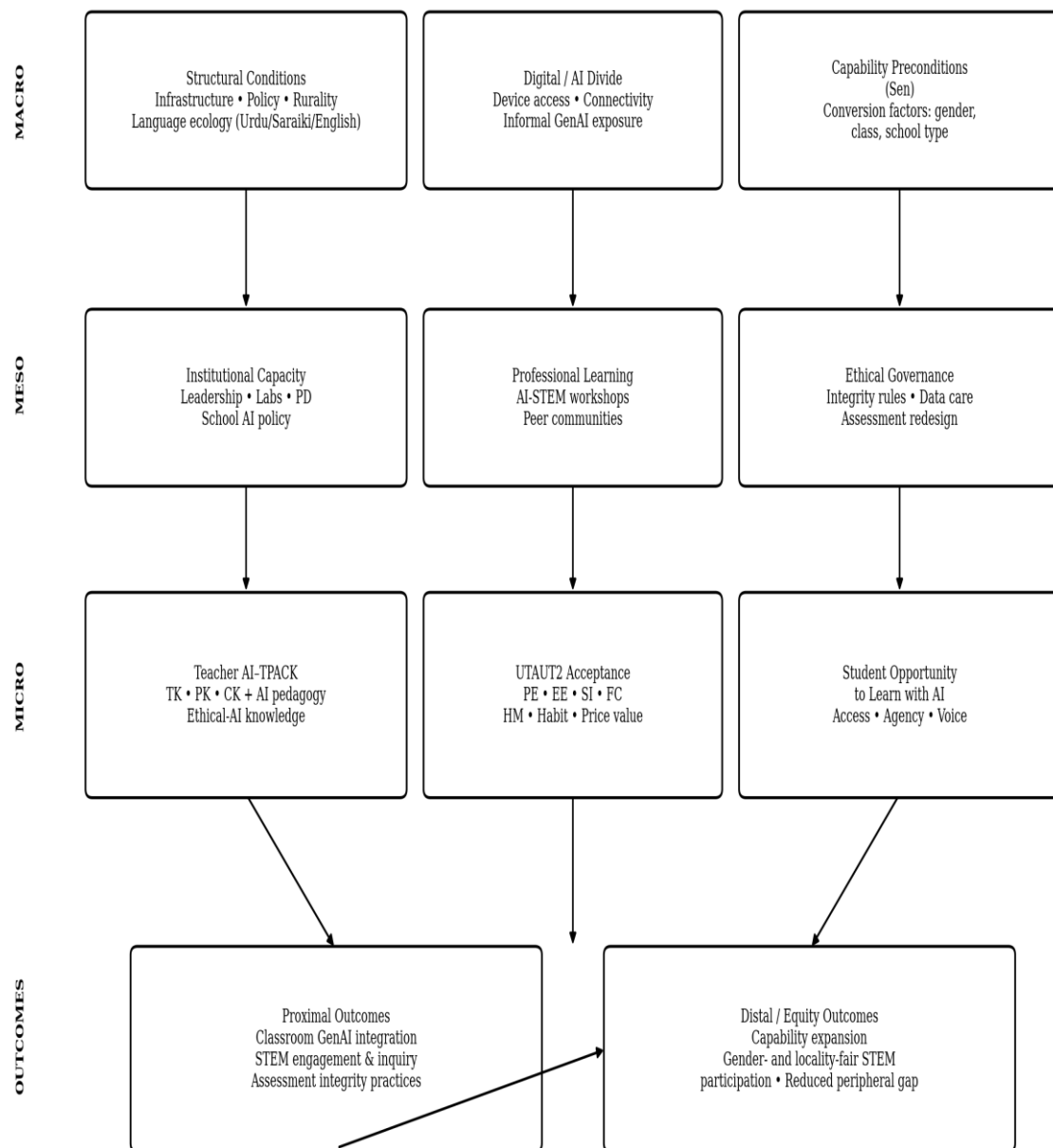
PASCF is a tiered model for peripheral systems with STEM. It doesn't eliminate TPACK or UTAUT2, it simply moves them. A digital and AI divide involving informal GenAI exposure is created by structural conditions such as electricity reliability, connectivity, transport, language ecology, and provincial policy as well as capability preconditions such as gender and socioeconomic status, school type, etc., at the macro level. At the meso level, institutional capacity (sense of leadership attention, laboratory function, professional development, and

written or tacit school rules) and professional learning opportunities, as well as ethical governance, mediate between the macro and school-level conditions to determine whether they will be realized in the school. At the micro level, teacher AI-TPACK, UTAUT2 acceptance and student opportunity to learn are proximal abilities. There are two sets of outcomes: proximal classroom practices (integration, inquiry-oriented STEM engagement, assessment-integrity routines) and distal equity outcomes (expanded functioning's, lessened gender and locality gaps).

PASCF is far from being a mere combination of established models, having three theoretical commitments. First, informal exposure is considered an antecedent which is capable of increasing performance expectancy and decreasing ethical readiness when taking the form of unexamined homework help. Second, AI-TPACK is hypothesized to be the main pedagogical converter: Without AI specific pedagogical content knowledge, the intention results in avoiding integrating AI with TCK or in uncritical tool display. Third, student opportunity to learn is an independent ability construct, rather than a reflection of teachers' self-reports. You might have a very confident computer teacher in a school and a very confident group of girls, but as long as you're in a rural campus, you have no private way of getting girls working with a model, no way to make it available in a language they can understand, and no way to make it available at the right time.

Figure 1. *Peripheral AI-STEM Capability Framework (PASCF).*

Figure 1
Peripheral AI-STEM Capability Framework (PASCF)



Moderators: locality (urban/peri-urban/rural), gender, public-private sector, STEM subject, prior informal AI use.

Note. PASCF integrates AI-TPACK, UTAUT2, and the capability approach for peripheral STEM systems.

Note. The framework nests teacher knowledge and acceptance inside institutional and structural conversion factors. Moderators include locality, gender, sector, STEM subject, and prior informal AI use.

Hypotheses

H1. Teacher AI-TPACK is positively predicted by teacher's institutional capacity. H2. Conditions facilitating positively predict behavioral intention. H3. The positive association of

informal GenAI exposure with ethical AI readiness is weak, and could be more strongly associated with uncritical use. H4. The relationship between AI–TPACK and behavioral intention, and the classroom integration is positive and significant. H5. AI–TPACK uniquely has the highest unique effect on classroom integration, while behavioral intention and ethical readiness equally positively predict classroom integration. H6. The classroom integration and opportunity to learn is less in rural schools and among female students, as compared to the sector. H7. Interview accounts will highlight factors for conversion (power, devices, language, assessment, and gendered mobility) that account for remaining differences that the survey model cannot explain.

Method

Design

We used a concurrent mixed-methods design (Creswell & Plano Clark, 2018) and equal emphasis was placed on a multi-level survey strand and a semi-structured interview strand. Integration was done at construct level, at the level of the participants in the interviews (in the surveyed schools), and at the level of interpretation (joint displays of path coefficients and themes). The design is cross-sectional with claims being associational rather than causal. The interview protocol included an explanatory sequential follow up which allowed probing of unexpected quantitative cells (high intention low integration, high informal use low ethical readiness) in the field.

Setting and Participants

The sampling frame included government and registered secondary (class 8) and higher secondary schools (class 12) imparting science-group education in nine districts of the three divisions of South Punjab: Multan, Khanewal, and Vehari (Multan Division); Bahawalpur, Bahawalnagar, and Rahim Yar Khan (Bahawalpur Division); Dera Ghazi Khan, Muzaffargarh, and Layyah (Dera Ghazi Khan Division). Districts were classified into strata based on both the official urban share and known laboratory density. In districts, schools were sampled according to the distribution of the number of students in their science groups and a sample oversampling girls' schools was obtained to facilitate gender analysis.

The teacher survey received 486 responses from 640 sent out (75.9%). Teachers taught physics (24.3%), chemistry (22.0%), biology (21.4%), mathematics (23.5%), and computer science or general science (8.8%). The mean teaching experience was 11.6 years (SD = 7.4). The female teachers were 46.1% of the teacher sample. Locality was 34.8% urban, 28.6% peri-urban, and 36.6% rural. 71.8% of the sample were public-sector teachers. A total of 812 of the 51.7% of the young women and 33.1% of the urban, 29.2% of the peri-urban, and 37.7% of the rural science-group students completed the student survey. Purposively sampled 24 STEM teachers, 8 headteachers/academic coordinators and 16 students (N = 48) who differed in locality, gender, sector and survey score profile.

Table 1. Sample Composition of the Teacher and Student Surveys.

Characteristic	Teachers (n = 486)	%	Students (n = 812)	%
Division: Multan	178	36.6	291	35.8
Division: Bahawalpur	162	33.3	274	33.7
Division: Dera Ghazi Khan	146	30.0	247	30.4
Locality: Urban	169	34.8	269	33.1
Locality: Peri-urban	139	28.6	237	29.2

Locality: Rural	178	36.6	306	37.7
Women / young women	224	46.1	428	52.7
Public sector	349	71.8	601	74.0
Private / other registered	137	28.2	211	26.0

Note. Percentages may not sum to 100 because of rounding. Student gender is reported as young women to reflect the secondary-school sample.

Instruments

Unless otherwise stated, teacher constructs were assessed on a 5-point Likert scale (1 = "strongly disagree" to 5 = "strongly agree"). The AI-TPACK dimensions (14 items) were adapted from the work of Celik (2023) on Intelligent-TPACK dimensions that encompass technological knowledge of GenAI, pedagogical knowledge for inquiry-oriented STEM, content-specific prompting, and ethical-AI knowledge. UTAUT2 was adapted to the context of GenAI in STEM lessons by using 4 items for each scale (Venkatesh et al., 2012): Performance expectancy, Effort expectancy, social influence, Facilitating conditions, Behavioral intention. Institutional capacity (8 items) included leadership support, laboratory reliability, access to professional development, and if any school policy had been established on AI. Ethical AI readiness (7 items) focused on plagiarism judgment, checking for hallucination, student-care of data, and fairness. Measuring Classroom GenAI-STEM integration was the planned use of the model, rather than its personal use (9 items). Informal GenAI exposure consisted of a 6-item frequency index (Never to Daily) for the use of ChatGPT or Gemini, WhatsApp bots, solvers applications and tools used outside class sessions.

The student instrument included the following measures: opportunity to learn with AI (12 items: device access, time, privacy, language comprehensibility, teacher mediation, permission to use tools for inquiry, not just product generation); STEM engagement (8 items); and a short ethics-awareness screen. Both were prepared in the English and Urdu languages and during administration, oral clarification was provided in Saraiki in Dera Ghazi Khan Division. The results of back-translation and a pilot of 36 teachers and 48 students in two non-sampled Multan schools were used to revise the items for the colloquial clarity (such as using the term "when the software invents a reference" or "when the software does a step").

Construct mapping was used to map interview protocols to PASC. Teacher interviews requested a recent lesson where AI was used or not used, for reports on student use of AI, and for assessments of board exams. The headteacher interviews were centered on issues of policy vacuum, pressure from parents and allocation of resources at school. Students' interviews, done in gender same pairs when specified, included the following questions: How does GenAI fit into homework, what languages work, and who handles the phone at home.

Table 2. *Measurement Properties of Teacher Survey Constructs.*

Construct	Items	M	SD	α	CR	AVE
AI-TPACK	14	3.38	0.74	.91	.92	.54
Performance expectancy	4	3.86	0.69	.88	.89	.67
Effort expectancy	4	3.31	0.81	.86	.87	.63
Social influence	4	3.12	0.77	.84	.85	.59
Facilitating conditions	4	2.92	0.94	.87	.88	.65
Behavioral intention	4	3.44	0.80	.90	.90	.70
Institutional capacity	8	2.86	0.88	.89	.90	.56
Ethical AI readiness	7	2.93	0.79	.85	.86	.52

Classroom integration	9	2.76	0.91	.90	.91	.55
Informal exposure index	6	3.21	0.96	.83	.84	.51

Note. α = Cronbach's alpha; CR = composite reliability; AVE = average variance extracted. All AVEs exceeded .50. Informal exposure was scored on a 1–5 frequency metric.

Procedure and Ethics

Data were gathered between January and February 2026 with approval from the district education authorities and school heads. Teacher surveys were given in staff rooms and/or after-school periods; student surveys were given in science classes as a whole class with a teacher in the room, but not wandering around the classroom. Interviews were conducted for 35–70 minutes, audio-recorded with the consent of the participants and transcribed into the language of the interview prior to bilingual coding. Participation was voluntary. Students might choose not to answer questions about control of household devices. School names are anonymized. The protocol did not consider the disclosure of student usage of AI as evidence of misconduct, and that was expressly written in consent scripts.

Analytic Strategy

The quantitative data analysis conducted was SPSS 29 and lavaan in R. Following the screening for missingness (the level of missingness for the largest construct was 2.4 percent; the item level missingness was estimated using the expectation-maximization approach for cases with less than 10 percent missing data), we estimated a confirmatory factor analysis (CFA) of the teacher measurement model and a structural path model corresponding to the micro–meso core of PASCf. Fit criteria followed conventional ranges (CFI and TLI $\geq$.90, RMSEA $\leq$.06, SRMR $\leq$.08; Hu & Bentler, 1999). The tests performed at the configural level and at the metric level were multi-group tests. Welch tests and factorial ANOVA were used for the group differences. The student opportunity-to-learn scores were modeled by using locality-by-gender ANOVA and school-level clustering checks.

Reflexive thematic analysis (Braun & Clarke, 2021) was used for the qualitative analysis. A 12-transcript subset was independently coded by two researchers, a codebook of 31 first-cycle codes was developed, under PASCf layers of coding, and then the codebook was applied to the entire corpus in NVivo. Theme generation was not codebook pre-determined. Standardized paths were displayed in conjunction with exemplar quotations for the theme. Disconfirming cases (high integration rural teachers, high fear and low use urban teachers) were kept in the narrative.

Results

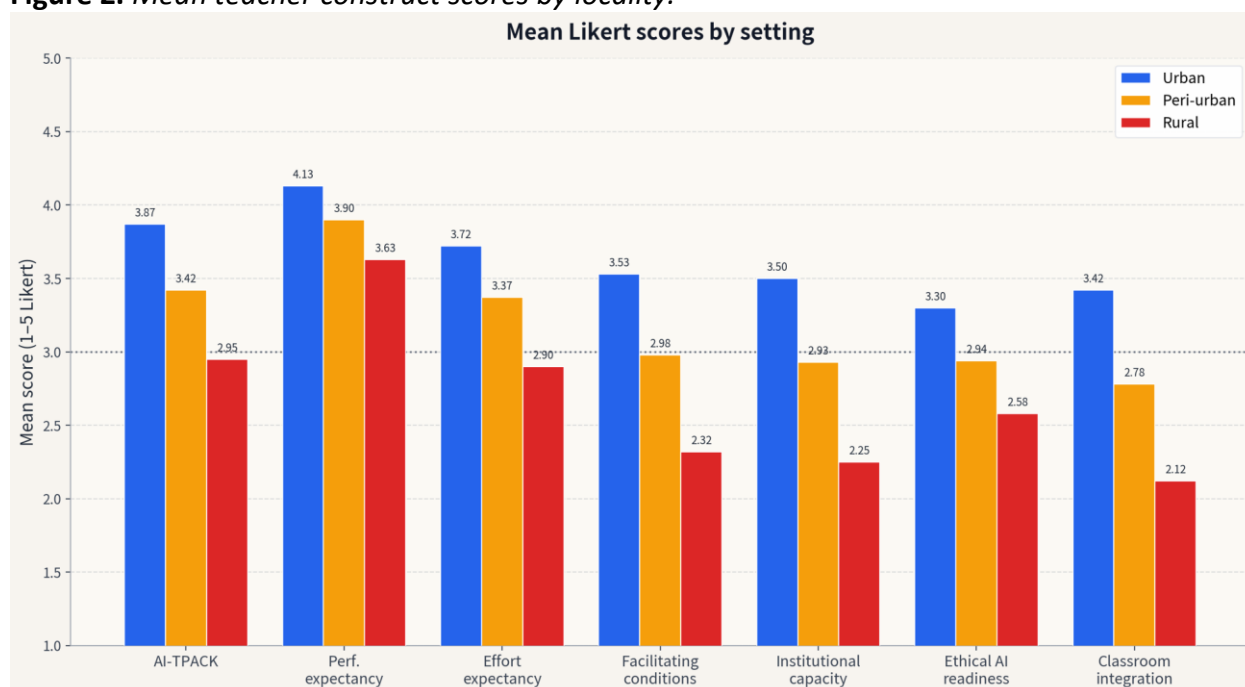
Measurement Model and Descriptive Pattern

CFA of the ten-construct teacher model yielded acceptable fit, $\chi^2(1947) = 3621.18$, CFI = .946, TLI = .931, RMSEA = .051 (90% CI [.048, .054]), SRMR = .047. The standardized loadings on them ranged from .62 to .84. The Fornell–Larcker comparisons and heterotrait–monotrait ratios were less than .85, which supported discriminant validity. All of the reliability coefficients were higher than the regular standards presented in Table 2.

Descriptive indicates describes a characteristic peripheral profile. The highest teacher construct was performance expectancy ($M = 3.86$) with teachers generally agreeing that GenAI would be useful for students to view alternative worked examples, to translate dense English explanations and to create practice items. The results for institutional capacity ($M = 2.86$), facilitating conditions ($M = 2.92$), ethical readiness ($M = 2.93$) and classroom integration ($M = 2.76$) were below the midpoint of the scale. The scores for AI–TPACK were on the middle level ($M = 3.38$); however, the scores of large locality gaps were hidden. The bottom line,

teachers felt that the tools were important, but not convinced, equipped, and/or prepared to use the tools to teach.

Figure 2. Mean teacher construct scores by locality.



Note. Error-free bars display cell means. The dotted reference line marks the scale midpoint (3.00). Rural means fall below the midpoint on facilitating conditions, institutional capacity, and classroom integration.

Table 3. Locality Differences on Focal Teacher Constructs.

Construct	Urban M (SD)	Peri-urban M (SD)	Rural M (SD)	F	η^2
AI-TPACK	3.86 (0.58)	3.41 (0.61)	2.94 (0.66)	98.14***	.29
Performance expectancy	4.12 (0.52)	3.88 (0.57)	3.62 (0.71)	31.07***	.11
Effort expectancy	3.71 (0.64)	3.36 (0.70)	2.89 (0.78)	58.22***	.19
Facilitating conditions	3.54 (0.71)	2.97 (0.76)	2.31 (0.81)	112.40** *	.32
Institutional capacity	3.48 (0.69)	2.92 (0.72)	2.24 (0.74)	127.18** *	.34
Ethical AI readiness	3.29 (0.70)	2.94 (0.71)	2.58 (0.76)	42.55***	.15
Classroom integration	3.41 (0.73)	2.78 (0.74)	2.12 (0.71)	136.92** *	.36

Note. Degrees of freedom for the locality ANOVA were 2, 483. *** $p < .001$. Welch tests confirmed the same rank order. Post hoc Games–Howell tests separated all three locality groups on integration, capacity, and facilitating conditions.

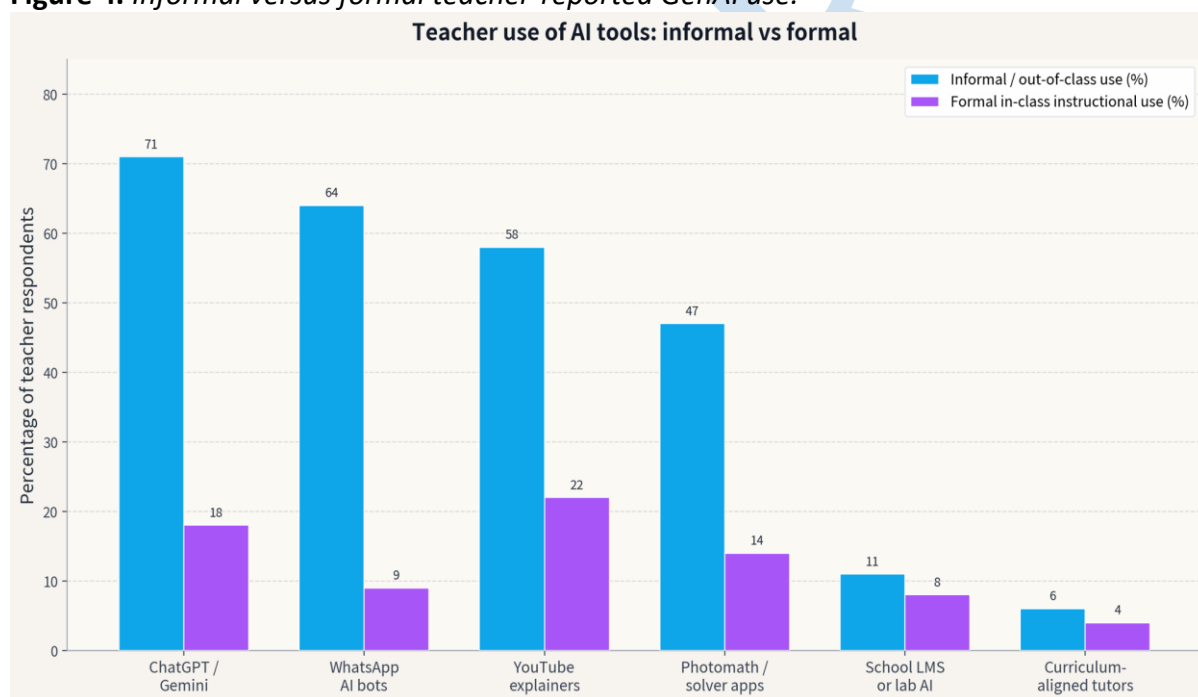
The differences between subjects were less than the differences between localities. The mean scores for the computer-science teachers were higher than the grand means of the other teachers for items corresponding to AI-TPACK ($M = 3.79$) and integration ($M = 3.22$), while the mean scores of the other teachers were close to the grand means, except for lower scores in ethical readiness for the mathematics teachers ($M = 3.22$), showing a higher performance expectancy but lower ethical readiness, which is consistent with the growing popularity of

solver applications. For women teachers there was no difference from men on performance expectancy ($p = .41$) but there was a difference in terms of facilitating conditions ($d = 0.28$), which we discuss again in the student data, and on integration in rural cells. Public–private differences were found to be positive for capacity ($d = 0.46$) and integration ($d = 0.39$), but diminished when locality was covaried, suggesting that a significant part of the public–private difference reflects urban private advantage.

The Informal–Formal Split

Figure 3 shows the median descriptive finding in practice. 71% of teachers reported that they had seen their students use ChatGPT or Gemini for informal purposes, 64% for WhatsApp AI bots, 58% for YouTube explainers, and 47% for solver applications. The percentage for formal use of the same tools in the classroom decreased to 18%, 9%, 22%, and 14% for the four families of tools respectively. It was uncommon to find school LMS or tutors aligned to the curriculum in both of these ecologies. The presence of school LMS or curriculum aligned tutors was uncommon in both ecologies. The personal use of a smartphone was most prevalent in the informal–formal gap, as personal use was relatively common in peri-urban public schools, while official permission and laboratory scheduling were not.

Figure 4. Informal versus formal teacher-reported GenAI use.



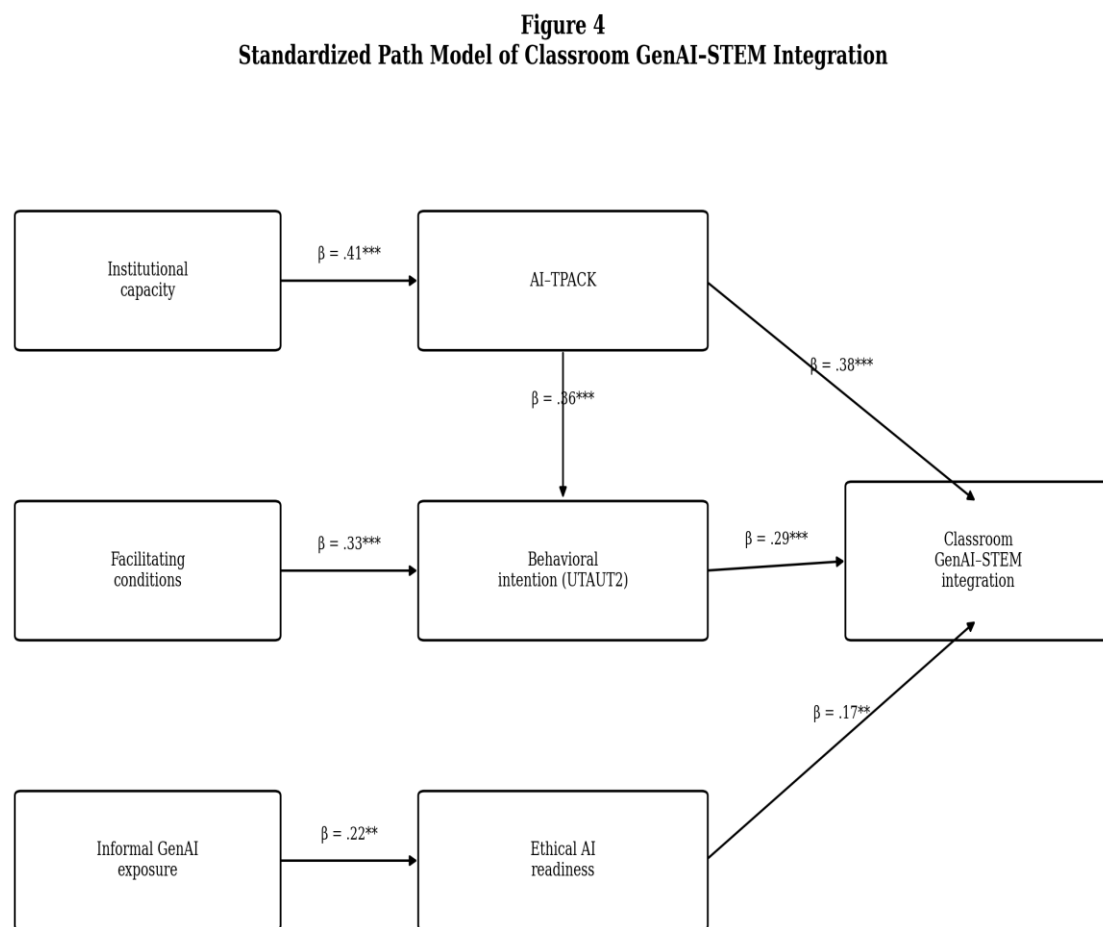
Note. Percentages are teacher self-reports of any use in the previous academic term. Informal use includes personal lesson planning and out-of-class student support; formal use requires in-period instructional design.

Structural Model

The structural model corresponding to H1–H5 fit adequately, CFI = .946, TLI = .931, RMSEA = .051, SRMR = .047. To support H1, institutional capacity was found to be a predictor of AI–TPACK ($\beta = .41$, $p < .001$). Facilitating conditions was positively associated with behavioral intention ($\beta = .33$, $p < .001$), supporting H2. The result for informal exposure was modest ($\beta = .22$, $p < .01$) and the same result was obtained in a supplementary model that substituted the ethical readiness with a short uncritical-use index ($\beta = .39$), lending some support to the caution expressed in H3. AI–TPACK predicted intention ($\beta = .36$, $p < .001$) and integration ($\beta = .38$, $p < .001$), supporting H4. Integration was predicted by both Intention ($\beta = .29$, $p < .001$)

and ethical readiness ($\beta = .17, p < .01$) with AI-TPACK having the largest unique coefficient, supporting H5. R² values for AI-TPACK, intention, ethical readiness and classroom integration were .44, .48, .19 and .57, respectively.

Figure 5. Standardized path model of classroom GenAI-STEM integration.



Note. N = 486 STEM teachers. **p < .01. ***p < .001. Model fit: CFI = .946, TLI = .931, RMSEA = .051, SRMR = .047.

Note. N = 486 STEM teachers. Indirect effect of institutional capacity on integration via AI-TPACK was significant (ab = .16, 95% bias-corrected CI [.11, .21]).

Table 4. Structural Paths Predicting Classroom GenAI-STEM Integration.

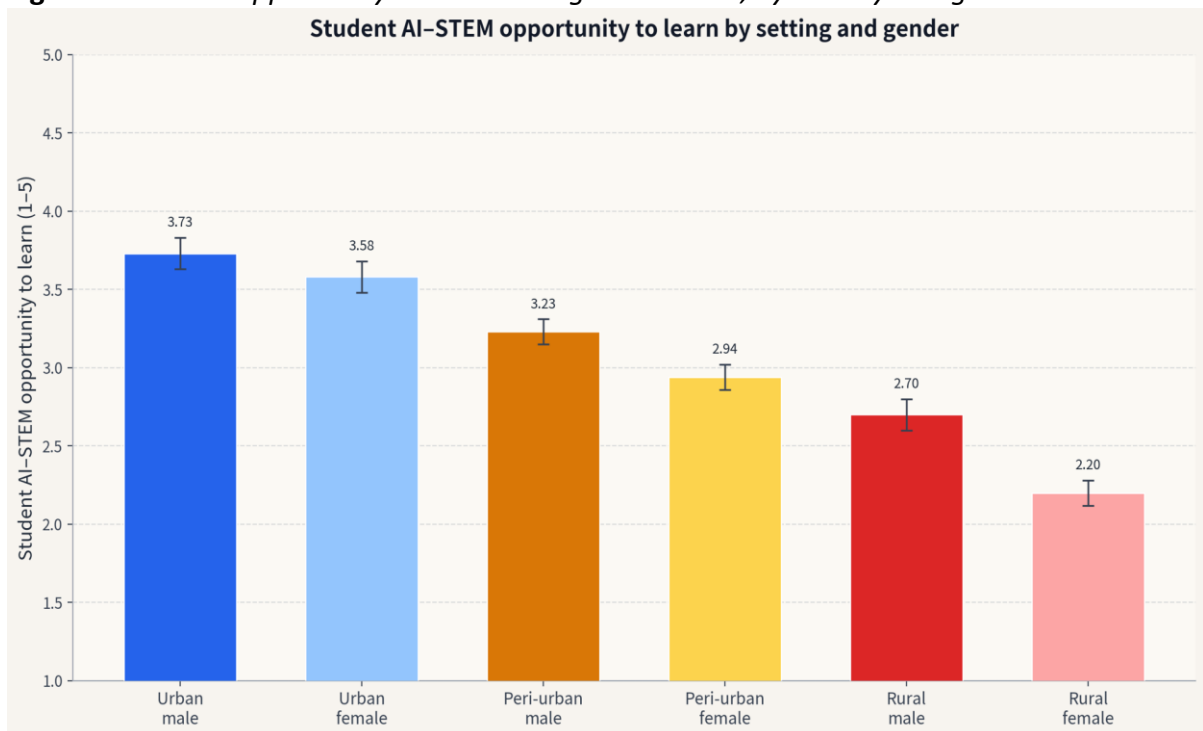
Path	β	SE	Z	p
Institutional capacity → AI-TPACK	.41	.04	10.22	< .001
Facilitating conditions → Intention	.33	.05	6.91	< .001
Informal exposure → Ethical readiness	.22	.05	4.18	< .001
AI-TPACK → Intention	.36	.05	7.44	< .001
AI-TPACK → Integration	.38	.05	7.90	< .001
Intention → Integration	.29	.05	6.02	< .001
Ethical readiness → Integration	.17	.04	3.88	< .001
Capacity → Integration (indirect via AI-TPACK)	.16	.03	5.71	< .001

Note. Standardized coefficients from the lavaan robust maximum-likelihood solution. Control paths from gender and years of experience to integration were small and not statistically significant.

Student Opportunity to Learn

Opportunity to learn (OTL) scores for students reproduced and accentuated the teacher locality gradient. A 3×2 ANOVA (locality by gender) yielded a large locality effect, $F(2, 806) = 94.66$, $p < .001$, $\eta^2 = .19$; a gender effect, $F(1, 806) = 28.41$, $p < .001$, $\eta^2 = .03$; and a locality-by-gender interaction, $F(2, 806) = 6.18$, $p = .002$, $\eta^2 = .02$. The interaction effect was focused in rural cells: the mean OTL of the rural young women was 2.19, which is quite low compared to the rural young men mean (2.68) and extremely low compared to the urban young women mean (3.58). The most pronounced gender differences were found at the item level with items such as “I can use a phone or computer to do science work without having to ask several people,” “I can read the answers from the tool in a language I understand,” and “My teacher tells me how to determine if the software is incorrect.”

Figure 6. Student opportunity to learn with generative AI, by locality and gender.



Note. Error bars represent standard errors. The interaction is driven by the rural young-women cell.

There was a positive and significant correlation between student STEM-engagement scores and OTL ($r = .46$, $p < .001$). It was not a replacement for achievement data, which was not collected (we did not take board marks), but rather a proxy for opportunity to be patterned onto motivational self-reports. 61% of the urban students, 49% of peri-urban students, and 33% of the rural students indicated that they used ChatGPT or similar tools to help with science or maths homework in the past month, and 24% of the rural young women indicated the same. Whereas, informal student use, similar to informal teacher use, exceeded any narration of school designed lessons around AI.

Interview Findings

Five themes organized the interview corpus. They do not merely illustrate the survey; they specify the conversion factors that survey residuals pointed toward.

Theme 1: Infrastructure as a Pedagogical Actor

The headteachers and rural teachers did not talk about the lack of electricity and devices as a backdrop, but as a canceller of lesson plans. A physics teacher from rural Muzaffargarh stated that he can prepare a prompt at night using his brother's cell phone, and, in third period, the

light turns off, they then have to write on the board and think that he was showcasing yesterday. Urban private coordinators used a different infrastructural language: bandwidth caps, shared laptop trolleys and filters that block the same sites that teachers recommend. In PASCf terms, infrastructure is a conversion factor that determines if AI-TPACK can be implemented within the same week that it is taken up in a questionnaire.

Theme 2: The Shadow Syllabus of GenAI

Students spoke of a shadow syllabus, a group on WhatsApp that they would use before the morning assembly to discuss a possible solution; screenshots of the Algebra problems on Photo math to solve and copy and paste as "explanation" into Urdu; and rewriting the explanation in English-to-Urdu format from the paragraphs of the Urdu textbook. An 11th grade chemistry student in Bahawalpur said, "Sir, we do not know ChatGPT. We know ChatGPT, we don't know how to ask it about the mechanism, we ask it for the long question." Those teachers who faced it first-hand when they attempted to use GenAI to create worksheets were aware of the pattern and still did not have a recognized answer. In interview terms the informal-formal split in Figure 5 is a split between an economy of homework run by the students as opposed to an economy of examination run by the teacher.

Theme 3: Language as a Gate, not a Medium

All the reports of tool use in South Punjab were framed within the language ecology of Saraiki and Punjabi, as languages in which the discourse occurred, Urdu, the language of the school, and English, the high-status language of STEM. Students highly appreciated models which could "make the English paragraph into simple Urdu." They had not only a contempt for outputs that continued to be in an idiomatic manner with citations but also a sense of distrust. While teachers relied on the convenience of translation to prepare for lessons, they also had the concerns that this convenience will further delay their students' scientific reading in English. The concept of language is therefore not considered as a demographic covariate, but rather as a conversion factor; in PASCf, increasing one's language capability increases one's capability when reading an explanation in one language, and vice versa, decreasing one's language capability decreases one's capability when reading an explanation in one language.

Theme 4: Assessment Gravity

Board examinations had the largest effect on practice, surpassing any of the belief items. Rahim Yar Khan's mathematics teacher just stated it: "Why should I teach the child to argue with the machine, if the board puts the same long question?" Headteachers were worried about complaints from parents over the use of AI, as they perceived it as a substitute for "real study." Ethics was not a discourse largely about data-privacy, hence this intelligibility of the ethical readiness scores being only mid-low. It was an integrity and marks discourse. Teachers sought a provincial directive that would clarify whether AI-generated homework is cheating or scaffolding, or both, and a set of tasks to assess that students could not complete independently of AI. If there is no redesign, high-performance expectancy will result in private teacher use and public classroom avoidance.

Theme 5: Gendered Conversion and Quiet Capability

In some rural and some peri-urban schools young girls spoke of the phone as an object of the house and not as a tool for learning. The entrance was provided by brothers, fathers or by having one of the family's cell phones available after dinner. A 10th grader from Layyah replied: "I use it when my brother goes to cricket, twenty minutes, I cannot watch the long video, and I ask for the short answer. Similar domestic negotiations were reported by female teachers when they were preparing lessons on their own personal devices. These accounts help to make sense of the fact that even if teacher performance expectancy is equal across

cell, the OTL gender gap is still greater in the rural cells. It's time, privacy and permissions that determine capability, and NOT the strength of the signal.

Disconfirming Cases

They implemented weekly “check the machine” routines using just one smartphone with a TV in two rural computer-science teachers from Khanewal and DG Khan. They took model error as the object of the lesson. Although their facilitating-conditions scores were low, their integration scores were high. The cases matter theoretically: In case of pedagogical ideas as interrogation, AI–TPACK can partially replace infrastructure. They do not misrepresent the capacity path, they make it clear that professional knowledge can extend a limited space, and that's a reason for professional learning and not an excuse to leave laboratories in the dark.

Discussion

This study was an attempt to find out what is GenAI doing on the periphery of Pakistani education system and the global AI–STEM research map in secondary STEM education? The answer is that, at present, AI is not a formal pedagogy and is not much of an informal practice in education. The longer answer is that in high resource systems, the presence of an equity problem is not something that can be observed with the use of the tool of tool effectiveness studies.

Relocating AI–STEM Theory

The core empirical argument of PASCf is that enthusiasm is not the primary pathway for institutionalizing classroom integration, but rather is the pathway that is taken at the conclusion of the process of institutionalization. Performance expectancy was high in all localities, with integration not. This dissociation is hard to take in for a single TAM or UTAUT story, and is predicted by a capability story. A tool can have a lot of worth for a teacher, but not be a functioning. The increased values of capacity, facilitating conditions and integration (η^2 from .32 to .36) are large and significantly larger than typical increments associated with the acceptance-models and consistent with the qualitative work on Pakistan that considers infrastructure as a gatekeeper (Hussain, 2025). They are also consistent with bibliometric concerns about models exported from the Global North, which might be expected to perform well if they are brought to a country where conditions are conducive rather than if they are brought to a country in which conditions are adverse (Yang et al., 2025).

The small link between informal exposure to ethics and ethical readiness, versus a larger link between the same exposure and uncritical use, is the most salient warning against generative-AI scholarship in the study. Previous studies on the topic of GenAI have been based on perceptual and short-term assessment. Through our data we get the idea of a mechanism: informal first contact trains fluency without training judgment. Using ChatGPT as a homework engine doesn't guarantee students and teachers that they'll get used to asking themselves about the origin of a mechanism, the source of a bogus citation, or the need to put a warrant in a proof. If informal fluency is assumed as capability, then human-flourishing agendas that call for broader than cognitive aims (OECD, 2024) will stall.

Equity Is Not a Residual

The young-women cell of the rural enjoys the status of the study's moral core. The documentation for ADB already provides the reasons for programs, such as gender and rural completion in the south. This study adds that GenAI, unguided by household allocation and informal markets, will follow along those same channels. A conversion factor includes device ownership, traversal by the night time, and traversal in English–Urdu–Saraiki. An in-every-classroom approach to AI that provides accounts or content libraries without reengineering

those factors will create urban male fluency and rural female exclusion, which will then be interpreted as a motivation problem.

Gender and rurality go together and so does language. High impact AI–STEM work doesn't really consider interface language as a scientific variable. It is a determining factor in South Punjab. It is not an accommodation to be able to switch from English scientific register to accessible Urdu; it's a condition of the enquiry. Meanwhile, translation without criticism can stifle students on the wrong side of scientific English, which is the language of many questions on the Board, university lectures and futures in science and technology all over the world. However, bilingual interrogation must be taught, and the choice of one language is not to be considered a solution.

The Informal–Formal Split as a Research Object

We believe the informal–formal dichotomy should be brought into the fold of Global South studies in AI–STEM research like the shadow education in comparative education research, and have a place in the mainstream of such research. Classroom integration is under-represented as it is only measured by the classroom. The measurement of pedagogy is an overcount, since it only captures self-reported use of ChatGPT. The dual representation of Figure 5 along with Themes 2 and 4 represents a system where the homework contract has already been rewritten by the students and the assessment contract has not been rewritten by schools. Private competent teachers and public non-changing lessons will be the result of AI–TPACK workshops until those contracts are renegotiated.

Relation to Prior South Punjab and Pakistani Work

Punjab-focused research at university has found teachers to be open to personalized AI learning and positive association between the use of digital tools and perceived skill development by students. The results of those studies are consistent with our high performance-expectancy estimates and high informal-use estimates. They are incomplete as a tool for engaging in secondary STEM policy, lacking, for example, the gravity of board exams, the unreliability of labs, and more constraining gendered household device regimes at the secondary level. The results of the single-district rural teacher survey results that identify TPACK as a strong predictor of enactment align with our path to support AI–TPACK; the current study expands this evidence across three divisions, introduces an additional strand (student OTL), and provides a formal–informal mechanism that is not visible to a teacher-only survey.

Implications

Implications for Theory

For future AI–STEM models for conversion factors should be explicitly called out as first class objects. The introduction of a “context” paragraph into a UTAUT2 paper is not identical to the theoretical analysis of institutional capacity, language traversal and gendered access, and testing for such moderation between the intention–use path. There is one specification for PASC, others can include self-determination theory or epistemic cognition. What is required is that the periphery not be viewed as a bunch of noise around a universal acceptance engine.

Implications for Practice and Professional Learning

Tool Tours should not be the starting point in professional development. It should start with the interrogation process that is applicable to one phone and one display—the disconfirming cases in this study already created such processes. The teachers will require some rehearsed moves: “Find the invented step”, “Give me the same mechanism in Urdu and tell me what disappeared”, “What experiment would falsify this generated claim?” The moves build AI–TPACK in the only environment that most schools in the South Punjab have at this time. In the

case of parallel tracks, assessment design is a critical issue because the teachers do not want students to be encouraged to do research on what they are asked to do, and the board may use the tools they are given as a means of cheating.

Implications for Policy

The ADB STEM secondary program and the provincial digital-skills programs provide a unique policy space. Sequence matters. Capability preconditions are requirements for the ability to build a successful AI curriculum, not a part of an AI curriculum. The need to differentiate between disclosure-based scaffolding and undisclosed product generation must be emphasized when formulating school-level integrity policies. With load shedding, content as well as tool selection ought to take into account scientific interfaces in both languages and offline-tolerant applications. Equity should be a goal and not a preamble to district professional-learning systems that should be delivered to EST and SST science cadres in rural girls' schools first. Lastly, informal trainers and e-Rozgar type centers should be considered as an integral part of the STEM ecology, rather than an alternative world of youth employment. Students are already transiting that border every night.

Limitations and Future Research

The design is cross sectional, so there is no cause and effect. Self-report was higher than observed practice, but interviews helped to resolve this issue. Standardized achievement and/or board-examination scores were not collected, so the distal learning claims are inferential. There was an underrepresentation of private unaided and madrasa attached science sections. Urdu instrumentation has emphasized some nuances of Saraiki at the cost of others, which might be attributed to the same. The student sample does not cover out-of-school adolescents, a non-trivial population in districts which have low secondary completion rates. Classroom observation and artifact analysis, ongoing follow-ups over an examination cycle, experimental trials of interrogation routines on "thin" infrastructure, and actual model log measurement (where ethically feasible) should be undertaken in future research. A comparative study with peripheral STEM systems, other interior STEM systems of Sindh and southern KP will help to understand if PASCf is travelable.

Conclusion

There are no lacks of tools, reviews, or optimism in AI-STEM research. It is devoid of theories and evidence developed within systems where a generator can be in the pocket of a student and not in the science period and in which the absence is patterned by rurality, gender, language, and examination gravity. South Punjab displays that condition. Teach educators in this study feel generative AI has a role to play; institutions don't design to use it; students, particularly young women from rural areas, inherit the least opportunity and the most shadow use. The Peripheral AI-STEM Capability Framework sees that pattern as the phenomenon to be explained, rather than more of a local nuisance on the road to the adoption curve. Pedagogical judgment, bilingual interrogation, integrity rules, and girls' actual time-on-device will not be installed with the forthcoming provincial and multi-lateral STEM investments, or will South Punjab get left behind by the generative-AI wave? It will be stratified in accordance with it.

References

- Asian Development Bank. (2025). Responsive, Ready, and Resilient Science, Technology, Engineering, and Mathematics Secondary Education in Punjab Program: Report and recommendation of the President (Project No. 54298-001).
- Braun, V., & Clarke, V. (2021). Thematic analysis: A practical guide. SAGE.
- Bureau of Statistics Punjab. (2024). Punjab development statistics.

- Bureau of Statistics Punjab. (2025). Punjab development statistics 2025.
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Chiu, T. K. F., & Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, 12(14), 5568.
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
- Express Tribune. (2025, May 4). Digital skills gain momentum in South Punjab.
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Iniesto, F., Sharples, M., & Scanlon, E. (2023). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hussain, M. (2025). The rural–urban divide in digital readiness for AI integration: A qualitative study of Pakistani teachers' perspectives. *Journal of Media Horizons*.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*.
- Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824–2838.
- 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.
- Organisation for Economic Co-operation and Development. (2024). *Education for human flourishing in an age of generative AI*.
- Programme Monitoring and Implementation Unit. (2024). *Annual school census 2023–24. Punjab Education Sector Reform Programme*.
- Programme Monitoring and Implementation Unit. (2025). *Annual school census 2024–25. Punjab Education Sector Reform Programme*.
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury.
- Selwyn, N. (2024). On the limits of artificial intelligence in education. *Nordic Journal of Studies in Educational Policy*.
- Sen, A. (1999). *Development as freedom*. Oxford University Press.
- Tikly, L. (2011). Towards a framework for researching the quality of education in low-income countries. *Comparative Education*, 47(1), 1–23.
- UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization.
- Unterhalter, E. (2009). What is equity in education? Reflections from the capability approach. *Studies in Philosophy and Education*, 28, 415–424.

- van Dijk, J. (2020). The digital divide. Polity.
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. <https://doi.org/10.2307/41410412>
- Walker, M., & Unterhalter, E. (Eds.). (2007). Amartya Sen's capability approach and social justice in education. Palgrave Macmillan.
- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 45(3), 223–235.
- Yang, Y., Sun, W., Sun, D., & Salas-Pilco, S. Z. (2025). Navigating the AI-enhanced STEM education landscape: A decade of insights, trends, and opportunities. *Research in Science & Technological Education*, 43(3), 693–717. <https://doi.org/10.1080/02635143.2024.2370764>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, Article 39.
- Zhai, X., & colleagues. (2025). Artificial intelligence in science education research: Current states and challenges. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-025-10239-8>