



JOURNAL ON EFFICIENCY
AND RESPONSIBILITY IN
EDUCATION AND SCIENCE

The ERIES Journal is being managed by an international editorial board as a regular scientific journal. A rigorous process of papers' reviews (double-blind peer review) is fully supported by a web-based submission system. The journal is published electronically four times a year, on March 31, June 30, September 30 and December 31 of the current year.

The journal is indexed in:

- BASE - Bielefeld Academic Search Engine
- Directory of Open Access Journals
- EBSCO database
- Emerging Sources Citation Index - Web of Science™ Core Collection
- ERIC - Education Resources Information Center
- ERIH PLUS
- Google Scholar
- Open Academic Journals Index
- ResearchBib
- SCOPUS
- The list of reviewed periodicals in the Czech Republic



www.eriesjournal.com

Scientific journal of the Czech University of Life Sciences Prague JOURNAL ON EFFICIENCY AND RESPONSIBILITY IN EDUCATION AND SCIENCE, distributed by the Faculty of Economics and Management. Published quarterly. Executive editors: Ing. Martin Flégl, Ph.D., Ing. Jiří Fejfar, Ph.D., PhDr. Michaela Cocca and Armando Cocca, PhD. Editorial Office: ERIES Journal, Czech University of Life Sciences Prague, CZ 165 21 Prague 6 - Suchbát, Czech Republic, email: editor@eriesjournal.com, tel: +420 224 382 355.

© Czech University of Life Sciences Prague, Czech Republic, 2026

volume 19
issue 3

2026

An international peer-reviewed journal published by
Faculty of Economics and Management
Czech University of Life Sciences Prague

editor@eriesjournal.com
www.eriesjournal.com
Online ISSN: 1803-1617
Printed ISSN: 2336-2375

EDITORIAL BOARD

EDITOR-IN-CHIEF

doc. Ing. Martin Pelikán, Ph.D.
Czech University of Life Sciences Prague, Czech Republic

EXECUTIVE EDITORS

Ing. Martin Flégl, Ph.D.
Tecnológico de Monterrey, Mexico
Ing. Jiří Fejfar, Ph.D.
Czech University of Life Sciences Prague, Czech Republic

EDITORIAL BOARD MEMBERS

Peter M. Bednar, PhD.
University of Portsmouth, United Kingdom
prof. RNDr. Helena Brožová, CSc.
Czech University of Life Sciences Prague, Czech Republic
Prof. Dr. Irem Comoglu
Dokuz Eylul University, Turkey
Assoc. prof. Anna Croon Fors
Umea University, Sweden
Dr. Martin Daumiller
University of Augsburg, Germany
doc. Ing. Peter Fandel, CSc.
Slovak University of Agriculture in Nitra, Slovak Republic
prof. Ing. Jakub Fischer, Ph.D.
Univesity of Economics Prague, Czech Republic
prof. Ing. Jana Hančlová, CSc.
Technical University of Ostrava, Czech Republic
Joan Harvey, PhD.
Newcastle University, United Kindgdom
doc. Ing. Milan Houška, Ph.D.
Czech University of Life Sciences Prague, Czech Republic
prof. PhDr. Tomáš Janík, Ph.D., M.Ed.
Masaryk University, Czech Republic
doc. Ing. Pavel Klouček, Ph.D.
Czech University of Life Sciences Prague, Czech Republic
prof. RNDr. Jindřich Klůfa, CSc.
University of Economics, Prague, Czech Republic
doc. PhDr. Luděk Kolman, CSc.
Czech University of Life Sciences Prague, Czech Republic
doc. Ing. Igor Krejčí, Ph.D.
Czech University of Life Sciences Prague, Czech Republic
PhDr. Kristýna Krejčová, Ph.D.
Charles University, Czech Republic
prof. PhDr. Michal Lošťák, Ph.D.
Czech University of Life Sciences Prague, Czech Republic
Ricardo Lozano, PhD.
Texas A&M International University, USA

TECHNICAL EDITORS

Mgr. Dominik Bláha*
Ing. Jiří Fejfar, Ph.D.*
Ing. Michal Hruška, Ph.D.*
* Czech University of Life Sciences Prague, Czech Republic

EDITORIAL OFFICE

ERIES Journal, Czech University of Life Sciences Prague (CZU), CZ 165 00 Prague 6 - Suchdol, Czech Republic
email: editor@eriesjournal.com * tel: +420 224 382 355
Registration number: MK ČR E 21414 © Czech University of Life Sciences Prague, Czech Republic, 2026

PhDr. Michaela Cocca
Czech University of Life Sciences Prague, Czech Republic
Armando Cocca, PhD.
Czech University of Life Sciences Prague, Czech Republic

Univ. prof. i. R. Dipl. Ing. Dr. Mikuláš Luptáčík
Matej Bel University, Slovakia
Dott. Ing. Tania Di Mascio, PhD
University of L'Aquila, Italy
doc. Ing. Stanislava Mildeová CSc.
University of Finance and Administration, Czech Republic
prof. RNDr. Eva Milková, Ph.D.
University of Hradec Králové, Czech Republic
prof. Ing. Zdeněk Molnár, CSc.
Czech Technical University in Prague, Czech Republic
Ing. Karel Němejč, Ph.D.
Czech University of Life Sciences Prague, Czech Republic
Ing. Jaromír Novák, Ph.D.
University of Economics in Bratislava, Slovakia
prof. RNDr. Jarmila Novotná, CSc.
Charles University, Czech Republic
prof. PhDr. Libor Pavera, CSc.
University of Economics, Prague, Czech Republic
David Pérez-Jorge, PhD
Universidad de La Laguna, Spain
Moufida Sadok, PhD.
University of Portsmouth, United Kingdom
prof. PhDr. RNDr. Antonín Slabý, CSc.
University of Hradec Králové, Czech Republic
PhDr. Ing. Petr Soukup, Ph.D.
Charles University, Czech Republic
doc. Ing. Tomáš Šubrt, Ph.D.
Czech Technical University in Prague, Czech Republic
prof. Ing. Milan Turčáni, CSc.
Constantine the Philosopher University in Nitra, Slovakia
Dr. Özgehan Uştuk
The Hong Kong Polytechnic University
Christine Welch, PhD.
University of Portsmouth Business School, United Kingdom
doc. Ing. Roman Zuzák, Ph.D.
University of the Applied Psychology, Terezín, Czech Republic

INSTRUCTIONS FOR AUTHORS

The Journal on Efficiency and Responsibility in Education and Science publishes papers of the following categories: full research papers, short communications, review studies and book reviews (on invitation only).

- FULL RESEARCH PAPERS
- SHORT COMMUNICATION
- REVIEW STUDY

Papers are published in English. A paper may comprise an empirical study using an acceptable research strategy, such as survey, case study, experiment, archival analysis, etc. It may contain a theoretical study aimed at advancing current theory or adapting theory to local conditions or it may arise from theoretical studies aimed at reviewing and/or synthesizing existing theory. Concepts and underlying principles should be emphasized, with enough background information to orient any reader who is not a specialist in the particular subject area.

Submission checklist

The paper. The paper is carefully formatted according to the template of the journal (see bellow). Special attention is paid to the exact application of the Harvard referencing convention to both continuous citations and list of references. If an electronic source has the DOI number assigned, also it will be provided in the list of references. Manuscripts are submitted via the editorial system in the DOC.

Research highlights. The core results, findings or conclusions of the paper are emphasized in 2-4 bullet points (max. 150 characters per bullet point including spaces). The highlights are submitted as a text into the submission form in the editorial system.

Copyright form. The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. The manuscript submitted is accompanied by the copyright form signed by the corresponding author who declares the agreement of all authors with the conditions in the Form. The Form is submitted into the editorial system in the PDF format.

Suggested reviewers. It is required to suggest two experts appropriate to evaluation of the paper. The experts should be out of the affiliation of the author(s), Czech University of Life Sciences Prague, and also both experts should be from different affiliations. The reviewers are submitted into the text fields in the submission form of the editorial system.

Preparation of the manuscript (technical notes)

Authors are responsible for applying all requirements that are specified in the journal's paper template in individual sections. Especially, the paper must provide a short review of current state in the area of the paper's aim in Introduction. The paper should refer significant sources, particularly scientific journals or monographs. Papers must be closely scrutinized for typographical and grammatical errors. If English is not author's first language then the paper should be proof-read by a native English-speaking person, preferably one with experience of writing for academic use. Spelling should follow the Oxford English Dictionary. Tables, graphs and illustrations should be drawn using a suitable drawing package. Colour may be used. Place all diagrams and tables where you wish them to appear in the paper. Ensure your diagrams fit within the margins and are resizable without distortion.

Review procedure

Following Editorial recommendation, papers are submitted to a double-blind peer review process before publication. Commentary by reviewers will be summarized and sent by email to authors, who can choose to revise their papers in line with these remarks. Re-submitted papers should be accompanied by the description of the changes and other responses to reviewers' comments (see above), so that the desk-editor can easily see where changes have been made.

Copyright

Authors are fully responsible for the paper's originality and for correctness of its subject-matter, language and formal attributes. Author's statement should be enclosed declaring that the paper has not been published anywhere else.

The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. Before publication, authors will be asked to complete a copyright release, giving the publisher permission to publish the paper in a specific issue of this Journal. Overall copyright ownership of the paper, however, remains with the author/s. It is the authors' responsibility to obtain written permission to quote material that has appeared in another publication.

In this third issue of 2026 (Vol. 19, No. 3), we are pleased to present eight articles representing diverse regions, including Uzbekistan, Indonesia, the Philippines, Colombia, Ghana, Zimbabwe, and Mexico. Collectively, these contributions examine how education systems and teacher preparation respond to challenges related to ethics, technology, equity, and climate resilience in rapidly evolving contexts.

The articles address a broad range of topics: bioethical-communicative judgments across medical training years in Uzbekistan; students' attitudes toward generative AI as a pathway to self-regulated learning in Indonesia; the parallel development of assessment and feedback literacy among preservice teachers in the Philippines; the relationship between socioeconomic vulnerability and digital connectivity in Colombian schools; vocational students' acceptance of learning management systems in Indonesia; postgraduate students' willingness to teach ethnomathematics in Ghana; the relevance of A-Level points for university admissions in Zimbabwe; and Social Studies teachers' assessment practices during heat-index class suspensions in the Philippines. Despite their varied contexts, these studies converge around several central themes: the need to align curricula and admissions practices with fairness and responsibility, to foster confidence and higher-order thinking, to leverage digital tools for equitable access, and to ensure that educational innovation leads to sustainable student outcomes. Taken together, the articles highlight the continuing transformation of education toward efficiency, inclusivity, and adaptability in the 21st century.

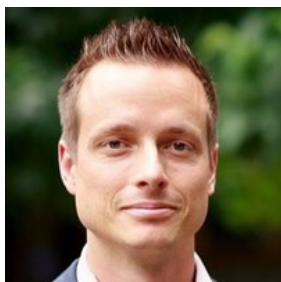
The first article, "Bioethical Communicative Judgments Across Training Years: Cross-Year Evidence from Uzbek Medical Education", by Gulnoza Salimova from Abu Ali ibn Sino Bukhara State Medical Institute, Uzbekistan, examines how medical students' ethical and communicative judgments evolve across different stages of training. Drawing on data from 100 undergraduates across six year groups, the study uses a scenario-based Situational Judgment Test to assess domains such as autonomy, consent, comprehension, emotional responsiveness, prognosis disclosure, and hierarchy. The findings reveal domain-specific and bidirectional variation rather than uniform developmental trends: Year 1 students more strongly endorsed autonomy and challenges to hierarchy, whereas students in later years placed greater emphasis on comprehension

and emotional responsiveness. These results underscore the interplay among cultural norms, family involvement, and professional expectations, while also identifying domains that warrant deeper longitudinal and qualitative investigation in Uzbek medical education.

The second article, "Students' Attitudes Toward Generative AI as a Mediating Pathway to Perceived Generative AI Support for Self-Regulated Learning in Higher Education", by Juli Sardi, Darmansyah, Oriza Candra, Nuzul Hidayat, and Devi Faizah Yuliana, examines how students' perceptions of generative AI shape its role in supporting self-regulated learning. Based on survey data from 264 undergraduates in Indonesia and analysed using structural equation modelling, the study tests relationships among perceived ease of use, perceived usefulness, attitudes toward AI, and perceived AI support.

Both ease of use and usefulness were positively associated with attitudes, which in turn mediated their relationship with perceived support. While usability showed a stronger direct effect, usefulness had a stronger indirect effect through attitudes. The authors emphasize that favourable attitudes involve not only willingness to use AI, but also the responsible evaluation of its outputs, thereby extending technology acceptance models into a learning-oriented framework.

The third article, "The Assessment-Feedback Literacy Nexus: a Longitudinal Mixed-Methods Study of Responsible Assessment Practice in Preservice Teacher Education", by Ronald Quileste from Xavier University-Ateneo de Cagayan, Philippines, investigates how these competencies develop together in teacher preparation. Using a two-wave, multi-cohort mixed-methods design with 250 preservice teachers, the study compares performance in Assessment of Learning 1 and 2 courses. Quantitative findings showed higher scores in both domains during the second phase, with strong predictive continuity from first-phase performance. Qualitative reanalysis revealed a shift from compliance-oriented views toward rubric use, feedback dialogue, fairness, and professional responsibility. These insights informed the proposed Assessment-Feedback Literacy Nexus (AFLN), which positions assessment literacy as the structural dimension of responsible practice and feedback literacy as the dialogic dimension through which evidence becomes meaningful. The study emphasizes that responsible assessment requires both defensible design and meaningful feedback engagement.



The fourth article, “Institutional Responsibility and Connectivity Efficiency in Public Schools During the COVID-19 Pandemic: a Comparative Analysis of the Digital Divide in Tunja, Colombia (2020–2021)”, by Alejandra Uscategui Gómez, Byron David Igua Muñoz, Juan Carlos Campos Medina, Joan Sebastian Igua Muñoz, and Jonathan Salomón Igua Muñoz, analyses the relationship between inequality and connectivity in Tunja, Colombia. Using census data from 14 public schools in 2020 and 2021, the study found that average connectivity improved by 24 percentage points, yet the most socioeconomically vulnerable schools—especially rural and peri-urban institutions—remained considerably disadvantaged. Despite government distribution of laptops and SIM cards, resource provision did not consistently translate into improved connectivity, underscoring the limits of device-focused interventions. The authors argue that digitally inclusive policies must go beyond infrastructure and address territorial disparities and pedagogical needs, ensuring that digital resources are usable, supported, and integrated into meaningful educational practices.

The fifth article, “Understanding LMS Acceptance in Vocational Secondary Education: Extending the Technology Acceptance Model Using SEM-GSCA”, by Lifa Farida Panduwinata, Prisilia Joyceline Atmojo, and Febrika Yogie Hermanto from Universitas Negeri Surabaya, Indonesia, examines how vocational students adopt learning management systems (LMS). Based on a survey of 252 students, the study extends the Technology Acceptance Model by incorporating external factors such as self-efficacy, social norms, enjoyment, accessibility, and playfulness. The results show that accessibility was the only external factor associated with ease of use, while perceived usefulness was the strongest predictor of attitudes, intention, and actual LMS use. Acceptance was driven more by functional utility and accessibility than by motivational influences. The authors conclude that sustainable LMS adoption requires prioritising usability, accessibility, and pedagogical relevance, with teachers integrating interactive and task-based approaches to strengthen perceived usefulness.

The sixth article, “From Knowledge to Instruction: Examining the Roles of Usefulness and Teaching Self-Efficacy on Postgraduate Students’ Willingness to Teach Ethnomathematics”, by Ebenezer Kwesi Lotey and Ebenezer Bonyah from Ghana’s University of Skills Training and Entrepreneurial Development, explores how postgraduate students translate ethnomathematics knowledge into teaching practice. Using survey data from 159

students and interviews with eight participants, the study found that pedagogical knowledge, perceived usefulness, and teaching self-efficacy strongly predicted willingness to teach, whereas content knowledge showed a small negative direct effect but a positive overall association through indirect pathways. Interviews revealed enthusiasm for the use of cultural artefacts in mathematics instruction, tempered by curricular and assessment constraints. The authors recommend integrating lesson design, guided practice, and microteaching into postgraduate programmes to strengthen culturally responsive mathematics education.

The seventh article, “Beyond A-Level Points: Evaluating Their Relevance to Responsible and Efficient University Admissions in the Digital Era”, by Desmond Mwembe, Brighton Mpofo, Nicholas J. Sibanda, and Aaron Nyoni from Zimbabwe’s National University of Science and Technology, assesses the role of A-Level points in predicting university outcomes. Analysing records from 1,512 graduates across five public universities, the study found that although A-Level points were associated with faculty placement, they did not directly predict final degree classification once faculty was considered. Gender was linked to faculty choice but not to entry scores or outcomes. Faculty differences were evident in both attainment and degree classifications, reflecting broader structural factors. The authors argue that A-Level points should not be used as a universal predictor of success and that admissions criteria should be discipline-specific, evidence-based, and sensitive to socioeconomic context.

The eighth article, “Assessment Practices and Challenges of Social Studies: Teachers During Heat-Index Class Suspensions”, by Czar Justine D. Navalta from Nueva Vizcaya State University, Philippines, examines how teachers adapted assessment during climate-related school disruptions. Using surveys and interviews with 129 Junior High School teachers, the study found that reflective and compassionate assessment was practised most frequently, followed by performance tasks and written work, with differences across age and years of service. Qualitative findings highlighted submission delays, retrieval difficulties, concerns about authenticity, and increased workload. Teachers reported extending deadlines, adjusting tasks, and strengthening home-school communication to sustain continuity. The study underscores the need for context-sensitive professional development and climate-resilient assessment planning, recommending targeted support in emergency assessment design, authentic assessment, and mentoring.

We extend our sincere gratitude to all authors who contributed their research to this issue of the ERIES Journal and to our reviewers for their dedicated efforts in ensuring the quality of the published work. Their commitment and expertise are essential to maintaining the journal's standards of

excellence. We hope that readers will find this third issue of the year both insightful and inspiring. To stay connected with the latest updates, including our most cited articles, upcoming events, and calls for special issues, we invite you to follow the ERIES Journal on LinkedIn.

Sincerely

A handwritten signature in blue ink, appearing to read 'M. Flégl', with a stylized, cursive script.

Martin Flégl

Executive Editor

ERIES Journal

www.eriesjournal.com

www.linkedin.com/company/eriesjournal/

www.erie.pef.czu.cz

Bioethical-Communicative Judgments Across Training Years: Cross-Year Evidence from Uzbek Medical Education Gulnoza Salimova	184
Students' Attitudes Toward Generative AI as a Mediating Pathway to Perceived Generative AI Support for Self-Regulated Learning in Higher Education Juli Sardi, Darmansyah, Oriza Candra, Nuzul Hidayat, Devi Faizah Yuliana	196
The Assessment–Feedback Literacy Nexus: a Longitudinal Mixed-Methods Study of Responsible Assessment Practice in Preservice Teacher Education Ronald Quileste	210
Institutional Responsibility and Connectivity Efficiency in Public Schools During the COVID-19 Pandemic: a Comparative Analysis of the Digital Divide in Tunja, Colombia (2020–2021) Alejandra Uscategui-Gómez, Byron David Igua Muñoz, Juan Carlos Campos Medina, Joan Sebastian Igua Muñoz, Jonathan Salomón Igua-Muñoz	222
Understanding LMS Acceptance in Vocational Secondary Education: Extending the Technology Acceptance Model Using SEM-GSCA Lifa Farida Panduwinata, Prisilia Joyceline Atmojo, Febrika Yogie Hermanto	237
From Knowledge to Instruction: Examining the Roles of Usefulness and Teaching Self-Efficacy on Postgraduate Students' Willingness to Teach Ethnomathematics Ebenezer Kwesi Lotey, Ebenezer Bonyah	254
Beyond A-Level Points: Evaluating their Relevance to Responsible and Efficient University Admissions in the Digital Era Desmond Mwembe, Brighton Mpofu, Nicholas J Sibanda, Aaron Nyoni	269
Assessment Practices and Challenges of Social Studies: Teachers during Heat-index Class Suspensions Czar Justine D. Nvalta	281

BIOETHICAL-COMMUNICATIVE JUDGMENTS ACROSS TRAINING YEARS: CROSS-YEAR EVIDENCE FROM UZBEK MEDICAL EDUCATION

Gulnoza Salimova

Abu Ali ibn Sino Bukhara State Medical
Institute, Uzbekistan

✉ salimova.gulnoza@bsmi.uz

ABSTRACT

Whether bioethical-communicative judgments vary across stages of medical education remains underexplored in Central Asian contexts. This pilot cross-sectional study examined scenario-based judgments among 100 undergraduate medical students from six year groups and seven universities in Uzbekistan using a purpose-designed six-scenario Situational Judgment Test (SJT). Five Likert-format scenarios were analysed separately, with S4 and S5 reverse-coded only for directional consistency, while S6 used a forced-choice format. After Holm correction, significant year-group differences remained for patient autonomy and consent (S3: $H(5) = 28.89$, $p < .001$, $\epsilon^2 = .254$), comprehension verification (S4: $H(5) = 14.52$, $p = .038$, $\epsilon^2 = .101$), and emotional responsiveness (S5: $H(5) = 18.72$, $p = .009$, $\epsilon^2 = .146$), but not for prognosis disclosure or hierarchical challenge. In Year 1 versus Years 2–6 comparisons, Year 1 students rated S2 and S3 higher, whereas Years 2–6 rated S4 and S5 higher; S1 did not differ significantly. S6 responses were divided among refusal, acceptance, and compromise. Overall, cross-year variation was domain-specific and bidirectional. Given the cross-sectional design, unequal group sizes, and exploratory instrument, the findings should be interpreted as hypothesis-generating rather than causal. They identify domains warranting further longitudinal, qualitative, and behaviourally validated investigation within Uzbek medical education settings.

KEYWORDS

Bioethical norms, Central Asian medical education, hidden curriculum, pragmatic socialization, professional responsibility, situational judgment test

HOW TO CITE

Salimova G. (2026) 'Bioethical-Communicative Judgments Across Training Years: Cross-Year Evidence from Uzbek Medical Education', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 184–195. <http://dx.doi.org/10.7160/eriesj.2026.190301>

Article history

Received

May 24, 2026

Received in revised form

August 10, 2026

Accepted

August 31, 2026

Available on-line

September 30, 2026

Highlights

- Cross-year patterns were domain-specific and bidirectional.
- Year 1 rated autonomy and hierarchical challenge higher in focused comparisons.
- Later-year students rated comprehension and emotional responsiveness higher.
- Findings are exploratory and do not establish causal training effects.

INTRODUCTION

Professional training can shape not only technical knowledge but also how practitioners judge context-sensitive communicative situations. In medical education, such judgments are particularly important in situations involving truth-telling, informed consent, professional hierarchy, patient comprehension, emotional distress, and professional boundaries (Ha and Longnecker, 2010). Pragmatic socialization research similarly suggests that sustained participation in institutional environments may influence communicative norms and expectations (Bardovi-Harlig, 1999; Duff, 2010). In the present study, however, scenario responses are treated as judgments of bioethical-communicative

appropriateness rather than direct measures of sociopragmatic competence or observed clinical behaviour.

Medical education research indicates that communicative and professional development does not necessarily follow a uniformly progressive trajectory. Previous studies have identified challenges involving empathy, hierarchy, speaking up, and patient-centred communication during training (Neumann et al., 2011; Salehi et al., 2020). Situational Judgment Tests (SJTs) have also been used to examine professionalism and context-sensitive decision-making among medical students, although their results should not be equated directly with actual behaviour (Sahota et al., 2023).

What remains insufficiently understood is whether bioethical-communicative judgments differ across year groups in culturally and institutionally distinctive medical education settings, and whether any observed differences are consistent across communicative domains. Evidence addressing these questions in Central Asian medical education remains particularly limited.

Accordingly, the objective of this pilot cross-sectional study was to examine whether scenario-based bioethical-communicative judgments differed across the six years of undergraduate medical education in Uzbekistan, to identify the communicative domains showing the strongest and weakest cross-year variation, and to determine whether the direction of these differences was consistent across dilemmas. Because the study compared different year-group cohorts at a single time point, its objective was comparative and exploratory rather than developmental or causal.

Clinical communication is also culturally and institutionally situated. Comparative research has shown that preferences concerning disclosure, patient involvement, family participation, and professional roles may vary across healthcare contexts, making it problematic to assume that a single communication model operates uniformly across settings (Surbone, 2006; Pun et al., 2018). In Central Asia, medical education has undergone substantial institutional change since the dissolution of the Soviet system, with reforms affecting curricula, training structures, accreditation, and professional education (Conaboy et al., 2005). Health-system analysis from Uzbekistan has likewise documented continuing reform in medical education, professional development, and quality-assurance structures (Ahmedov et al., 2014). These bodies of literature provide important contextual background, but they do not establish how Uzbek medical students evaluate specific bioethical-communicative dilemmas across year groups.

Three research questions guided the study:

- RQ1: Do scenario-based appropriateness ratings differ across year groups among Uzbek medical students?
- RQ2: Which scenarios show the strongest and weakest cross-year variation, and are the observed patterns consistent in direction?
- RQ3: What do these scenario-specific patterns suggest, cautiously, about the interaction between internationally promoted bioethical-communicative norms and culturally or institutionally salient expectations?

The study is relevant to questions of educational effectiveness and professional responsibility. From an educational perspective, it examines whether students at different stages of training evaluate key communicative dilemmas similarly or differently. From a professional perspective, it considers how principles such as informed consent, transparent communication, patient comprehension, and professional boundaries are judged within a specific medical education context. Because the design is cross-sectional, the study does not establish that training causes the observed differences.

The remainder of the paper reviews relevant literature on pragmatic socialization, the hidden curriculum, SJTs, and medical communication; describes the study design, instrument, sample, and statistical analysis; presents the scenario-level

results; discusses their implications and limitations; and concludes with directions for future research.

Pragmatic Socialization, Fossilization, and Professional Discourse

Pragmatic socialization describes how communicative norms are learned through participation in social and institutional settings (Duff and Talmy, 2011). Pragmatic development is context-sensitive and may follow non-linear trajectories, while persistent non-target forms have been discussed as pragmatic fossilization (Kasper and Rose, 2002; Trillo, 2002). In this study, these concepts serve as theoretical analogies rather than demonstrated mechanisms of medical training.

Medical education similarly places learners within role-specific discourse communities where formal professional expectations may coexist with locally situated interactional practices (Lave and Wenger, 1991). Such tensions may concern authority, patient participation, disclosure, and professional communication. However, the present cross-sectional design cannot determine whether pragmatic socialization or fossilization explains year-group differences. Establishing such processes would require longitudinal and qualitative evidence.

The Hidden Curriculum and Communicative Competence

The hidden curriculum comprises informal values and behavioural norms transmitted alongside formal instruction (Hafferty, 1998). In medical education, these influences may operate through role modelling, informal feedback, hierarchy, and observed clinical practice. Previous research has associated hidden-curriculum processes with empathy, communication, professional behaviour, and ethical reasoning (Neumann et al., 2011; Salehi et al., 2020; Ullah et al., 2025).

The hidden curriculum therefore provides a relevant interpretive framework, but it was not directly measured in this study. Cross-sectional scenario ratings cannot establish whether exposure to competing institutional norms produces pragmatic re-socialization. Longitudinal assessment, qualitative inquiry, and observation of clinical communication would be required to test this possibility.

Situational Judgment Tests in Undergraduate Medical Education

Situational Judgment Tests are used in medical education to examine professionalism, interpersonal behaviour, and context-sensitive decision-making. A systematic review of 24 studies in undergraduate medical training found applications across professionalism, clinical skills, personality attributes, and later educational outcomes, while also highlighting heterogeneous evidence and the need for stronger prospective validation (Sahota et al., 2023). SJT responses should therefore be interpreted as judgments about hypothetical situations rather than direct demonstrations of behaviour.

Supporting evidence includes associations between lower professionalism-based SJT scores and recorded professionalism concerns (Sahota and Taggar, 2020), as well as acceptable reliability for an adapted professionalism-focused SJT across health-professions students (Hosseinpour and Keshmiri,

2024). These findings support SJTs as exploratory measures of professional judgment without equating them with validated communicative competence.

Family-Mediated Decision-Making, Hierarchy, and Emotional Responsiveness

Several study scenarios reflect communicative tensions documented in other healthcare contexts. Family involvement in disclosure and decision-making is one example. Wang et al. (2013) reported substantial family support for withholding a cancer diagnosis because of concern about psychological harm, whereas Wu et al. (2021) found stronger preferences for full disclosure among patients than among relatives or medical professionals. These findings illustrate tensions between patient autonomy and family-protective reasoning.

Professional hierarchy presents a distinct challenge. Trainees may recognise safety or professionalism concerns while remaining reluctant to challenge senior colleagues. Martinez et al. (2017) reported reduced willingness to speak up to attending physicians, while Carrillo et al. (2024) linked speaking-up behaviour with psychological safety. Hierarchical power relations may therefore influence participation and communication in medical education (Salehi et al., 2020; Vanstone and Grierson, 2022; Singh et al., 2025).

Emotional responsiveness is also distinct from general empathy. LaNoue and Roter (2018) linked emotionally responsive communication with patient-centred interaction, whereas Brodahl et al. (2021) showed that biomedical tasks may take priority when patients express emotional concerns. Gift acceptance represents another distinct domain because gratitude and reciprocity may conflict with professional boundaries and obligation (Zolkefli, 2021). These findings support analysing the scenarios as distinct communicative judgments rather than indicators of a single competence.

Year-Group Differences in Pragmatic and Sociolinguistic Awareness

Research on pragmatic awareness across educational levels suggests that development is neither linear nor uniform. Matsumura (2003) showed that proficiency and exposure interact non-linearly, with contextual factors shaping pragmatic trajectories. However, few studies compare scenario-based communicative judgments across all stages of undergraduate medical training. The present study addresses this gap by comparing all six year groups in an Uzbek medical education sample while treating the observed differences as cross-sectional rather than developmental.

Sociopragmatic Competence in the Uzbek Medical Education Context

Research on bioethical-communicative judgments among medical students in Uzbekistan remains limited. Existing Central Asian literature has focused mainly on post-Soviet reforms in curricula, accreditation, professional training, and institutional capacity (Conaboy et al., 2005). Uzbekistan has also undergone reforms in medical education, professional development, treatment standards, and quality assurance (Ahmedov et al., 2014). At the same time, substantial

variation among former Soviet health systems cautions against treating the region as a single explanatory cultural category (Rechel et al., 2014).

The specific gap concerns how Uzbek medical students evaluate dilemmas involving autonomy, family participation, professional hierarchy, comprehension, emotional responsiveness, and professional boundaries, and whether these judgments differ consistently across year groups. Evidence from other Asian settings shows variation in family involvement, disclosure, patient participation, and culturally situated communication expectations (Pun et al., 2018; Surbone, 2006), but such findings cannot be assumed to apply directly to Uzbekistan.

The present study uses the five MedSPC components as an analytical scaffold: Sociopragmatic Knowledge (Kasper and Rose, 2002), Discourse Organizational Competence (Bardovi-Harlig, 1999), Interactional Role Competence (Lave and Wenger, 1991), Face Management Competence (Brown and Levinson, 1987), and Pragmatic Adaptability (Duff, 2010). MedSPC (Salimova, 2026) is used heuristically rather than treated as an independently validated theoretical or psychometric model.

Accordingly, the SJT should be regarded as an exploratory scenario-based instrument, not a validated psychometric scale. Its five Likert-format scenarios represent distinct communicative domains and are analysed separately rather than assumed to measure a common latent construct. The present study does not establish structural, criterion-related, temporal, or behavioural validity. Further validation should include cognitive interviewing, test-retest assessment, independent behavioural or educational criteria, larger item pools, and more balanced samples.

MATERIALS AND METHODS

Research Design

This pilot study used a cross-sectional quantitative design with a purpose-built Situational Judgment Test (SJT) as the primary research instrument. SJTs have been used in medical education to examine context-sensitive professional judgment, interpersonal reasoning, and professionalism-related decision-making, although evidence of validity depends on the specific instrument, population, and external criterion against which performance is evaluated (Patterson et al., 2016; Sahota et al., 2023). The present SJT was therefore used as an exploratory scenario-based instrument rather than as an established measure of clinical competence. Given the study's applied-linguistic component, the methodological decisions and statistical reporting were made explicit in accordance with broader recommendations concerning transparency and quality in quantitative applied-linguistics research (Plonsky and Gass, 2011). A cross-sectional design captures differences between year groups at a single point in time; it does not track individuals as they progress through training. This matters for interpretation because the observed differences are compatible with several plausible explanations, including cohort effects, selection bias, and measurement artefact. They are therefore treated throughout as exploratory, hypothesis-generating findings rather than evidence of developmental change.

Ethical Considerations

The study was conducted with reference to the ethical principles of the Declaration of Helsinki (World Medical Association, 2013) and the British Psychological Society (2021) Code of Human Research Ethics. Participation was voluntary and anonymous, and no personally identifying data were collected. No deception was used. Formal institutional ethics committee approval was not obtained before data collection. This governance limitation is acknowledged explicitly, and future replication studies should obtain prospective ethics review through an appropriate institutional mechanism where available.

Instrument

The Situational Judgment Test (SJT) was developed as a purpose-designed exploratory instrument for the present study rather than as an established or previously validated psychometric scale. Scenario construction was theory-driven and was guided by the five components represented in the MedSPC framework, together with established literature on sociopragmatic competence, pragmatic socialization, face management, interactional roles, and discourse organization in professional communication. The final instrument comprised six scenarios designed to sample distinct bioethical-communicative tensions relevant to medical interaction: prognosis disclosure, hierarchical challenge, patient autonomy and consent, comprehension verification, emotional management, and professional gift acceptance. The six scenarios were intended to provide focused coverage of theoretically salient communicative dilemmas rather than exhaustive representation of the full domain of medical sociopragmatic competence.

Scenario Development and Content Review

Each scenario was constructed to represent a clinically recognizable situation in which an internationally promoted bioethical or communicative principle could come into tension with a culturally or institutionally salient competing expectation. Scenario content was mapped onto the MedSPC components before administration. S1 addressed prognosis disclosure and family-mediated communication; S2 addressed hierarchical challenge and professional accountability; S3 addressed patient autonomy in the presence of spousal objection; S4 addressed judgments about the adequacy of communication when patient comprehension had not been actively verified; S5 addressed judgments about continuing information delivery while disregarding a patient's acute emotional distress; and S6 addressed professional boundaries in the context of culturally meaningful gift giving.

The scenario set was independently reviewed by two bilingual academics whose combined expertise encompassed applied linguistics, medical education, clinical medicine, and intercultural communication. Review focused on four criteria: cultural appropriateness for the Uzbek medical education context, conceptual clarity, comprehensibility of the scenario wording, and consistency between each scenario's intended communicative dilemma and the response task presented to participants. Feedback from the reviewers was considered

during revision of the instrument before data collection, and wording was refined where necessary to improve clarity and cultural appropriateness.

The review process was formative rather than a formal quantitative content-validity study. Reviewers did not provide standardised numerical ratings of item relevance, and no content validity index or inter-rater agreement coefficient was calculated. Accordingly, the expert review provides preliminary support for the conceptual and cultural plausibility of the scenarios but should not be interpreted as formal evidence of content validity. Future validation studies should employ a predefined candidate-item pool, structured expert-rating procedures, quantitative content-validity indices, cognitive interviewing, and independent replication before the instrument is used as a validated assessment measure.

This scenario-to-domain mapping was retained throughout the analysis and interpretation of the scenario-level results.

S1-S5 used a five-point Likert-type appropriateness scale, whereas S6 used three nominal forced-choice options and was analysed separately. Because S4 and S5 were worded in the opposite evaluative direction to S1-S3, they were reverse-coded as 6 minus the original response solely to provide a consistent direction for scenario-level interpretation. No aggregate SJT score was calculated.

The SJT was developed and administered in Uzbek. Corresponding English renderings were prepared for international reporting and reviewed by bilingual academics for conceptual clarity and semantic correspondence. Formal independent back-translation was not undertaken because English was not the administered instrument. Cross-language equivalence should therefore not be assumed, and future cross-language validation should include independent translation, blinded back-translation, reconciliation, and cognitive interviewing. The final Uzbek instrument was administered via Google Forms with mandatory response settings.

The complete SJT instrument is provided in Supplementary Appendix A, including the administered Uzbek wording of all six scenarios, the corresponding English renderings prepared for international reporting, the response anchors for S1-S5, and all forced-choice response options for S6.

One interpretive boundary applies throughout: responses to the SJT represent scenario-based judgments of communicative appropriateness rather than behaviourally demonstrated sociopragmatic competence. Higher or lower scenario ratings should therefore not be interpreted as direct measures of clinical communication ability. The instrument remains exploratory, and the observed scenario-level patterns require replication and further validation before they can support educational or policy decisions.

Participants

Undergraduate medical students were recruited through purposive sampling across seven medical universities in Uzbekistan: Bukhara State Medical Institute, Tashkent State Medical University, Tashkent Medical Academy, Samarkand State Medical University, Andijan State Medical Institute, Fergana Public Health Medical Institute, and the First Moscow State Medical University (Sechenov) branch in Tashkent.

Recruitment across multiple institutions was intended to broaden institutional representation within the sample rather than to support institution-level comparisons. The study was not powered to examine differences between individual universities. Eligibility was limited to undergraduate medical students enrolled in the invited student groups at the participating institutions. Only completed questionnaire submissions were included in the analysis.

Participation was voluntary and anonymous throughout. Recruitment was facilitated through academic contacts across the participating institutions, who distributed the survey invitation to 11 existing student groups comprising 176 eligible undergraduate medical students in total. Of these, 100 students completed the questionnaire and were included in the final analysis, yielding an overall response rate of 56.8%. The final sample represented all six years of undergraduate medical training: Year 1 ($n = 43$), Year 2 ($n = 8$), Year 3 ($n = 10$), Year 4 ($n = 16$), Year 5 ($n = 10$), and Year 6 ($n = 13$).

Year 1 students constituted 43% of the final sample, reflecting differential availability across the participating student groups rather than a deliberate sampling preference. This imbalance has important analytical implications because statistical power is concentrated disproportionately in comparisons involving Year 1, including the focused comparison between Year 1 and Years 2-6. Effect-size estimates derived from these comparisons should therefore be interpreted cautiously given the unequal and relatively small subgroup sizes. Future replication studies should employ stratified recruitment procedures aimed at achieving substantially more balanced representation across year groups and should determine the required sample size prospectively through an appropriate power analysis.

Scoring and Statistical Analysis

Responses to S1-S5 were analysed at the individual scenario level rather than aggregated into a total score. Ratings for S1-S3 were retained in their original 1-to-5 coding. Because S4 and S5 were phrased in the opposite evaluative direction, these two items were reverse-coded using the transformation 6 minus the original rating. Following this direction standardisation, higher values on all five scenarios consistently indicated stronger endorsement of the target bioethical-communicative position represented by that scenario. Reverse-coding was used solely to facilitate consistent interpretation across scenarios and did not create an aggregate psychometric scale.

For each Likert-format scenario, descriptive statistics were calculated for the full sample and by year group. Because each scenario consisted of a single ordinal Likert-type rating and several year groups were small and unequal in size, inferential comparisons were conducted using rank-based procedures. Differences across the six year groups were examined separately for S1-S5 using the Kruskal-Wallis test, with epsilon-squared (ϵ^2) reported as the effect-size measure for each scenario. To control the family-wise error rate arising from the five scenario-level omnibus tests, p values were adjusted using the Holm procedure.

For scenarios whose Holm-adjusted omnibus test remained statistically significant, exploratory pairwise comparisons across the six year groups were conducted using Dunn's

test. Holm adjustment was applied separately across the 15 pairwise comparisons within each scenario. Full pairwise matrices, including non-significant contrasts, are reported in Supplementary Table S1.

A substantively motivated focused comparison between Year 1 and Years 2-6 combined was also examined separately for each of the five Likert-format scenarios using two-sided Mann-Whitney U tests. Rank-biserial correlation was reported as the effect-size measure for these comparisons, and Holm adjustment was applied across the five scenario-level contrasts. This focused analysis was treated as exploratory and complementary to, rather than a substitute for, the six-group omnibus tests. No internal-consistency coefficient, item-total correlation, aggregate SJT score, or sensitivity analysis based on removal of individual items was used because the scenarios were treated as distinct judgments rather than indicators of a single unidimensional scale.

For the forced-choice S6 scenario, the association between year group and response category was initially examined using Pearson's chi-square test of independence, with Cramér's V reported as the effect-size measure. Because several expected cell frequencies were below five, the asymptotic chi-square result was supplemented by a Monte Carlo test conditional on the observed row and column marginal totals. The Monte Carlo procedure was based on 100,000 randomly generated contingency tables using a fixed random seed of 20260807. An exploratory post-hoc comparison of Year 1 with Years 3-5 was subsequently conducted because Years 3-5 showed a common descriptive pattern in which Options B and C predominated. Responses were collapsed into Option A versus Options B and C combined for this comparison. Statistical significance for this 2×2 comparison was assessed using a two-sided Fisher's exact test, with the odds ratio reported as an effect-size estimate. Because the contrast was selected after inspection of the descriptive pattern, its p value was not treated as confirmatory evidence and no multiplicity-adjusted claim is made for this post-hoc comparison.

Given the pilot nature of the study, the unequal year-group sizes, and the relatively small numbers of participants in several senior-year groups, all scenario-level inferential results and effect-size estimates were interpreted cautiously and treated as exploratory. All statistical tests were two-sided, with statistical significance evaluated at $p < .05$. Statistical analyses were conducted in Python 3.11 using NumPy and SciPy. Monte Carlo contingency tables were generated using `scipy.stats.random_table` with a fixed random seed of 20260807, and the exploratory 2×2 comparison was evaluated using `scipy.stats.fisher_exact`. The commented analysis script and de-identified participant-level dataset are provided as supplementary reproducibility materials.

RESULTS

Scenario-Level Descriptive Statistics

Following direction standardisation, the five Likert-format scenarios showed clearly different response profiles (Table 1). Mean ratings were lower for S1, prognosis disclosure ($M = 1.86$, $SD = 1.03$), S2, hierarchical challenge ($M = 2.20$, $SD = 1.13$),

and S3, patient autonomy and consent ($M = 2.15$, $SD = 1.24$). In contrast, the direction-standardised means were higher for S4, comprehension verification ($M = 3.84$, $SD = 1.16$), and S5, emotional responsiveness ($M = 4.04$, $SD = 1.02$). For S4 and S5, the displayed values reflect reverse-coded

ratings, such that higher scores consistently indicate stronger endorsement of the target bioethical-communicative position. These descriptive differences indicate substantial variation across communicative domains and reinforce the rationale for reporting the five scenarios separately.

Scenario	Theme	M	SD
S1	Prognosis disclosure vs. family wishes	1.86	1.03
S2	Hierarchical challenge and patient safety	2.20	1.13
S3	Patient autonomy vs. spousal objection	2.15	1.24
S4	Comprehension verification	3.84	1.16
S5	Emotional responsiveness	4.04	1.02

Note: Ratings ranged from 1 to 5. S1-S3 retain their original coding. S4 and S5 were reverse-coded using 6 minus the original response so that higher values consistently indicate stronger endorsement of the target bioethical-communicative position. No total score was calculated across the five scenarios.

Table 1: Scenario-Level Descriptive Statistics for Direction-Standardised Ratings ($N = 100$), 2026 (source: own calculation)

Item-Level Response Distributions

Response distributions further demonstrated that endorsement varied substantially by scenario (Table 2). For S1-S3, direction-standardised ratings of 1-2 predominated, accounting for 75%, 72%, and 67% of responses, respectively. In contrast, after reverse-coding S4 and S5, ratings of 4-5 accounted for 73% of responses to S4 and

80% of responses to S5. Thus, the response pattern differed in direction across communicative domains: target positions relating to prognosis disclosure, hierarchical challenge, and autonomous consent received relatively limited endorsement, whereas target positions relating to verification of patient comprehension and responsiveness to emotional distress received substantially stronger endorsement.

Scenario	M	SD	% Rating 1-2	% Rating 3	% Rating 4-5
S1: Prognosis disclosure	1.86	1.03	75%	16%	9%
S2: Hierarchical challenge	2.20	1.13	72%	12%	16%
S3: Patient autonomy/consent	2.15	1.24	67%	16%	17%
S4: Comprehension verification	3.84	1.16	14%	13%	73%
S5: Emotional responsiveness	4.04	1.02	9%	11%	80%

Note: S1-S3 retain their original coding, whereas S4 and S5 were reverse-coded for directional consistency. Across all five scenarios, ratings of 1-2 therefore indicate weaker endorsement and ratings of 4-5 indicate stronger endorsement of the target bioethical-communicative position. Percentages may not sum to 100% because of rounding.

Table 2: Direction-Standardised Scenario Response Distributions ($N = 100$), 2026 (source: own calculation)

Year-Group Differences Across the Likert-Format Scenarios

Kruskal-Wallis tests showed that evidence of year-group differences varied across the five Likert-format scenarios (Table 3). After Holm correction for five scenario-level comparisons, statistically significant year-group differences remained for S3, patient autonomy and consent, $H(5) = 28.89$, Holm-adjusted $p < .001$, $\epsilon^2 = .254$; S4, comprehension verification, $H(5) = 14.52$, Holm-adjusted $p = .038$, $\epsilon^2 = .101$; and S5, emotional responsiveness, $H(5) = 18.72$, Holm-adjusted $p = .009$, $\epsilon^2 = .146$.

The omnibus difference for S2, hierarchical challenge, was statistically significant before multiplicity adjustment, $H(5) = 11.78$, unadjusted $p = .038$, but did not remain significant after Holm correction (adjusted $p = .076$, $\epsilon^2 = .072$). No statistically significant year-group difference was observed for S1, prognosis disclosure, $H(5) = 7.50$, adjusted $p = .186$, $\epsilon^2 = .027$. These results indicate that cross-year variation was scenario-specific rather than uniformly distributed across the five communicative domains.

Scenario	$H(5)$	Unadjusted p	Holm-adjusted p	ϵ^2
S1: Prognosis disclosure	7.50	.186	.186	.027
S2: Hierarchical challenge	11.78	.038	.076	.072
S3: Patient autonomy/consent	28.89	< .001	< .001	.254
S4: Comprehension verification	14.52	.013	.038	.101
S5: Emotional responsiveness	18.72	.002	.009	.146

Note. $df = 5$ for all Kruskal-Wallis tests. Holm-adjusted p values control the family-wise error rate across the five scenario-level omnibus tests. S4 and S5 were reverse-coded before analysis. Epsilon-squared (ϵ^2) is reported as the effect-size measure.

Table 3: Kruskal-Wallis Tests of Year-Group Differences Across Direction-Standardised SJT Scenarios, 2026 (source: own calculation)

Exploratory Dunn-Holm comparisons for S3, S4, and S5 are reported in Supplementary Table S1. For S3, Year 1 ratings were higher than those of Years 3, 4, and 5 (Holm-adjusted $p = .001$, $.005$, and $.005$, respectively). For S4, Year 4 ratings were higher than Year 1 ratings (adjusted $p = .012$). For S5, Years 4 and 5 ratings were higher than Year 1 ratings (adjusted $p = .013$ and $.047$, respectively). No other pairwise comparison remained statistically significant after Holm adjustment.

Scenario-Level Analysis by Year Group

Descriptive year-group means showed that cross-year differences were not uniform in direction across the five Likert-format scenarios (Table 4). For S1-S3, Year 1 students showed higher direction-standardised mean ratings than students in Years 2-6 combined. The largest descriptive difference occurred for S3, patient autonomy and consent, with a Year 1

mean of 2.84 compared with 1.63 in Years 2-6, corresponding to a difference of 1.21 points. The corresponding differences were 0.87 points for S2, hierarchical challenge, and 0.49 points for S1, prognosis disclosure.

A contrasting pattern was observed for S4 and S5 after direction standardisation. For S4, comprehension verification, the Year 1 mean was 3.35 compared with 4.21 in Years 2-6, corresponding to a difference of -0.86 points. For S5, emotional responsiveness, the respective means were 3.60 and 4.37, corresponding to a difference of -0.76 points. Thus, later-year students showed lower endorsement than Year 1 students for the target positions represented by S1-S3, but higher endorsement for those represented by S4 and S5. This pattern indicates domain-specific cross-year variation rather than a common directional shift across all communicative scenarios.

Scenario	Year 1 ($n = 43$)	Year 2 ($n = 8$)	Year 3 ($n = 10$)	Year 4 ($n = 16$)	Year 5 ($n = 10$)	Year 6 ($n = 13$)	Years 2-6 ($n = 57$)	Difference Y1 - Y2 - 6
S1	2.14	1.62	1.70	1.31	1.70	2.00	1.65	+0.49
S2	2.70	2.12	1.70	1.69	1.80	1.92	1.82	+0.87
S3	2.84	2.25	1.20	1.50	1.30	2.00	1.63	+1.21
S4	3.35	3.75	4.10	4.50	4.40	4.08	4.21	-0.86
S5	3.60	3.88	4.30	4.56	4.60	4.31	4.37	-0.76

Note. All values are direction-standardised item means on a 1-to-5 scale. S1-S3 retain their original coding, whereas S4 and S5 were reverse-coded using 6 minus the original response. Higher values therefore consistently indicate stronger endorsement of the target bioethical-communicative position. The Difference column represents the Year 1 mean minus the Years 2-6 combined mean. Positive values indicate higher Year 1 endorsement, whereas negative values indicate higher endorsement among Years 2-6. Descriptive comparisons should be interpreted cautiously given the unequal year-group sizes.

Table 4: Direction-Standardised Mean Scenario Ratings by Year Group, 2026 (source: own calculation)

The focused Year 1 versus Years 2-6 comparisons further showed that the direction and magnitude of group differences varied across scenarios. After Holm correction for five comparisons, Year 1 students showed significantly higher ratings for S2, $U = 1668.0$, adjusted $p = .002$, rank-biserial $r = .361$, and S3, $U = 1878.5$, adjusted $p < .001$, rank-biserial $r = .533$. The difference for S1 was smaller and did not reach statistical significance after adjustment, $U = 1482.0$, adjusted $p = .055$, rank-biserial $r = .209$.

The direction was reversed for S4 and S5. Students in Years 2-6 showed significantly higher direction-standardised ratings than Year 1 students for S4, $U = 770.5$, adjusted $p = .002$, rank-biserial $r = -.371$, and S5, $U = 696.0$, adjusted $p < .001$, rank-biserial $r = -.432$. Taken together, these focused contrasts provide no evidence of a uniform Year 1 advantage across the SJT scenarios. Instead, they show opposing cross-year patterns across different communicative domains.

Scenario	Mann-Whitney U	Unadjusted p	Holm-adjusted p	Rank-biserial r
S1: Prognosis disclosure	1482.0	.055	.055	.209
S2: Hierarchical challenge	1668.0	.001	.002	.361
S3: Patient autonomy/consent	1878.5	< .001	< .001	.533
S4: Comprehension verification	770.5	< .001	.002	-.371
S5: Emotional responsiveness	696.0	< .001	< .001	-.432

Note: Two-sided Mann-Whitney U tests were used for all comparisons. Holm-adjusted p values control the family-wise error rate across the five focused scenario-level comparisons. Rank-biserial correlation is reported as the effect-size measure. Positive values indicate higher ratings in Year 1, whereas negative values indicate higher ratings in Years 2-6. S4 and S5 were reverse-coded before analysis.

Table 5: Focused Year 1 Versus Years 2-6 Comparisons Across Direction-Standardised SJT Scenarios, 2026 (source: own calculation)

S1 showed the least evidence of cross-year differentiation. Although the Year 1 mean was descriptively higher than the Years 2-6 mean by 0.49 points, the focused rank-based comparison was not statistically significant (unadjusted $p = .055$; Holm-adjusted $p = .055$). This finding is consistent with the nonsignificant six-group omnibus result for S1

and suggests comparatively limited year-group variation in responses to the prognosis-disclosure scenario.

S3 showed the most pronounced cross-year variation among the scenarios in which Year 1 students had higher ratings. The mean was 2.84 in Year 1, compared with 2.25 in Year 2, 1.20 in Year 3, 1.50 in Year 4, 1.30 in Year

5, and 2.00 in Year 6. Year 3 produced the lowest mean for this scenario ($M = 1.20$, $SD = 0.42$), while the Year 6 mean was higher than those observed in Years 3-5 but remained below the Year 1 level. The six-group difference remained statistically significant after Holm adjustment, and the focused Year 1 versus Years 2-6 comparison produced the largest positive rank-biserial effect among the five scenarios ($r = .533$). Because the data are cross-sectional, however, these differences should not be

interpreted as evidence of within-person decline, recovery, or developmental change.

Forced-Choice Scenario (S6): Gift Acceptance

For S6, which addressed gift acceptance from a grateful patient's family, no single response option achieved majority endorsement across the full sample: Option A was selected by 30% of participants ($n = 30$), Option B by 37% ($n = 37$), and Option C by 33% ($n = 33$) (Table 6).

Year	<i>n</i>	Option A: Refuse (%)	Option B: Accept (%)	Option C: Compromise (%)
Year 1	43	49	23	28
Year 2	8	38	25	38
Year 3	10	0	60	40
Year 4	16	19	50	31
Year 5	10	10	60	30
Year 6	13	15	38	46
Total	100	30	37	33

Note: Percentages represent the proportion of respondents within each year group selecting each response option and are rounded to the nearest whole number; totals may therefore differ slightly from 100%. Option a = refuse on professional ethics grounds; Option B = accept in accordance with culturally embedded reciprocity norms; Option C = pragmatic compromise through acceptance for departmental use. S6 used a nominal forced-choice response format and was analysed separately from the five Likert-format scenarios. Percentages for the smaller year groups should be interpreted cautiously.

Table 6: S6 Gift-Acceptance Response Distribution by Year Group, 2026 (source: own calculation)

Response-option distributions varied across the six year groups, $\chi^2 (10, N = 100) = 18.53$, $p = .047$, Cramér's $V = .304$. However, 13 of the 18 expected cell frequencies were below five, with a minimum expected count of 2.40, making reliance on the asymptotic chi-square approximation alone inappropriate. The supplementary Monte Carlo procedure, based on 100,000 randomly generated contingency tables with a fixed seed of 20260807, produced a similar result (Monte Carlo $p = .043$). The association between year group and S6 response category should nevertheless be interpreted cautiously because of the sparse contingency table and small sizes of several year groups.

An exploratory post-hoc comparison was conducted between Year 1 and Years 3-5 because Years 3-5 showed a common descriptive pattern in which Options B and C predominated. For this comparison, responses were collapsed into Option A versus Options B and C combined. Options B and C were selected by 51% of Year 1 students (22 of 43) compared with 89% of students in Years 3-5 (32 of 36). A two-sided Fisher's exact test yielded $p < .001$, with an odds ratio of 7.64. The odds of selecting Option A rather than Options B or C were approximately 7.6 times higher among Year 1 students than among students in Years 3-5. Because this contrast was selected post hoc after inspection of the descriptive pattern, the p value is exploratory and unadjusted and should not be interpreted as confirmatory evidence. The small subgroup sizes should also be considered when interpreting the effect-size estimate.

Descriptively, Option A was the most frequently selected response among Year 1 students (49%, $n = 21$), whereas Options B and C predominated in Years 3-5. No Year 3 participant selected Option A, while 60% selected Option B and 40% selected Option C. Year 6 showed the highest

proportion selecting Option C (46%, $n = 6$). These year-specific percentages should be interpreted cautiously because several groups contained relatively few participants.

DISCUSSION

Principal Findings and Research Questions

The present pilot study identified domain-specific and bidirectional cross-year variation in Uzbek medical students' scenario-based bioethical-communicative judgments. After multiplicity adjustment, significant omnibus differences were observed for S3, patient autonomy and consent, S4, comprehension verification, and S5, emotional responsiveness, whereas evidence was weaker for S1, prognosis disclosure, and S2, hierarchical challenge. In the focused Year 1 versus Years 2-6 comparison, Year 1 students rated S2 and S3 higher, whereas Years 2-6 rated S4 and S5 higher; S1 showed no significant difference.

These findings address the first two research questions by showing that cross-year differences were scenario-specific rather than uniformly directional. S3 showed the strongest Year 1 versus later-year difference, while S1 showed comparatively little differentiation. The opposite patterns for S4 and S5 further indicate that later-year students were not uniformly less aligned with the target bioethical-communicative positions.

This bidirectional pattern is consistent with literature showing that disclosure, family participation, patient involvement, and professional role expectations vary across and within cultural settings (Surbone, 2006; Pun et al., 2018). Medical SJT research likewise suggests that scenario-based judgments can inform professional reasoning but remain dependent on scenario content and require validation against independent educational or behavioural outcomes (Patterson et al., 2016;

Sahota et al., 2023). The present findings extend this literature by identifying distinct cross-year patterns across autonomy, hierarchy, comprehension, emotional responsiveness, and professional boundaries in an Uzbek medical education sample. The third research question concerns how these patterns may relate to culturally or institutionally salient expectations. The present data identify where judgments differ but cannot establish why. Because participants provided appropriateness ratings rather than qualitative explanations or observed behaviour, interpretations involving hierarchy, family orientation, institutional socialization, or cultural norms remain theoretically informed possibilities rather than demonstrated mechanisms.

Prognosis Disclosure, Family Involvement, and Patient Autonomy

S1 and S3 both involved tensions between patient autonomy and family participation but produced different cross-year patterns. S1 ratings were relatively low across year groups and showed little systematic differentiation, suggesting that family-mediated approaches to difficult disclosure may remain salient across training stages.

This interpretation is consistent with evidence that disclosure preferences differ among patients, relatives, and clinicians. Wang et al. (2013) reported substantial family support for withholding a cancer diagnosis because of concern about psychological harm, whereas Wu et al. (2021) found stronger preferences for full disclosure among patients than among relatives or medical professionals. Surbone (2006) similarly emphasizes that truth-telling involves interactions among autonomy, relationships, individual preferences, and cultural context. These studies do not establish the reasoning of Uzbek students, but they show that disclosure and family involvement are multidimensional communicative issues. The relatively stable S1 pattern therefore warrants further qualitative investigation.

S3 showed a different pattern, with Year 1 students more strongly endorsing the patient's own consent despite spousal objection. This scenario showed the clearest cross-year differentiation among the autonomy-related items, although the mechanism underlying this pattern remains unresolved. Taken together, S1 and S3 indicate that family involvement does not produce a single response pattern. Disclosure, consent, prognosis, and spousal involvement may be evaluated according to different pragmatic and ethical considerations, cautioning against treating collectivism or family orientation as a single explanatory variable.

Hierarchy and Speaking Up

S2 examined the appropriateness of openly challenging a senior physician's potentially serious error in front of a patient. Year 1 students rated this scenario higher than Years 2–6 in the focused comparison, although the six-group omnibus result did not remain significant after multiple-testing correction. This mixed pattern warrants caution but is consistent with evidence that hierarchy can influence speaking-up behaviour in clinical settings.

Martinez et al. (2017) found that trainees were less likely

to raise professionalism-related concerns with attending physicians, while Carrillo et al. (2024) linked speaking-up behaviour with psychological safety. More broadly, hierarchical power relations may shape participation, communication, and professional interaction in medical education (Salehi et al., 2020; Vanstone and Grierson, 2022; Singh et al., 2025).

Importantly, S2 assessed the appropriateness of a public challenge rather than actual speaking-up behaviour. Lower ratings may therefore reflect disagreement with the manner or setting of the challenge rather than unwillingness to address a clinical error. The finding should consequently be treated as a basis for further investigation rather than evidence of deficient patient-safety communication.

Comprehension Verification and Emotional Responsiveness

S4 and S5 showed the opposite pattern to S2 and S3. After direction standardisation, later-year students more strongly endorsed verification of patient comprehension and responsiveness to emotional distress, arguing against a general decline in patient-centred communicative judgments across training.

For S4, higher ratings reflected stronger rejection of communication that did not verify patient understanding. This pattern may indicate greater sensitivity to the practical importance of comprehension, particularly where linguistic, educational, or contextual barriers are present (Ha and Longnecker, 2010). The finding, however, concerns scenario-based judgment rather than observed clinical behaviour.

S5 similarly showed stronger later-year endorsement of responding to patient distress rather than continuing information delivery without acknowledging the patient's emotional state. Previous research links emotionally responsive communication with patient-centred interaction, while also showing that biomedical task demands may compete with attention to emotional concerns (LaNoue and Roter, 2018; Brodahl et al., 2021).

Together, S4 and S5 complicate a simple hidden-curriculum explanation. Later-year students were less supportive of some target positions but more supportive of others, indicating that cross-year variation was domain-specific and bidirectional rather than uniformly directional.

Gift Acceptance and Professional Boundaries

S6 used a forced-choice rather than Likert format, and responses were distributed relatively evenly across refusal, acceptance, and pragmatic compromise, with no majority option. This suggests that gift acceptance was a contested judgment within the sample rather than a domain with a single preferred response.

Patient gifts may involve competing considerations of gratitude, reciprocity, professional boundaries, obligation, and institutional policy (Zolkefli, 2021). The divided responses are therefore consistent with multiple professional and relational considerations. However, because participants did not explain their choices, Option B should not be assumed to represent cultural reciprocity, nor Option A stronger professional ethics. Response distributions also differed across year groups.

The exploratory Year 1 versus Years 3-5 comparison showed greater selection of refusal among Year 1 students, but this finding should be interpreted cautiously because of the small subgroup sizes, sparse cells, and data-driven post-hoc selection of the contrast. The unadjusted result is therefore best treated as hypothesis-generating rather than as evidence of a stable training-stage pattern.

Interpreting Cross-Year Variation

Taken together, the scenario-level results do not support a single developmental account of cross-year variation. The bidirectional patterns across S1-S5 make a uniform process of pragmatic erosion, re-socialization, or normative decline difficult to sustain. Hidden-curriculum theory remains one possible interpretive framework because informal role modelling, institutional hierarchy, and clinical practices can influence professional learning (Hafferty, 1998; Ullah et al., 2025), but the present study did not measure such exposure directly.

Several alternative explanations remain plausible, including cohort composition, prior experience, selection processes, unequal group sizes, institutional variation, and differences in how students interpreted the scenarios. Year group should therefore be treated as a cross-sectional grouping variable rather than as a proxy for within-person development. Longitudinal follow-up, qualitative investigation of reasoning, and observation of clinical communication are required before claims about professional socialization or developmental change can be evaluated more confidently.

Implications for Applied Linguistics and Sociopragmatic Research

The findings have two tentative implications for sociopragmatic research. First, professional pragmatic development may be more domain-specific than a single aggregate competence trajectory would imply. The opposing patterns observed across autonomy, hierarchy, comprehension, and emotional responsiveness illustrate why theoretically distinct communicative judgments should not automatically be combined into a single developmental score.

Second, scenario-based pragmatic assessment is necessarily situated within a normative framework. The present SJT operationalised particular internationally promoted bioethical-communicative positions, but lower endorsement should not automatically be interpreted as weaker communicative competence. A learner may reason effectively within locally salient relational or institutional expectations while diverging from the target position embedded in a particular scenario. This reinforces the importance of distinguishing scenario-based normative judgment from behaviourally demonstrated competence.

The MedSPC framework provides one theoretical structure for organising these domains, but the present study should be understood as an initial empirical application rather than a validation of the framework. Further research is required to establish the dimensional structure of the construct, its relationship to observed communication, and its applicability across linguistic and cultural settings.

Implications for Medical Education

The current findings do not provide sufficient evidence to recommend specific curricular interventions. They do, however, identify several domains that merit closer educational investigation: autonomous consent in the presence of family opposition, communication across professional hierarchies, verification of patient comprehension, responsiveness to emotional distress, and professional boundaries around gifts.

For educators, the most appropriate next step would therefore be needs assessment rather than immediate curriculum redesign. Qualitative interviews, focus groups, observed simulations, or structured debriefing could help determine why students evaluate these scenarios differently and whether the underlying reasoning reflects ethical knowledge, cultural expectations, clinical experience, role hierarchy, or pragmatic considerations not captured by the SJT. Only after these mechanisms are better understood would it be possible to determine whether particular communication-training approaches require modification.

Simulation, role-play, and other experiential approaches may eventually offer useful means of exploring competing communicative expectations, but the present pilot data do not establish their effectiveness in the Uzbek context. Any educational intervention derived from these findings should therefore be evaluated prospectively and should distinguish changes in scenario judgments from changes in actual communicative behaviour.

Limitations

The most important limitation is the cross-sectional design. Students from different year groups were assessed at one point in time, so the study cannot determine whether individual students' judgments changed as they progressed through medical school. Cohort effects, differences in prior experience, selection processes, and other unmeasured factors remain plausible explanations for the observed cross-year patterns.

The unequal year-group distribution is also substantial. Year 1 participants accounted for 43% of the sample, whereas several later-year groups contained only 8 to 16 students. This imbalance reduces the stability of year-specific estimates and is a central reason why all inferential findings should be considered exploratory. Recruitment was purposive and was conducted through 11 existing student groups across seven universities, with an overall response rate of 56.8%. Non-response bias therefore cannot be excluded. Although recruitment across several institutions broadened representation, the study was not designed to estimate institution-specific effects or establish national representativeness. Because institution-by-year composition was not modelled, some observed year-group differences may also reflect differences in institutional composition rather than year of training alone.

A further limitation concerns the instrument. The six-scenario SJT was purpose-built for this pilot study and has not yet undergone full psychometric validation. The five Likert-format scenarios represented distinct communicative judgments and were therefore analysed separately rather than combined into a total score. S4 and S5 required reverse-coding because their administered wording was evaluatively opposite to S1-S3. Direction standardisation facilitated interpretation but does

not establish that the five scenarios represent a common latent construct. The forced-choice S6 scenario was also structurally different and was analysed separately.

Cross-language equivalence has not yet been independently established. The Uzbek instrument was the version administered to participants, whereas the English rendering was prepared for international reporting. Although bilingual review was used to examine conceptual clarity and semantic correspondence, formal back-translation and cognitive interviewing were not conducted. The English rendering should therefore not be treated as a separately validated version of the instrument.

The SJT measured reported judgments about hypothetical situations rather than actual clinical communication. Judgment-behaviour correspondence remains unknown, and social desirability may have influenced responses despite anonymous administration. The study also did not collect open-ended explanations for participants' choices, which limits the extent to which observed ratings can be attributed to particular cultural, ethical, or institutional orientations. Five Likert-format scenarios and one forced-choice scenario cannot capture the full range of sociopragmatic demands encountered in clinical practice.

Finally, the findings derive from a relatively small, unevenly distributed sample within one national educational context. They should not be generalised to all Uzbek medical students or to other post-Soviet, Central Asian, or collectivist contexts without replication. Larger studies with balanced year-group recruitment, independent samples, behavioural measures, qualitative reasoning data, and longitudinal follow-up are needed before stronger developmental, theoretical, or curricular claims can be made.

CONCLUSION

This pilot study provides exploratory cross-year evidence on bioethical-communicative judgments among medical students in Uzbekistan. Cross-year variation was domain-specific and bidirectional rather than uniformly developmental. The strongest differences concerned patient autonomy and consent, comprehension verification, and emotional responsiveness, whereas prognosis disclosure showed limited differentiation. Responses to the forced-choice gift-acceptance scenario were also divided, suggesting variation in how professional boundaries and relational considerations are evaluated.

These findings should be interpreted cautiously. The study was cross-sectional, year-group sizes were unequal, and the SJT measured judgments about hypothetical situations rather than observed behaviour. The results therefore cannot establish causal effects of clinical training or within-person change over time. Cohort composition, prior experience, institutional variation, selection processes, and measurement factors remain plausible explanations. MedSPC should likewise be viewed here as an exploratory analytical framework rather than a validated model.

The study's main contribution is to identify communicative domains that warrant more rigorous investigation in Uzbek medical education. Future research should use larger and more balanced samples, longitudinal and qualitative designs, direct observation of clinical communication, and further validation of the SJT before stronger conclusions are drawn about professional socialization, communicative development, or curriculum design.

REFERENCES

- Ahmedov, M., Azimov, R., Mutalova, Z., Huseynov, S., Tsoyi, E. and Rechel, B. (2014) 'Uzbekistan: Health system review', *Health Systems in Transition*, Vol. 16, No. 5, pp. 1–137.
- Salimova, G. (2026) 'Sociopragmatic Competence in Medical Higher Education: Structure, Criteria, and Measurement Methods', *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, Iss. 1, article accepted and in press. <https://doi.org/10.22049/jalda.2026.31530.1982>
- Bardovi-Harlig, K. (1999) 'Exploring the interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics', *Language Learning*, Vol. 49, No. 4, pp. 677–713. <http://dx.doi.org/10.1111/0023-8333.00105>
- British Psychological Society (2021) *BPS Code of Human Research Ethics*, Leicester: British Psychological Society. <https://doi.org/10.53841/bpsrep.2021.inf180>
- Brodahl, K. Ø., Finset, A., Storøy, H.-L. E. and Pedersen, R. (2021) 'Medical students' expressions of empathy: A qualitative study of verbal interactions with patients expressing emotional issues in a medical interview', *Patient Education and Counseling*, Vol. 104, No. 12, pp. 2936–2943. <http://dx.doi.org/10.1016/j.pec.2021.03.042>
- Brown, P. and Levinson, S. C. (1987) *Politeness: Some universals in language usage*, Cambridge: Cambridge University Press.
- Carrillo, I., Serpa, P., Landa-Ramírez, E., Guilbert, M., Gómez-Ayala, Y., López-Pineda, A. and Mira, J. J. (2024) 'Speaking up about patient safety, withholding voice and safety climate in clinical settings: A cross-sectional study among Ibero-American healthcare students', *International Journal of Public Health*, Vol. 69, p. 1607406. <http://dx.doi.org/10.3389/ijph.2024.1607406>
- Conaboy, K. A., Nugmanova, Z., Yeguebaeva, S., Jaeger, F. and Daugherty, R. M. (2005) 'Central Asian republics: A case study for medical education reform', *Journal of Continuing Education in the Health Professions*, Vol. 25, No. 1, pp. 52–64. <http://dx.doi.org/10.1002/chp.9>
- Duff, P. A. (2010) 'Language socialization into academic discourse communities', *Annual Review of Applied Linguistics*, Vol. 30, pp. 169–192. <http://dx.doi.org/10.1017/S0267190510000048>
- Duff, P. A. and Talmy, S. (2011) 'Language socialization approaches to second language acquisition: Social, cultural, and linguistic development in additional languages', in: Atkinson, D. (ed.), *Alternative approaches to second language acquisition*, London: Routledge, pp. 95–116. <https://doi.org/10.4324/9780203830932-5>
- Ha, J. F. and Longnecker, N. (2010) 'Doctor-patient communication: A review', *Ochsner Journal*, Vol. 10, No. 1, pp. 38–43.
- Hafferty, F. W. (1998) 'Beyond curriculum reform: Confronting medicine's hidden curriculum', *Academic Medicine*, Vol. 73, No. 4, pp. 403–407. <http://dx.doi.org/10.1097/00001888-199804000-00013>

- Hosseinpour, A. and Keshmiri, F. (2024) 'Professionalism assessment of students in clinical education through Situational Judgment Test (SJT)', *BMC Medical Education*, Vol. 24, No. 1, p. 1291. <http://dx.doi.org/10.1186/s12909-024-06205-4>
- Kasper, G. and Rose, K. R. (2002) *Pragmatic development in a second language*, Malden, MA: Blackwell Publishers.
- LaNoue, M. D. and Roter, D. L. (2018) 'Exploring patient-centeredness: The relationship between self-reported empathy and patient-centered communication in medical trainees', *Patient Education and Counseling*, Vol. 101, No. 6, pp. 1143–1146. <http://dx.doi.org/10.1016/j.pec.2018.01.016>
- Lave, J. and Wenger, E. (1991) *Situated learning: Legitimate peripheral participation*, Cambridge: Cambridge University Press.
- Martinez, W., Lehmann, L. S., Thomas, E. J., Etchegaray, J. M., Shelburne, J. T., Hickson, G. B., Brady, D. W., Schleyer, A. M., Best, J. A., May, N. B. and Bell, S. K. (2017) 'Speaking up about traditional and professionalism-related patient safety threats: A national survey of interns and residents', *BMJ Quality & Safety*, Vol. 26, No. 11, pp. 869–880. <http://dx.doi.org/10.1136/bmjqs-2016-006284>
- Matsumura, S. (2003) 'Modelling the relationships among interlanguage pragmatic development, L2 proficiency, and exposure to L2', *Applied Linguistics*, Vol. 24, No. 4, pp. 465–491. <http://dx.doi.org/10.1093/applin/24.4.465>
- Neumann, M., Edelhäuser, F., Tauschel, D., Fischer, M. R., Wirtz, M., Woopen, C., Haramati, A. and Scheffer, C. (2011) 'Empathy decline and its reasons: A systematic review of studies with medical students and residents', *Academic Medicine*, Vol. 86, No. 8, pp. 996–1009. <http://dx.doi.org/10.1097/ACM.0b013e318221e615>
- Patterson, F., Rowett, E., Hale, R., Grant, M., Roberts, C., Cousans, F. and Martin, S. (2016) 'The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia', *BMC Medical Education*, Vol. 16, No. 1, p. 87. <http://dx.doi.org/10.1186/s12909-016-0606-4>
- Plonsky, L. and Gass, S. (2011) 'Quantitative research methods, study quality, and outcomes: The case of interaction research', *Language Learning*, Vol. 61, No. 2, pp. 325–366. <http://dx.doi.org/10.1111/j.1467-9922.2011.00640.x>
- Pun, J. K. H., Chan, E. A., Wang, S. and Slade, D. (2018) 'Health professional-patient communication practices in East Asia: An integrative review of an emerging field of research and practice in Hong Kong, South Korea, Japan, Taiwan, and Mainland China', *Patient Education and Counseling*, Vol. 101, No. 7, pp. 1193–1206. <http://dx.doi.org/10.1016/j.pec.2018.01.018>
- Rechel, B., Richardson, E. and McKee, M. (eds.) (2014) *Trends in health systems in the former Soviet countries, Observatory Studies Series*, No. 35, Copenhagen: European Observatory on Health Systems and Policies.
- Sahota, G. S. and Taggar, J. S. (2020) 'The association between Situational Judgement Test (SJT) scores and professionalism concerns in undergraduate medical education', *Medical Teacher*, Vol. 42, No. 8, pp. 937–943. <http://dx.doi.org/10.1080/0142159X.2020.1772466>
- Sahota, G. S., Fisher, V., Patel, B., Ju, J. K. and Taggar, J. S. (2023) 'The educational value of situational judgement tests (SJTs) when used during undergraduate medical training: A systematic review and narrative synthesis', *Medical Teacher*, Vol. 45, No. 9, pp. 997–1004. <http://dx.doi.org/10.1080/0142159X.2023.2168183>
- Salehi, P. P., Jacobs, D., Suhail-Sindhu, T., Judson, B. L., Azizzadeh, B. and Lee, Y. H. (2020) 'Consequences of medical hierarchy on medical students, residents, and medical education in otolaryngology', *Otolaryngology-Head and Neck Surgery*, Vol. 163, No. 5, pp. 906–914. <http://dx.doi.org/10.1177/0194599820926105>
- Singh, P. K., Singh, S., Ahmad, S., Singh, V. K. and Kumar, R. (2025) 'Navigating power dynamics and hierarchies in medical education: Enhancing faculty experiences and institutional culture', *Journal of Postgraduate Medicine*, Vol. 71, No. 2, pp. 82–90. http://dx.doi.org/10.4103/jpgm.jpgm_728_24
- Surbone, A. (2006) 'Telling the truth to patients with cancer: What is the truth?', *The Lancet Oncology*, Vol. 7, No. 11, pp. 944–950. [http://dx.doi.org/10.1016/S1470-2045\(06\)70941-X](http://dx.doi.org/10.1016/S1470-2045(06)70941-X)
- Trillo, J. R. (2002) 'The pragmatic fossilization of discourse markers in non-native speakers of English', *Journal of Pragmatics*, Vol. 34, No. 6, pp. 769–784. [http://dx.doi.org/10.1016/S0378-2166\(02\)00022-X](http://dx.doi.org/10.1016/S0378-2166(02)00022-X)
- Ullah, H., Huma, S., Tahir, N., Yasin, G., Ashraf, M., Shabana, H., Gabr, B. M., Yunus, M. and Khan, J. S. (2025) 'Effects of hidden curriculum on medical education and strategies to reshape hidden curriculum as a curriculum promotor: A systematic review', *Advances in Medical Education and Practice*, Vol. 16, pp. 2087–2098. <http://dx.doi.org/10.2147/AMEP.S561248>
- Vanstone, M. and Grierson, L. (2022) 'Thinking about social power and hierarchy in medical education', *Medical Education*, Vol. 56, No. 1, pp. 91–97. <http://dx.doi.org/10.1111/medu.14659>
- Wang, D.-C., Peng, X., Guo, C.-B. and Su, Y.-J. (2013) 'When clinicians telling the truth is de facto discouraged, what is the family's attitude towards disclosing to a relative their cancer diagnosis?', *Supportive Care in Cancer*, Vol. 21, No. 4, pp. 1089–1095. <http://dx.doi.org/10.1007/s00520-012-1629-y>
- World Medical Association (2013) *Declaration of Helsinki: Ethical principles for medical research involving human subjects*, [Online], Available at: <https://www.wma.net/wp-content/uploads/2024/10/DoH-Oct2013.pdf> [7 August 2026].
- Wu, J., Wang, Y., Jiao, X., Wang, J., Ye, X. and Wang, B. (2021) 'Differences in practice and preferences associated with truth-telling to cancer patients', *Nursing Ethics*, Vol. 28, No. 2, pp. 272–281. <http://dx.doi.org/10.1177/0969733020945754>
- Zolkefli, Y. (2021) 'Healthcare dilemma towards gift giving by patients', *Malaysian Journal of Medical Sciences*, Vol. 28, No. 5, pp. 137–141. <http://dx.doi.org/10.21315/mjms2021.28.5.14>

STUDENTS' ATTITUDES TOWARD GENERATIVE AI AS A MEDIATING PATHWAY TO PERCEIVED GENERATIVE AI SUPPORT FOR SELF-REGULATED LEARNING IN HIGHER EDUCATION

Juli Sardi¹✉
Darmansyah¹
Oriza Candra¹
Nuzul Hidayat¹
Devi Faizah Yuliana²

¹Universitas Negeri Padang, Indonesia

²IAIN Kerinci, Indonesia

✉ julisardi@ft.unp.ac.id

ABSTRACT

Generative artificial intelligence (GAI) is increasingly integrated into higher education, creating opportunities for flexible and personalised learning while raising questions about AI-supported learning. This study examines whether students' attitudes toward GAI mediate the relationships between perceived ease of use (PEU), perceived usefulness (PU), and perceived Generative AI support for self-regulated learning (GAI-SRL). A cross-sectional survey was conducted with 264 undergraduate students at a public university in Indonesia. Data were analysed using partial least squares structural equation modelling (PLS-SEM). PEU and PU were positively associated with attitudes toward GAI, while attitudes were positively associated with perceived GAI support for self-regulated learning. PEU and PU also showed significant direct relationships with GAI-SRL, and their indirect relationships through attitudes were significant, indicating partial statistical mediation. These findings suggest that students' attitudes toward GAI constitute an intermediate statistical pathway linking technology-related perceptions with perceived GAI support for self-regulated learning. The study extends technology acceptance research by positioning attitudes toward GAI within a learning-oriented framework and highlights responsible, learner-centred use of GAI while preserving students' learning agency.

KEYWORDS

Generative artificial intelligence, higher education, perceived AI support, responsibility, self-regulated learning, technology acceptance

HOW TO CITE

Sardi J., Darmansyah, Candra O., Hidayat N., Yuliana D. F. (2026) 'Students' Attitudes Toward Generative AI as a Mediating Pathway to Perceived Generative AI Support for Self-Regulated Learning in Higher Education', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 196–209. <http://dx.doi.org/10.7160/eriesj.2026.190302>

Article history

Received

February 11, 2026

Received in revised form

August 15, 2026

Accepted

August 27, 2026

Available on-line

September 30, 2026

Highlights

- Attitudes toward GAI form an intermediate pathway to perceived AI support for SRL.
- Attitudes partially mediate PEU and PU relationships with perceived AI support for SRL.
- Responsible AI use should preserve students' learning agency.

INTRODUCTION

Generative Artificial Intelligence (GAI) is rapidly becoming part of higher education, changing how students access information, obtain feedback, and approach academic tasks. Applications such as ChatGPT, Gemini, and Claude can provide explanations, generate examples, offer alternative solutions, and support students throughout the learning process. These capabilities have expanded the role of AI from a supplementary digital resource to a more active component

of contemporary learning environments (Madrid Akpovo, 2025; Chatzopoulos, 2026; Kim and Thille, 2026). Recent research further highlights the potential of GAI to support personalised learning, academic productivity, problem solving, and flexible access to learning resources (Chang et al., 2023; Tariq, 2025; Hou et al., 2026). At the same time, the growing use of GAI has raised concerns about over-reliance, uncritical acceptance of AI-generated information, and the potential weakening of learner autonomy when AI

is used without adequate reflection (Choi et al., 2024; Zhai, Wibowo and Li, 2024; Zhang et al., 2024).

However, the educational value of GAI does not depend solely on what the technology can provide. It also depends on how students perceive and use it during learning. Students must decide when AI assistance is appropriate, how much support to seek, and whether AI-generated information should be accepted, questioned, or verified. These decisions involve more than technical competence; they reflect students' beliefs and attitudes toward the role of AI in their learning (Svenningsson et al., 2022; Yanto et al., 2023; Acosta-Enriquez et al., 2024). Students with similarly favourable perceptions of GAI may therefore engage with the technology in fundamentally different ways. Some may use AI strategically to clarify concepts, compare explanations, and support independent learning, whereas others may rely on AI primarily to complete academic tasks with limited reflection or evaluation (Mahniza et al., 2024; Tariq, 2025; Rahimi and Faskhodi, 2026). This distinction suggests that students' attitudes toward GAI may have a broader educational role than simply predicting technology acceptance or intention to use.

The Technology Acceptance Model (TAM) explains how perceived usefulness (PU) and perceived ease of use (PEU) shape users' evaluations of technology. TAM has also been applied to Generative AI in higher education, where these perceptions have been linked to attitudes, acceptance, and intention to use AI (Hwang, Lee and Lee, 2025; Lee, Ramasamy and Subbarao, 2025). Consistent with this framework, studies have reported that students who perceive GAI as useful and easy to use tend to develop more favourable attitudes and stronger intentions to incorporate AI into academic activities (Karan and Chakma, 2025; Mustofa et al., 2025; Ozili, 2025). Research has also identified related factors, including trust and perceived competence, as relevant to students' acceptance and engagement with AI-supported technologies (Dahri et al., 2024; Park, Lee and Do, 2025; Shahzad, Xu and Asif, 2025). These studies have substantially advanced understanding of why students accept and adopt AI-based technologies.

Technology acceptance, however, does not fully explain how students engage with GAI after adopting it. Educational use also depends on how learners direct, evaluate, and integrate AI assistance into their learning processes. Recent research has therefore highlighted learner agency, critical evaluation, metacognition, and responsible engagement in AI-supported learning (Chauncey and McKenna, 2023; Kim et al., 2025; Ge, 2026). Students' attitudes may consequently provide an important psychological pathway through which perceptions of AI are associated with students' perceived support for learning regulation.

Self-regulated learning (SRL) offers a complementary perspective because it focuses on how learners set goals, plan, monitor progress, use strategies, and reflect on learning (Zimmerman, 2013; Nilson and Zimmerman, 2023). These processes are particularly relevant to higher education, where students are expected to manage learning with increasing autonomy (Schunk and Zimmerman, 2012). GAI can provide explanations, feedback, alternative strategies, and on-demand assistance that may support these processes (Guan, Li and Gu, 2024; Sardi et al., 2025a; Zhang and Chen, 2026).

Recent studies have similarly reported that AI-supported learning can facilitate metacognitive and self-regulatory processes when

students remain actively involved in directing their learning (Molenaar, 2022; Joo, Han and Park, 2026; Melanou and Beege, 2026). However, access to GAI does not necessarily mean that students perceive it as supporting their self-regulated learning. Students still need to evaluate and decide how AI-generated information should be used, particularly because excessive reliance may reduce opportunities for independent reasoning and cognitive engagement (Wang, 2024; Bueno-Vesga et al., 2026). Thus, learners remain the principal agents of regulation, while GAI functions as a resource whose educational value depends partly on how students approach and use it.

These observations reveal a conceptual intersection between technology acceptance and SRL that remains insufficiently explained. TAM helps explain why students perceive and accept GAI, whereas SRL explains how students regulate their learning. Less clear is how students' attitudes toward GAI connect these two perspectives once AI becomes part of the learning process. Existing research has established relationships between perceived usefulness, perceived ease of use, and technology acceptance (Runge, Hebibi and Lazarides, 2025; Wang, Li and Zou, 2025), while research on AI-supported learning has increasingly examined how GAI can assist self-regulatory and metacognitive processes (Fitriati and Willian, 2025; EL Hosayny, Razkane and Yeou, 2026). However, limited research has examined whether students' attitudes toward GAI mediate the relationships between perceived ease of use, perceived usefulness, and perceived Generative AI support for self-regulated learning. This gap leaves the role of students' attitudes insufficiently understood at the intersection between technology-related perceptions and AI-supported learning regulation.

The present study addresses this gap by examining students' attitudes toward GAI as a mediator between technology-related perceptions and perceived GAI support for self-regulated learning. Students' attitudes are conceptualised as an evaluative orientation toward the use of GAI for academic learning, whereas GAI-SRL captures students' perceptions of how GAI supports learning-regulation processes. Because the study uses cross-sectional data, the proposed relationships are examined as statistical associations rather than temporal or causal processes.

The study extends technology acceptance research by examining attitudes toward GAI in a learning-oriented context rather than focusing only on acceptance or behavioural intention. It also connects PEU and PU with perceived GAI support for self-regulated learning within a single mediation model. The findings provide implications for AI-supported learning, particularly regarding student agency, critical evaluation, and responsible use of GAI. The remainder of the paper presents the methods, results, discussion, limitations, and conclusion.

MATERIALS AND METHODS

Research Design

This study employed a cross-sectional quantitative survey design to examine the structural relationships among perceived ease of use (PEU), perceived usefulness (PU), students' attitudes toward Generative Artificial Intelligence (ATT), and perceived Generative AI support for self-regulated learning (GAI-SRL). The analysis examined whether attitudes

toward GAI statistically mediate the relationships between technology-related perceptions and perceived GAI support for self-regulated learning (Creswell, 2012).

The conceptual model integrates the Technology Acceptance Model (TAM) and Self-Regulated Learning (SRL) theory. PEU and PU were specified as exogenous constructs, ATT as the mediating construct, and GAI-SRL as the endogenous construct. In contrast to interpreting SRL as a direct measure of students' self-regulated learning ability, GAI-SRL captures students' perceptions of how Generative AI supports specific learning-regulation processes.

Partial Least Squares Structural Equation Modelling (PLS-SEM) was used to estimate the measurement and structural models and to examine the direct and indirect statistical relationships among the constructs. PLS-SEM was selected because the analytical objective combined estimation of the hypothesised structural relationships with an emphasis on explained variance and predictive assessment of the endogenous constructs. Because the study used cross-sectional data, the estimated relationships and mediation effects were interpreted as statistical associations within the proposed conceptual model rather than as evidence of temporal or causal effects.

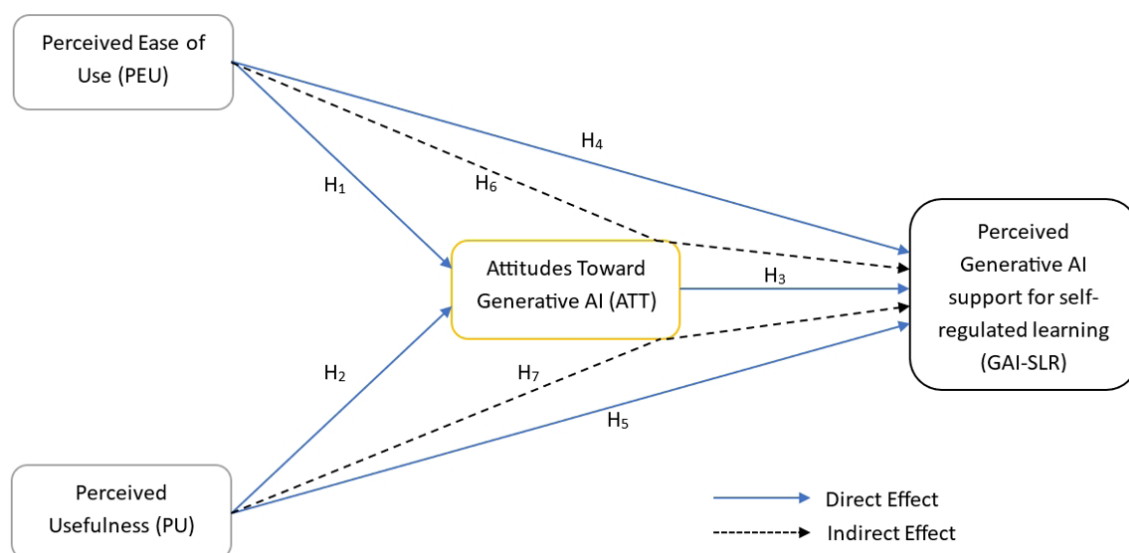


Figure 1: Proposed Conceptual Model of Students' Attitudes Toward Generative AI and Perceived Generative AI Support for Self-Regulated Learning

Proposed Conceptual Model and Hypotheses

The conceptual model proposed in this study is derived from the integration of the Technology Acceptance Model (TAM) and Self-Regulated Learning (SRL) theory. As discussed in the Introduction, TAM provides a robust explanation of how learners develop favourable attitudes toward emerging technologies through their perceptions of usefulness and ease of use. In contrast, SRL theory explains how learners regulate their cognitive, motivational, and behavioural processes to achieve meaningful learning outcomes. Building upon these complementary perspectives, the present study proposes that students' attitudes toward Generative Artificial Intelligence (GAI) conceptually connect technology-related perceptions with self-regulated learning.

Figure 1 illustrates the proposed conceptual model. Within the model, perceived ease of use (PEU) and perceived usefulness (PU) function as exogenous constructs representing students' evaluations of the functional characteristics of Generative AI. Students' attitudes toward GAI (ATT) are positioned as the central mediating construct, reflecting learners' evaluative orientation toward the use of AI in academic learning. Perceived Generative AI Support for Self-Regulated Learning (GAI-SRL) serves as the endogenous construct, representing students' perceptions of how Generative AI supports learning-regulation processes such as goal setting, planning, monitoring, evaluation, strategy formulation, and learning engagement.

Consistent with TAM, students who perceive Generative AI as useful and easy to use are expected to develop more favourable attitudes toward AI-supported learning. These positive attitudes are subsequently expected to be associated with stronger perceptions of Generative AI support for self-regulated learning. In addition to these indirect relationships, the conceptual model also posits that perceived usefulness and perceived ease of use are directly associated with perceived Generative AI support for self-regulated learning. This specification recognises that students' technology-related perceptions may contribute to learning regulation both independently and through their attitudes toward AI. Consequently, the proposed mediation model enables examination of whether students' attitudes provide empirical support for the conceptual perspective advanced in this study.

Based on this conceptual model, the following hypotheses were formulated:

- **H1:** Perceived Ease of Use (PEU) is positively associated with students' attitudes toward Generative AI (ATT).
- **H2:** Perceived Usefulness (PU) is positively associated with students' attitudes toward Generative AI (ATT).
- **H3:** Students' attitudes toward Generative AI (ATT) are positively associated with perceived Generative AI support for self-regulated learning (GAI-SRL).
- **H4:** Perceived Ease of Use (PEU) is positively associated with perceived Generative AI support for self-regulated learning (GAI-SRL).

- **H5:** Perceived Usefulness (PU) is positively associated with perceived Generative AI support for self-regulated learning (GAI-SRL).
- **H6:** Students' attitudes toward Generative AI (ATT) mediate the relationship between Perceived Ease of Use (PEU) and perceived Generative AI support for self-regulated learning (GAI-SRL).
- **H7:** Students' attitudes toward Generative AI (ATT) mediate the relationship between Perceived Usefulness (PU) and perceived Generative AI support for self-regulated learning (GAI-SRL).

Participants and Sampling Procedure

This study was conducted at a public university in Indonesia. The target population comprised undergraduate students who had prior experience using Generative Artificial Intelligence applications, such as ChatGPT, Gemini, or comparable AI-based tools, for academic learning purposes. Participants were recruited using purposive sampling based on predefined eligibility criteria. Participants were required to (1) be currently enrolled as undergraduate students, (2) have used Generative AI for academic learning purposes during the current academic year, and (3) voluntarily agree to participate in the study. Restricting participation to students with prior experience using Generative AI ensured that respondents were able to evaluate AI-supported learning based on actual educational experience rather than hypothetical expectations.

Data were collected through an online questionnaire during the first semester of the 2025/2026 academic year. A total of 278 questionnaires were returned. During the predefined eligibility screening procedure, 9 responses were excluded because the personal information provided could not be verified as indicating active undergraduate student status. A further 5 responses were excluded because the respondents reported very limited use of Generative AI for academic learning and therefore did not meet the study's predefined eligibility criterion concerning prior experience with Generative AI. Thus, 14 responses were excluded and 264 valid responses were retained for the final analysis.

The final sample consisted of 112 male students (42.4%) and 152 female students (57.6%), with a mean age of 20.8 years ($SD = 1.34$). Participation was voluntary, and before completing the questionnaire, respondents were informed about the purpose of the study and provided electronic informed consent. Respondents' names and student identification numbers (NIM) were collected solely for the purpose of verifying active undergraduate student status. These identifying data were not included in the analytical dataset and were deleted before statistical analysis. Accordingly, the statistical analyses were conducted using de-identified data, while responses were treated confidentially.

Instrument Development and Measurement

The research instrument was developed using a theory-informed adaptation and contextualisation approach. The measurement items were selected from established empirical studies and adapted to the context of Generative AI use for academic learning in higher education. The adaptation was guided by the conceptual definitions of the constructs and the requirements of the proposed structural model.

The questionnaire was administered primarily in Indonesian, which was the language used by the respondents. An English equivalent was provided in parentheses for each item in the research documentation and manuscript to facilitate correspondence with the international research literature. The English versions were not used as a separate measurement form.

The adaptation sources were mapped to the constructs as follows: Perceived Ease of Use (PEU; PEU1–PEU5) and Perceived Usefulness (PU; PU1–PU5) were primarily informed by Yanto et al. (2023), Students' Attitudes toward Generative AI (ATT; ATT1–ATT5) were informed by Svenningsson et al. (2022), and Perceived Generative AI Support for Self-Regulated Learning (GAI-SRL; SRL1–SRL7) was informed by Mahniza et al. (2024). The source items were contextualised so that references to technology or learning activities specifically referred to students' academic use of Generative AI.

The adaptation process involved contextual wording adjustments intended to improve clarity, readability, and relevance for Indonesian undergraduate students while preserving the intended conceptual meaning of the source items. The adapted questionnaire was subsequently reviewed by five experts with expertise in educational technology, instructional design, learning psychology, educational measurement, and higher education pedagogy. The experts evaluated the relevance, representativeness, clarity, and contextual appropriateness of the items, and their comments were used to refine wording before the main data collection.

A formal independent back-translation procedure was not conducted because the questionnaire was administered in Indonesian and the English versions were provided only as corresponding English renderings. Likewise, no separate formal cognitive-interview study was conducted before the main survey. Item clarity and contextual suitability were instead examined through the expert-review process. These item structures are reflected in the measurement indicators reported in Table 1.

The Perceived Ease of Use (PEU) construct consisted of five indicators measuring students' perceptions of the usability, accessibility, and comprehensibility of Generative AI in academic learning. These indicators assessed the extent to which students perceived AI systems as easy to learn, straightforward to operate, and usable without requiring specialised technical skills. Likewise, the Perceived Usefulness (PU) construct comprised five indicators representing students' evaluations of the educational benefits of Generative AI, including its contribution to learning effectiveness, efficiency, knowledge acquisition, independent learning, time management, and understanding of complex learning materials. Together, these two constructs represent students' technology-related perceptions, consistent with the Technology Acceptance Model.

The Students' Attitudes toward Generative AI (ATT) construct was operationalised as students' overall evaluative orientation toward the use of Generative AI for academic learning. The construct encompasses complementary cognitive, affective, confidence-related, and responsibility-oriented aspects of students' orientation toward GAI. ATT1 and ATT2 primarily represent students' openness and positive evaluation of GAI as a learning tool. ATT4 reflects confidence in using

Construct	Indicators	Adaptation source
PEU	PEU1. AI technology is easy to use	Yanto et al. (2023)
	PEU2. AI technology is easy to access	
	PEU3. AI technology interfaces are user-friendly and intuitive	
	PEU4. AI technology is easy to understand	
	PEU5. Using AI technology does not require specific skills	
PU	PU1. AI technology helps to improve my knowledge	Yanto et al. (2023)
	PU2. AI technology helps improve learning efficiency and effectiveness	
	PU3. AI technology helps me to be self-reliant	
	PU4. AI technology helps save time	
	PU5. AI helps me understand complex and in-depth learning materials	
ATT	ATT1. I have an open and accepting attitude towards AI for use in learning	Svenningsson et al. (2022)
	ATT2. I have a positive perception of AI as a learning tool	
	ATT3. I take responsibility for verifying AI-generated content before using it	
	ATT4. I am more confident when using AI technology in learning	
	ATT5. I enjoyed using AI technology as a learning tool	
GAI-SRL	SRL1. Generative AI assists in setting clear and achievable learning goals	Mahniza et al. (2024)
	SRL2. Generative AI supports the development of effective learning plans	
	SRL3. Generative AI helps evaluate and reflect on the studied materials	
	SRL4. Generative AI enables consistent monitoring of learning progress	
	SRL5. Generative AI guides the formulation of personalised learning strategies	
	SRL6. Generative AI boosts motivation and willingness to engage in learning activities	
	SRL7. Generative AI makes the learning process more enjoyable and engaging	

Table 1: Constructs, Indicators, and Adaptation Sources (Source: Authors' adaptation based on Yanto et al. (2023), Svenningsson et al. (2022), and Mahniza et al. (2024))

GAI for learning, while ATT5 captures the affective dimension of enjoyment. ATT3 reflects a responsibility-oriented disposition toward verifying AI-generated information before using it academically. This indicator was therefore interpreted as an attitudinal orientation toward responsible AI use rather than as a direct behavioural measure. The five indicators were modelled reflectively as manifestations of an overall evaluative orientation toward GAI. Although they cover complementary facets of attitude, their conceptual coherence and empirical results supported their treatment as a single reflective construct. The broader conceptualisation of ATT remains a limitation to be examined in future research.

The Perceived Generative AI Support for Self-Regulated Learning (GAI-SRL) construct was operationalised to capture students' perceptions of how Generative AI assists, supports, or enables their learning-regulation processes. The seven indicators address goal setting, learning planning, evaluation and reflection, monitoring of learning progress, formulation of personalised learning strategies, motivational engagement, and learning engagement. Accordingly, GAI-SRL does not represent a direct measure of students' self-regulated learning ability. Rather, it represents students' perceived contribution of Generative AI to processes associated with self-regulated learning. This distinction is important because the present study examines students' perceptions of AI-supported learning rather than objectively measuring changes in their self-regulatory competence. All measurement items were assessed using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Content Validity and Expert Review

To strengthen the methodological rigour of the measurement instrument, content validity was assessed through expert judgement using the Content Validity Index (CVI). This procedure was conducted before the main data collection to evaluate whether the adapted measurement items adequately represented their intended constructs and were appropriate for the context of Generative AI-supported learning in higher education.

A panel of five experts was purposively invited to participate in the validation process. The panel represented complementary areas of expertise, including educational technology, instructional design, learning psychology, educational measurement, and higher education pedagogy. Before the evaluation, each expert received the questionnaire, operational definitions of the constructs, and the proposed conceptual model.

Experts independently evaluated the relevance of each indicator using a four-point rating scale ranging from 1 (not relevant) to 4 (highly relevant). I-CVI values were calculated as the proportion of experts assigning a relevance rating of 3 or 4, and S-CVI/Ave was calculated as the average of the item-level CVIs. Given the small size of the five-expert panel, I-CVI values were interpreted cautiously. Following Polit, Beck, and Owen (2007), I-CVI values of approximately 0.78 or higher may be regarded as evidence of good content validity, with the observed values in the present study ranging from 0.80 to 1.00. The S-CVI/Ave values ranged from 0.92 to 0.96 across the four constructs (Table 2).

Latent Construct	Number of Items	Range of I-CVI	S-CVI/Ave
Perceived Ease of Use (PEU)	5	0.80 – 1.00	0.92
Perceived Usefulness (PU)	5	0.80 – 1.00	0.94
Attitudes toward GAI (ATT)	5	0.80 – 1.00	0.96
Perceived GAI Support for Self-Regulated Learning (GAI-SRL)	7	0.80 – 1.00	0.93

Table 2: Content Validity Index (CVI) Results by Latent Construct (Source: Authors' own analysis based on expert judgement)

Data Analysis Procedure

Data analysis was conducted using SmartPLS 3.0 following a two-stage PLS-SEM procedure. The first stage involved assessment of the measurement model, followed by evaluation of the structural model and mediation relationships. Given the cross-sectional research design, all estimated relationships were interpreted as statistical associations rather than causal effects (Hair et al., 2019).

The measurement model was assessed in terms of internal consistency reliability, convergent validity, and discriminant validity. Internal consistency reliability was evaluated using Cronbach's Alpha (CA) and Composite Reliability (CR). Convergent validity was assessed using indicator outer loadings and Average Variance Extracted (AVE). Discriminant validity was examined using the Fornell–Larcker criterion and the Heterotrait–Monotrait Ratio (HTMT).

The structural model was evaluated using path coefficients, coefficient of determination (R^2), collinearity assessment, effect sizes (f^2), predictive relevance (Q^2), and out-of-sample predictive assessment using PLSpredict. Effect size f^2 was used to assess the contribution of each exogenous construct to the explained variance of the endogenous construct. Predictive relevance was examined using the blindfolding procedure, with positive Q^2 values interpreted as evidence of predictive relevance. PLSpredict was additionally used to assess out-of-sample predictive performance using the resulting Q^2 predict values.

Mediation effects were assessed through indirect effects estimated using bootstrapping with 5,000 resamples. The indirect effects of PEU and PU on GAI-SRL through ATT were examined using the bootstrapped estimates and confidence intervals. Consistent with the cross-sectional design, statistically significant indirect effects were interpreted

as evidence of statistical mediation within the proposed model rather than evidence of a temporal or causal mechanism.

Procedural Considerations for Common-Method Variance

Because the predictor, mediator, and endogenous constructs were measured using self-report items within the same online questionnaire and at the same measurement occasion, the potential for common-method variance was considered. Several procedural safeguards were applied, including voluntary participation, electronic informed consent, confidentiality of responses, and de-identification of the analytical dataset.

Names and student identification numbers were collected only to verify active student status and were deleted before statistical analysis. The analytical dataset therefore did not contain direct personal identifiers.

The outer and inner VIF values reported in the structural model were used to assess statistical collinearity and were not treated as a formal test of common-method bias. Therefore, the VIF results should not be interpreted as demonstrating the absence of common-method variance. Given the single-source, self-report measurement approach, the possibility of common-method variance cannot be completely excluded and is considered in the interpretation of the findings.

RESULTS

Measurement Model Assessment

The measurement model demonstrated satisfactory reliability and validity across PEU, PU, ATT, and GAI-SRL. Indicator collinearity, internal consistency, convergent validity, and discriminant validity were within the specified criteria. As shown in Table 3, outer VIF values ranged from 2.849 to 4.135, indicating no excessive indicator collinearity.

Indicator	VIF	Indicator	VIF
PEU1	2.941	ATT1	3.568
PEU2	2.964	ATT2	3.410
PEU3	3.486	ATT3	4.135
PEU4	3.075	ATT4	2.849
PEU5	3.347	ATT5	3.288
PU1	3.346	SRL1	2.967
PU2	3.597	SRL2	3.506
PU3	3.610	SRL3	3.571
PU4	3.699	SRL4	3.339
PU5	3.642	SRL5	3.540
		SRL6	3.783
		SRL7	3.444

Table 3: Outer VIF Values Analysis (Source: own data processing)

Multicollinearity was also assessed at the structural-model level using inner VIF values. As shown in Table 4, all inner VIF values ranged from 1.191 to 1.330, indicating that collinearity among the predictor constructs was also within acceptable limits. The VIF analyses were conducted to assess statistical

collinearity and should not be interpreted as a formal test of common-method bias. Because all constructs were measured through the same self-report questionnaire, the potential for common-method variance is addressed separately as a methodological limitation of the study.

Construct	ATT	GAI-SRL
PEU	1.191	1.316
PU	1.191	1.330
ATT	–	1.279

Table 4: Inner VIF Values Analysis (Source: own data processing)

Internal Consistency Reliability and Convergent Validity

Following the assessment of indicator multicollinearity, internal consistency reliability and convergent validity were evaluated. Internal consistency reliability was examined using Cronbach's Alpha, rho_A, and Composite Reliability (CR), while convergent validity was assessed using indicator outer loadings and Average Variance Extracted (AVE).

As shown in Table 5, outer loadings ranged from 0.860 to 0.917, while Cronbach's Alpha, rho_A, and Composite Reliability ranged from 0.933 to 0.953, 0.936 to 0.953, and 0.949 to 0.961, respectively. AVE values ranged from 0.778 to 0.815. All values exceeded the specified thresholds,

supporting indicator reliability, internal consistency, and convergent validity.

Discriminant Validity

Discriminant validity was assessed to determine whether the four latent constructs represented empirically distinguishable dimensions within the proposed model. The assessment used the Fornell–Larcker criterion and the Heterotrait–Monotrait Ratio (HTMT). As presented in Table 6, the square root of the AVE for each construct exceeded its correlations with the other constructs. The square roots of AVE were 0.888 for ATT, 0.897 for PU, 0.903 for PEU, and 0.882 for GAI-SRL. These results satisfy the Fornell–Larcker criterion and indicate satisfactory discriminant validity.

Construct	Indicator Loading	Cronbach's Alpha (> 0.7)	rho_A (> 0.7)	CR (> 0.7)	AVE (> 0.5)	Convergent Validity
PEU	0.896–0.906	0.943	0.944	0.956	0.815	Valid
PU	0.876–0.917	0.939	0.939	0.953	0.804	Valid
ATT	0.879–0.901	0.933	0.936	0.949	0.789	Valid
GAI-SRL	0.860–0.899	0.953	0.953	0.961	0.778	Valid

Table 5: Reliability and Convergent Validity of the Measurement Model (Source: own data processing)

Constructs	ATT	PU	PEU	GAI-SRL
ATT	0.888			
PU	0.429	0.897		
PEU	0.389	0.401	0.903	
GAI-SRL	0.418	0.456	0.459	0.882

Table 6: Discriminant Validity Assessment Using the Fornell–Larcker Criterion (Source: own data processing)

The HTMT criterion was subsequently examined to provide additional evidence of construct distinctiveness. As shown in Table 7, all HTMT values were below 0.90, with values

ranging from 0.436 to 0.492. Together, the Fornell–Larcker and HTMT results supported discriminant validity among the four constructs.

Constructs	ATT	PU	PEU	GAI-SRL
ATT	—			
PU	0.473	—		
PEU	0.436	0.458	—	
GAI-SRL	0.462	0.489	0.492	—

Table 7: Heterotrait–Monotrait Ratio (HTMT) (Source: own data processing)

Structural Model Assessment

Following confirmation of the adequacy of the measurement model, the structural model was evaluated to examine the hypothesised relationships among PEU, PU, ATT, and GAI-SRL. The assessment considered direct path

coefficients, statistical significance, coefficient of determination (R^2), effect sizes (f^2), predictive relevance (Q^2), and out-of-sample predictive performance using PLSpredict. The mediating role of ATT was subsequently examined through bootstrapped indirect effects. Because

the study employed a cross-sectional design, the structural results are interpreted as statistical associations within the proposed conceptual model. In particular, statistically significant indirect effects are interpreted as evidence of statistical mediation rather than as evidence of a temporal or causal mechanism.

Direct Effects

The structural relationships were evaluated using the bootstrapped path coefficients, standard errors, t-values, p-values, and 95% confidence intervals. As presented in Table 8, all five hypothesised direct relationships were positive and statistically significant, with all 95% confidence intervals excluding zero.

Hypothesis	Path	β	SE	t	p	95% CI	Decision
H ₁	PEU → ATT	0.259	0.053	4.902	< 0.001	[0.152, 0.357]	Supported
H ₂	PU → ATT	0.325	0.051	6.353	< 0.001	[0.221, 0.420]	Supported
H ₃	ATT → GAI-SRL	0.199	0.059	3.339	0.001	[0.077, 0.315]	Supported
H ₄	PEU → GAI-SRL	0.278	0.056	4.954	< 0.001	[0.160, 0.379]	Supported
H ₅	PU → GAI-SRL	0.259	0.053	4.845	< 0.001	[0.150, 0.360]	Supported

Table 8: Direct Effects and Hypothesis Testing (Source: own data processing)

Perceived ease of use was positively associated with attitudes toward GAI ($\beta = 0.259$, $SE = 0.053$, $t = 4.902$, $p < 0.001$, 95% CI [0.152, 0.357]), supporting H1. Perceived usefulness showed a stronger positive association with ATT ($\beta = 0.325$, $SE = 0.051$, $t = 6.353$, $p < 0.001$, 95% CI [0.221, 0.420]), supporting H2. ATT was positively associated with perceived GAI support for self-regulated learning (GAI-SRL) ($\beta = 0.199$, $SE = 0.059$, $t = 3.339$, $p = 0.001$, 95% CI [0.077, 0.315]), supporting H3. PEU also demonstrated a positive association with GAI-SRL ($\beta = 0.278$, $SE = 0.056$, $t = 4.954$, $p < 0.001$, 95% CI [0.160, 0.379]), supporting H4. Similarly, PU was positively associated with GAI-SRL ($\beta = 0.259$, $SE = 0.053$, $t = 4.845$, $p < 0.001$, 95% CI [0.150, 0.360]), supporting H5. Overall, the results support all five hypothesised direct relationships. Given the cross-sectional design, these coefficients indicate associations within the proposed model and should not be interpreted as causal effects.

Explanatory Power of the Model

The explanatory power of the structural model was assessed using the coefficient of determination (R^2). The R^2 values represent the proportion of variance in each endogenous construct explained by its predictor constructs. As presented in Table 9, the model explained 24.0% of the variance in ATT ($R^2 = 0.240$) and 32.9% of the variance in perceived GAI support for self-regulated learning (GAI-SRL) ($R^2 = 0.329$). Thus, the proposed model accounted for a larger proportion of variance in GAI-SRL than in ATT.

These values indicate that PEU and PU jointly explain part of the observed variation in students' attitudes toward GAI, while PEU, PU, and ATT collectively explain part of the variation in students' perceived GAI support for self-regulated learning. The R^2 values should be interpreted as explanatory evidence within the present model rather than as evidence of causal influence.

Endogenous Construct	R^2	Interpretation
ATT	0.240	24.0% of variance explained
GAI-SRL	0.329	32.9% of variance explained

Table 9: Coefficient of Determination (R^2) (Source: own data processing)

Effect Size (f^2)

Effect sizes (f^2) were examined to assess the contribution of each predictor to the explained variance of the corresponding endogenous construct. As presented in Table 10, all structural relationships showed small individual effect sizes.

The largest effect was observed for PEU on ATT ($f^2 = 0.117$), followed by PU on GAI-SRL ($f^2 = 0.090$), PEU on GAI-SRL ($f^2 = 0.075$), PU on ATT ($f^2 = 0.074$), and ATT on GAI-SRL ($f^2 = 0.045$). Although the corresponding structural paths were statistically significant, their individual contributions to the explained variance were relatively modest.

Structural Relationship	f^2	Interpretation
PEU → ATT	0.117	Small
PU → ATT	0.074	Small
PEU → GAI-SRL	0.075	Small
PU → GAI-SRL	0.090	Small
ATT → GAI-SRL	0.045	Small

Table 10: Effect Size (f^2) of the Structural Relationships (Source: own data processing)

Predictive Relevance (Q^2)

Blindfolding produced positive Q^2 values for ATT (0.186) and GAI-SRL (0.251), indicating predictive relevance for

both endogenous constructs (Table 11). The higher value for GAI-SRL indicates stronger predictive relevance within the estimated model.

Endogenous Construct	Q^2	Predictive Relevance
ATT	0.186	Yes
GAI-SRL	0.251	Yes

Table 11: Predictive Relevance (Q^2) of the Endogenous Constructs (Source: own data processing)

Out-of-Sample Predictive Assessment

PLSpredict showed positive Q^2_{predict} values for all indicators of ATT and GAI-SRL (Table 12). Values ranged from 0.087 to 0.185 for ATT and from 0.134 to 0.223 for GAI-SRL.

Because the present analysis does not report a comparison of

prediction errors with a linear-model benchmark, the positive Q^2_{predict} values are reported descriptively as evidence that the PLS predictions outperform the naïve benchmark used in the Q^2_{predict} calculation. No claim is made regarding low, medium, or high predictive power without comparison of prediction-error metrics against the linear-model benchmark.

Construct	Indicator	Q^2_{predict}
ATT	ATT1	0.176
ATT	ATT2	0.157
ATT	ATT3	0.140
ATT	ATT4	0.087
ATT	ATT5	0.185
GAI-SRL	SRL1	0.206
GAI-SRL	SRL2	0.210
GAI-SRL	SRL3	0.223
GAI-SRL	SRL4	0.176
GAI-SRL	SRL5	0.201
GAI-SRL	SRL6	0.200
GAI-SRL	SRL7	0.134

Table 12: PLSpredict Results (Source: own data processing)

Mediating Role of Students' Attitudes toward Generative AI

The mediating role of students' attitudes toward Generative AI was examined using bootstrapped specific indirect effects. As shown in Table 13, the indirect relationship between PEU and perceived GAI support for self-regulated learning through ATT was positive and statistically significant ($\beta = 0.051$, $SE = 0.017$, $t = 2.954$, $p = 0.003$, 95% $CI [0.021, 0.090]$). Because the corresponding direct relationship between PEU and GAI-SRL remained positive and statistically significant ($\beta = 0.278$, $p < 0.001$), the result indicates partial statistical mediation, supporting H6.

The indirect relationship between PU and GAI-SRL through ATT was also positive and statistically significant ($\beta = 0.065$, $SE = 0.023$, $t = 2.837$, $p = 0.005$, 95% $CI [0.025, 0.117]$). Because the direct relationship between PU and GAI-SRL remained significant ($\beta = 0.259$, $p < 0.001$), H7 was supported and the mediation was partial.

Both indirect effects were statistically significant, with confidence intervals excluding zero. Given the cross-sectional design, these mediation results indicate statistical mediation within the proposed model but do not establish temporal ordering or causality.

Hypothesis	Indirect Path	β	SE	t	p	95% CI	Direct (β)	Mediation Decision
H ₆	PEU → ATT → GAI-SRL	0.051	0.017	2.954	0.003	[0.021, 0.090]	0.278	Partial mediation
H ₇	PU → ATT → GAI-SRL	0.065	0.023	2.837	0.005	[0.025, 0.117]	0.259	Partial mediation

Table 13: Specific Indirect Effects and Statistical Mediation (Source: own data processing)

DISCUSSION

Interpretation of the Structural Relationships

The structural coefficients provide further insight into the relative strength of the relationships within the model. PU had the strongest association with ATT, whereas PEU showed the slightly stronger direct association with GAI-SRL. The indirect association through ATT was somewhat stronger for PU than for PEU. This pattern indicates that perceived usefulness may be particularly relevant to students' evaluations of GAI, while ease of use has a somewhat stronger direct association with perceived AI support for self-regulated learning.

The model explained 24.0% of the variance in ATT and 32.9% in GAI-SRL. However, all individual f^2 values were small (0.045–0.117), indicating that the observed relationships represent modest contributions within a broader set of factors. Thus, although the structural relationships were statistically significant, the contribution of each individual predictor was relatively modest. This finding supports a cautious interpretation in which perceived AI-supported self-regulated learning is likely associated with multiple factors beyond the constructs included in the present model.

PEU and PU in AI-Supported Self-Regulated Learning

The significant direct relationships of PEU and PU with GAI-SRL indicate that students' perceptions of the usability and educational value of GAI are statistically associated with how strongly they perceive AI as supporting their learning-regulation processes. PEU showed a positive association with GAI-SRL ($\beta = 0.278$), while PU also demonstrated a positive association ($\beta = 0.259$). These findings are consistent with the broader logic of the Technology Acceptance Model, in which perceived usefulness and perceived ease of use are important determinants of users' evaluations of technology. In the present learning-oriented model, however, these perceptions are examined in relation to students' perceived AI support for learning regulation rather than technology adoption alone.

The slightly stronger relationship between PEU and GAI-SRL suggests that usability may be particularly relevant when students incorporate GAI into learning activities. When AI tools are perceived as easy to access, understand, and operate, students may perceive fewer barriers to obtaining explanations, feedback, examples, or other forms of learning assistance. These functions may support processes associated with planning, monitoring, reflection, and strategy formulation. Previous research has similarly highlighted the potential of Generative AI to provide accessible explanations, personalised assistance, feedback, and flexible learning support (Banihashem et al., 2024; Kizilcec et al., 2024; Ge, 2026).

PU also demonstrated a significant positive association with GAI-SRL. Students who see clear educational value in GAI may also view it as useful for learning-related processes. Its value may lie in clarifying difficult concepts, generating alternative explanations, providing immediate assistance, and supporting academic tasks. These affordances may be relevant to self-regulated learning processes because students can potentially use them when deciding what to learn, how to approach a task, and how to evaluate their understanding (Xiaolei and Teng, 2024; Kim and Thille, 2026; Melanou and Beege, 2026).

The findings do not show that making GAI easier to use or more useful will improve students' self-regulated learning ability. They instead indicate an association between technology-related perceptions and perceived AI support for learning regulation. Because GAI-SRL captures perceived assistance rather than objectively observed regulatory competence, the findings should be understood as evidence concerning students' perceptions of AI-supported learning processes.

Students' Attitudes toward Generative AI as a Mediating Pathway

The mediation results indicate that ATT mediates the relationships between technology-related perceptions and perceived GAI support for self-regulated learning. The indirect relationship between PEU and GAI-SRL through ATT was significant ($\beta = 0.051$, $p = 0.003$), while the indirect relationship between PU and GAI-SRL through ATT was also significant ($\beta = 0.065$, $p = 0.005$). Because the corresponding direct relationships remained significant, both relationships represent partial statistical mediation.

The stronger indirect relationship observed for PU suggests that perceived usefulness may have a somewhat greater association with GAI-SRL when operating through students' attitudes toward GAI. This pattern is theoretically plausible because perceptions of educational usefulness may contribute to favourable evaluations of the technology, which are subsequently associated with stronger perceptions of AI-supported learning regulation. Nevertheless, the cross-sectional design does not establish that usefulness perceptions temporally precede attitudes or that attitudes subsequently cause changes in learning regulation. The mediation findings therefore represent statistical associations consistent with the proposed conceptual pathway rather than evidence of a causal process.

ATT represented a broad evaluative orientation toward GAI, including positive evaluation, confidence, enjoyment, and responsibility-oriented evaluation of AI-generated information. The positive association between ATT and GAI-SRL therefore suggests that students' overall orientation toward GAI is relevant to how they perceive AI support for learning-regulation processes. The responsibility-oriented component is particularly important in the context of Generative AI. Students may obtain explanations, suggestions, or information from AI systems without necessarily evaluating their accuracy or appropriateness. Previous research has emphasised the importance of critical evaluation, verification, and responsible engagement when Generative AI is used for educational purposes (Chauncey and McKenna, 2023; Asio, 2024; Tang et al., 2024; Akbay et al., 2025). A favourable attitude toward GAI should therefore not be equated with simple willingness to use the technology. It may also involve how students evaluate the role and appropriate use of AI within their learning processes.

Theoretical Contribution

Theoretically, the findings extend the application of the Technology Acceptance Model by connecting technology-related perceptions with a learning-oriented outcome. Previous applications of TAM to Generative AI have primarily focused on acceptance, intention to use, and technology-related evaluations (Dahri et al., 2024; Hwang et al., 2025; Mustofa et al., 2025). The findings extend the Technology Acceptance Model by linking technology-related perceptions with a learning-oriented outcome: perceived GAI support for self-regulated learning. Rather than treating acceptance as the endpoint, the model examines how attitudes toward GAI relate to students' perceptions of AI-supported learning processes. The partial mediation results indicate that ATT accounts for part, but not all, of the relationships between PEU, PU, and GAI-SRL. This extension should be interpreted as a conceptual and statistical contribution rather than evidence of a causal regulatory process.

Pedagogical and Institutional Implications

The results point to several considerations for AI-supported learning in higher education. The significant direct and indirect

relationships suggest that effective integration of Generative AI should address not only students' access to AI tools and perceptions of their usefulness, but also the ways in which students evaluate and incorporate AI into their learning processes. In particular, the partial mediation pattern suggests that favourable attitudes toward GAI are relevant to perceived AI support for self-regulated learning, although they do not represent the only pathway.

Instructors can ask students to establish learning goals before using GAI, formulate purposeful prompts, compare AI-generated responses with course materials, and explain how they verified the information. These activities keep students responsible for planning, monitoring, evaluating, and making learning decisions. This approach is consistent with the view of AI as a metacognitive and learning scaffold rather than an autonomous learning agent (Xia et al., 2023; Chiu, 2024; Weng et al., 2024). Such practices may help students use GAI as a support for learning rather than as a substitute for cognitive engagement (Li et al., 2025; Sardi et al., 2025b; Bueno-Vesga et al., 2026; Ding and Song, 2026).

At the institutional level, guidance can address transparency, verification, attribution, and appropriate AI use in learning and assessment. Faculty development should support instructors in translating these principles into learning activities and assessment practices.

Limitations and Directions for Future Research

Several limitations should be considered when interpreting the findings. First, the data came from a single higher education institution in Indonesia. Institutional and disciplinary contexts, as well as prior GAI experience, may shape students' perceptions and attitudes. Replication across institutions and disciplines is therefore needed to assess the transferability of the model.

Second, the cross-sectional design limits temporal and causal interpretation. Although mediation analysis identified significant indirect relationships, the data do not establish that PEU or PU precede ATT or that ATT subsequently changes learning regulation. Longitudinal and experimental studies are needed to examine these temporal relationships.

Third, the study relied on self-reported questionnaire data collected through a single online questionnaire. Although this approach is appropriate for examining perceptions and attitudes, it may not fully capture students' actual use of Generative AI or their observable learning-regulation processes. In particular, GAI-SRL reflects students' perceived support from Generative AI rather than direct evidence of self-regulated learning competence or observed behaviour. Future studies could combine questionnaires with AI-use logs, learning analytics, task-based assessments, observations, or interviews to capture actual AI use and learning-regulation processes.

Fourth, the use of a single self-report measurement occasion creates the potential for common-method variance. Procedural safeguards, including voluntary participation, informed consent, confidentiality, and de-identification of the analytical dataset, were applied; however, these procedures cannot completely eliminate common-method variance. The VIF

values were used to assess statistical collinearity, not as a formal test of common-method bias. Future research could reduce this concern through temporal separation of measurements, multiple data sources, behavioural indicators, or methodologically heterogeneous measures.

Finally, the model includes only four constructs and therefore does not capture all factors that may shape perceived AI support for self-regulated learning. AI literacy, prior experience, trust, perceived reliability, ethical awareness, task characteristics, and actual patterns of AI use may also be relevant. Future studies could incorporate these variables and use longitudinal, multi-institutional, and multimethod designs to examine how students' perceptions and attitudes develop over time.

CONCLUSION

This study examined the relationships among perceived ease of use, perceived usefulness, students' attitudes toward Generative AI, and perceived Generative AI support for self-regulated learning. PEU and PU were positively associated with ATT and GAI-SRL, while ATT was positively associated with GAI-SRL. Significant indirect relationships through ATT, together with significant direct relationships, indicated partial statistical mediation. The findings extend technology acceptance research by placing attitudes toward GAI within a learning-oriented framework focused on perceived AI support for self-regulated learning. They also highlight the importance of purposeful AI use, verification of AI-generated information, and continued student responsibility for learning decisions. Given the cross-sectional, self-report design and single-institution sample, the findings should be interpreted as statistical associations rather than causal effects. Future longitudinal, experimental, and multimethod research can further examine how these relationships develop over time.

ETHICAL AND INSTITUTIONAL CONSIDERATIONS

Formal approval from a separate institutional ethics committee or Institutional Review Board (IRB) was not obtained because a separate ethics review was not required under the institutional procedure applicable to this educational survey at the time of data collection. The study instead followed the institutional research-permission procedure of Universitas Negeri Padang, and written institutional permission to conduct the research was obtained through the applicable institutional procedure. Participation was voluntary, and respondents were informed about the purpose of the study before providing electronic informed consent. Respondents' names and student identification numbers (NIM) were collected solely to verify active undergraduate student status. These identifying data were not included in the analytical dataset and were deleted before statistical analysis. The analytical dataset was therefore de-identified, and the responses were treated confidentially.

ACKNOWLEDGMENT

The authors would like to thank the Institute for Research and Community Service (Lembaga Penelitian dan Pengabdian Masyarakat), Universitas Negeri Padang, for funding this work under Contract No. 1886/UN.35.15/LT/2026.

- Acosta-Enriquez, B. G., Arbulú Pérez Vargas, C. G., Huamani Jordan, O., Arbulú Ballesteros, M. A. and Paredes Morales, A. E. (2024) 'Exploring attitudes toward ChatGPT among college students: An empirical analysis of cognitive, affective, and behavioral components using path analysis', *Computers and Education: Artificial Intelligence*, Vol. 7, p. 100320. <https://doi.org/10.1016/j.caeai.2024.100320>
- Akbay, B., Karakoç Topal, Ö. and Bostan Sarioğlu, A. (2025) 'Artificial Intelligence Literacy and Anxiety Levels of Pre-service Science Teachers: A Mixed Method Study', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 18, No. 4, pp. 318–333. <https://doi.org/10.7160/eriesj.2025.180407>
- Asio, J. M. R. (2024) 'AI Literacy, Self-Efficacy, and Self-Competence among College Students: Variances and Interrelationships among Variables', *Malaysian Online Journal of Educational Sciences (MOJES)*, Vol. 12, No. 3, pp. 44–60. <https://doi.org/10.22452/aldad.vol12no3.4>
- Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J. and Drachsler, H. (2024) 'Feedback sources in essay writing: peer-generated or AI-generated feedback?', *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 1. <https://doi.org/10.1186/s41239-024-00455-4>
- Buabeng-Andoh, C. (2025) 'Investigating student-teachers' continuous intention to use mobile learning management system: the technology acceptance model and expectation confirmatory model', *Discover Education*, Vol. 4, No. 1, p. 76. <https://doi.org/10.1007/s44217-025-00447-0>
- Bueno-Vesga, J. A., Xu, X., Gu, Y., He, H., Li, S. and Duan, Y. (2026) 'The Role of Cognitive Load and Engagement in Students' Adoption of an AI-supported VR Training System', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 32, No. 5, pp. 4617–4625. <https://doi.org/10.1109/TVCG.2026.3680684>
- Chang, D. H., Lin, M. P. C., Hajian, S. and Wang, Q. Q. (2023) 'Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization', *Sustainability (Switzerland)*, Vol. 15, No. 17, p. 12921 <https://doi.org/10.3390/su151712921>
- Chatzopoulos, A. (2026) 'Using generative AI at the university: Do ChatGPT and other AI tools promote computational thinking?', in Kalogiannakis, M. and Sotiropoulos, D. (eds.) *Generative AI-enhanced science education from early childhood to higher education*, Hershey, PA: IGI Global Scientific Publishing, pp. 155–240. <https://doi.org/10.4018/979-8-3373-9074-1.ch005>
- Chauncey, S. A. and McKenna, H. P. (2023) 'A framework and exemplars for ethical and responsible use of AI Chatbot technology to support teaching and learning', *Computers and Education: Artificial Intelligence*, Vol. 5, p. 100182. <https://doi.org/10.1016/j.caeai.2023.100182>
- Chiu, T. K. F. (2024) 'A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: a case of ChatGPT', *Educational Technology Research and Development*, Vol. 72, No. 4, pp. 2401–2416. <https://doi.org/10.1007/s11423-024-10366-w>
- Choi, G. W., Kim, S. H., Lee, D. and Moon, J. (2024) 'Utilizing Generative AI for Instructional Design: Exploring Strengths, Weaknesses, Opportunities, and Threats', *TechTrends*, Vol. 68, No. 4, pp. 832–844. <https://doi.org/10.1007/s11528-024-00967-w>
- Creswell, J. W. (2012) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*, 4th edition, Boston, MA: Pearson.
- Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., Shutaleva, A. and Soomro, R. B. (2024) 'Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study', *Heliyon*, Vol. 10, No. 8, p. e29317. <https://doi.org/10.1016/j.heliyon.2024.e29317>
- Ding, J. and Song, Y. (2026) 'Descriptive writing using generative AI as a cognitive scaffold in the metaverse environment: university students' perceptions, learning engagement, and performance', *Educational Psychology*, Vol. 46, No. 1, pp. 183–207. <https://doi.org/10.1080/01443410.2025.2568189>
- EL Hosayny, O., Razkane, H. and Yeou, M. (2026) 'Metacognitive regulation and ChatGPT-assisted writing: introducing an interactive model for cultivating strategic engagement', *Discover Education*, Vol. 5, No. 1, p. 246. <https://doi.org/10.1007/s44217-026-01282-7>
- Fitriati, S. W. and Williyani, A. (2025) 'AI-enhanced self-regulated learning: EFL learners' prioritization and utilization in presentation skills development', *Journal of Pedagogical Research*, Vol. 9, No. 2, pp. 22–37. <https://doi.org/10.33902/JPR.202530647>
- Ge, Z.-G. (2026) 'The mechanisms of AI feedback's influence on college students' learning engagement in human-AI collaborative learning', *Innovations in Education and Teaching International*, pp. 1–14. <https://doi.org/10.1080/14703297.2026.2620543>
- Guan, L., Li, S. and Gu, M. M. (2024) 'AI in informal digital English learning: A meta-analysis of its effectiveness on proficiency, motivation, and self-regulation', *Computers and Education: Artificial Intelligence*, Vol. 7, p. 100323. <https://doi.org/10.1016/j.caeai.2024.100323>
- Hair, J. F., Howard, M. C. and Nitzl, C. (2020) 'Assessing measurement model quality in PLS-SEM using confirmatory composite analysis', *Journal of Business Research*, Vol. 109, pp. 101–110. <https://doi.org/10.1016/j.jbusres.2019.11.069>
- Hair, J. F., Risher, J. J., Sarstedt, M. and Ringle, C. M. (2019) 'When to use and how to report the results of PLS-SEM', In *European Business Review*, Vol. 31, No. 1, pp. 2–24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hou, C., Zhu, G., Liu, Y., Sudarshan, V., Chong, J. L. L., Zhang, F. Y., Tan, M. Y. H. and Ong, Y. S. (2026) 'The effects of critical thinking intervention on reliance behaviors, problem-solving quality, and creativity during human-Generative AI collaborative learning', *Computers and Education*, Vol. 247, p. 105576. <https://doi.org/10.1016/j.compedu.2026.105576>
- Hwang, M., Lee, E. and Lee, H. K. (2025) 'Exploring EFL Learners' Acceptance of ChatGPT: Application of the Extended Technology Acceptance Model', *English Teaching (South Korea)*, Vol. 80, No. 1, p. 45–69. <https://doi.org/10.15858/engtea.80.1.202503.45>
- Jiang, Y., Wang, P., Li, Q. and Li, Y. (2022) 'Students' Intention toward Self-Regulated Learning under Blended Learning Setting: PLS-SEM Approach', *Sustainability (Switzerland)*, Vol. 14, No. 16, p. 10140. <https://doi.org/10.3390/su141610140>
- Joo, S., Han, I. and Park, I. (2026) 'The effect of metacognitive prompts using generative AI on cognitive load, task performance, and self-efficacy in online self-regulated learning', *Educational Psychology*, pp. 1–24. <https://doi.org/10.1080/01443410.2026.2618717>

- Karan, B. and Chakma, C. (2025) 'Influence of higher education students' perceived behaviour on their artificial intelligence acceptance: an empirical investigation using technology acceptance model', *Journal of Applied Research in Higher Education*, Vol. 18, No. 3, pp. 745–760. <https://doi.org/10.1108/JARHE-11-2023-0535>
- Kim, J. H., Moldavan, A. M., Terry, R. and Massey, C. C. (2025) 'Enhancing student engagement with generative AI', in Kyei-Blankson, L. and Ntuli, E. (eds.) *Transformative AI practices for personalized learning strategies*, Hershey, PA: IGI Global Scientific Publishing, pp. 69–106. <https://doi.org/10.4018/979-8-3693-8744-3.ch004>
- Kim, Y. and Thille, C. (2026) 'Personalized learning experiences with AI', in Ives, K. S., Cini, M. and Schroeder, R. (eds.) *AI applications in online higher education administration: Strategies for maximizing returns and improving outcomes*, New York: Routledge, pp. 75–87. <https://doi.org/10.4324/9781003586388-10>
- Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S. and Radulescu, C. (2024) 'Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States', *Computers and Education: Artificial Intelligence*, Vol. 7, p. 100269. <https://doi.org/10.1016/j.caeai.2024.100269>
- Lai, J. W. (2024) 'Adapting Self-Regulated Learning in an Age of Generative Artificial Intelligence Chatbots', *Future Internet*, Vol. 16, No. 6, p. 218. <https://doi.org/10.3390/fi16060218>
- Lee, A. T., Ramasamy, R. K. and Subbarao, A. (2025) 'Understanding psychosocial barriers to healthcare technology adoption: A review of TAM Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology and UTAUT frameworks', *Healthcare*, Vol. 13, No. 3, p. 250. <https://doi.org/10.3390/healthcare13030250>
- Li, Y., He, X., Li, Y., Xiao, X., Li, R., Fang, J. and Zhou, R. (2025) 'Exploring dynamics in learners' cognitive engagement associated with learning activities using EEG microstate approach', *Thinking Skills and Creativity*, Vol. 58, p. 101911. <https://doi.org/10.1016/j.tsc.2025.101911>
- Madrid Akpovo, S. (2026) 'Generative AI in the intercultural learning process: An exploration of its usefulness in university virtual exchange programs', in Dervin, F. and R'boul, H. (eds.) *An introduction to AI and intercultural communication education*, New York: Routledge, pp. 64–76. <https://doi.org/10.4324/9781003675983-7>
- Mahniza, M., Sari, R. E., Suci, P. H., Saputra, I. and Putri, E. Y. (2024) 'AI-Driven Learning: Mediating and Moderating Dynamics in Self-Regulated Learning', *Journal of Educational Science and Technology (EST)*, Vol. 10, No. 3, p. 229. <https://doi.org/10.26858/est.v10i3.68254>
- Melanou, C. and Beege, M. (2026) 'Scaffolding Generative AI as a Tutor: A Quasi-Experimental Study of Learning Outcomes and Motivational, Cognitive and Metacognitive Processes', *Education Sciences*, Vol. 16, No. 4, p. 651. <https://doi.org/10.3390/educsci16040651>
- Molenaar, I. (2022) 'The concept of hybrid human-A regulation: Exemplifying how to support young learners' self-regulated learning', *Computers and Education: Artificial Intelligence*, Vol. 3, p. 100070. <https://doi.org/10.1016/j.caeai.2022.100070>
- Moya, S. and Camacho, M. (2024) 'Leveraging AI-powered mobile learning: A pedagogically informed framework', *Computers and Education: Artificial Intelligence*, Vol. 7, p. 100276. <https://doi.org/10.1016/j.caeai.2024.100276>
- Mustofa, R. H., Kuncoro, T. G., Atmono, D., Hermawan, H. D. and Sukirman. (2025) 'Extending the technology acceptance model: The role of subjective norms, ethics, and trust in AI tool adoption among students', *Computers and Education: Artificial Intelligence*, Vol. 8, p. 100379. <https://doi.org/10.1016/j.caeai.2025.100379>
- Nilson, L. B. (2023) *Creating self-regulated learners: Strategies to strengthen students' self-awareness and learning skills*, New York: Routledge. <https://doi.org/10.4324/9781003443803>
- Ozili, P. K. (2025) 'Technology impact model: a transition from the technology acceptance model', *AI and Society*, Vol. 40, No. 2, p. 1–3. <https://doi.org/10.1007/s00146-024-01896-1>
- Park, J. H., Lee, C. W. and Do, C. (2025) 'Examining Users' Acceptance Intention of Health Applications Based on the Technology Acceptance Model', *Healthcare (Switzerland)*, Vol. 13, No. 6, p. 596. <https://doi.org/10.3390/healthcare13060596>
- Polit, D. F. and Beck, C. T. (2006) 'The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations', *Research in Nursing and Health*, Vol. 29, No. 5, pp. 489–497. <https://doi.org/10.1002/nur.20147>
- Rahimi, A. R. and Faskhodi, A. A. (2026) 'Shaping AI-driven language learners' learning approaches, and motivation through computational thinking skills in computer-assisted language learning (CTCALL): A bisymmetric research design', *Smart Learning Environments*, Vol. 13, No. 1, p. 27. <https://doi.org/10.1186/s40561-026-00451-3>
- Runge, I., Hebibi, F. and Lazarides, R. (2025) 'Acceptance of Pre-Service Teachers Towards Artificial Intelligence (AI): The Role of AI-Related Teacher Training Courses and AI-TPACK Within the Technology Acceptance Model', *Education Sciences*, Vol. 15, No. 2, p. 167. <https://doi.org/10.3390/educsci15020167>
- Saqr, M., Cheng, R., López-Pernas, S. and Beck, E. D. (2024) 'Idiographic artificial intelligence to explain students' self-regulation: Toward precision education', *Learning and Individual Differences*, Vol. 114, p. 102499. <https://doi.org/10.1016/j.lindif.2024.102499>
- Sardi, J., Darmansyah, Candra, O., Yuliana, D. F., Habibullah, Yanto, D. T. P. and Eliza, F. (2025) 'How Generative AI Influences Students' Self-Regulated Learning and Critical Thinking Skills? A Systematic Review', *International Journal of Engineering Pedagogy*, Vol. 15, No. 1, pp. 94–108. <https://doi.org/10.3991/ijep.v15i1.53379>
- Scherer, R. (2025) 'Is the Technology Acceptance Model just old wine in new wineskins? Exploring issues for further model development', In *Journal of University Teaching and Learning Practice*, Vol. 21, No. 08. <https://doi.org/10.53761/var55x96>
- Schunk, D. H. and Zimmerman, B. J. (eds.) (2012) *Motivation and self-regulated learning: Theory, research, and applications*, New York: Routledge. <https://doi.org/10.4324/9780203831076>
- Shahzad, M. F., Xu, S. and Asif, M. (2025) 'Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model', *British Educational Research Journal*, Vol. 51, No. 2, pp. 489–513. <https://doi.org/10.1002/berj.4084>
- Svenningsson, J., Höst, G., Hultén, M. and Hallström, J. (2022) 'Students' attitudes toward technology: exploring the relationship among affective, cognitive and behavioral components of the attitude construct', *International Journal of Technology and Design Education*, Vol. 32, No. 3, pp. 1531–1551. <https://doi.org/10.1007/s10798-021-09657-7>

- Tang, T., Sha, J., Zhao, Y., Wang, S., Wang, Z. and Shen, S. (2024) 'Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria', *Thinking Skills and Creativity*, Vol. 53, p. 101607. <https://doi.org/10.1016/j.tsc.2024.101607>
- Tariq, M. U. (2025a) 'AI-driven personalized learning: Revolutionizing student engagement and teacher efficiency in Education 5.0', in Mobo, F. D. (ed.) *AI and emerging technologies for emergency response and smart cities*, Hershey, PA: IGI Global Scientific Publishing, pp. 75–96. <https://doi.org/10.4018/979-8-3693-9770-1.ch004>
- Tariq, M. U. (2025b) 'AI-enhanced project-based learning: Revolutionizing commerce education with generative AI', in ElSayary, A. (ed.) *Prompt engineering and generative AI applications for teaching and learning*, Hershey, PA: IGI Global Scientific Publishing, pp. 125–142. <https://doi.org/10.4018/979-8-3693-7332-3.ch008>
- van den Berg, G. (2024) 'Generative AI and Educators: Partnering in Using Open Digital Content for Transforming Education', *Open Praxis*, Vol. 16, No. 2, pp. 130–141. <https://doi.org/10.55982/openpraxis.16.2.640>
- Wang, C., Li, X. and Zou, B. (2025) 'Revisiting Integrated Model of Technology Acceptance Among the Generative AI-Powered Foreign Language Speaking Practice: Through the Lens of Positive Psychology and Intrinsic Motivation', *European Journal of Education*, Vol. 60, No. 1. <https://doi.org/10.1111/ejed.70054>
- Wang, Y. (2024) 'Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing', *System*, Vol. 126, p. 103505. <https://doi.org/10.1016/j.system.2024.103505>
- Weng, X., Xia, Q., Ahmad, Z. and Chiu, T. K. F. (2024) 'Personality traits for self-regulated learning with generative artificial intelligence: The case of ChatGPT', *Computers and Education: Artificial Intelligence*, Vol. 7, p. 100315. <https://doi.org/10.1016/j.caeai.2024.100315>
- Xia, Q., Chiu, T. K. F., Chai, C. S. and Xie, K. (2023) 'The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot', *British Journal of Educational Technology*, Vol. 54, No. 4, pp. 967–986. <https://doi.org/10.1111/bjet.13305>
- Xiaolei, S. and Teng, M. F. (2024) 'Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing', *Thinking Skills and Creativity*, Vol. 52, p. 101524. <https://doi.org/10.1016/j.tsc.2024.101524>
- Yanto, D. T. P., Ganefri, Hastuti, Candra, O., Kabatiah, M., Andrian and Zaswita, H. (2023) 'The Affecting Factors of Students' Attitudes Toward the Use of a Virtual Laboratory: A Study in Industrial Electrical Engineering', *International Journal of Online and Biomedical Engineering (IJOE)*, Vol. 19, No. 13, pp. 4–16. <https://doi.org/10.3991/ijoe.v19i13.41219>
- Zhai, C., Wibowo, S. and Li, L. D. (2024) 'The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review', *Smart Learning Environments*, Vol. 11, No. 1. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhang, L. and Chen, S. C. (2026) 'Towards human-AI collaborative learning: synergizing self-regulation and artificial-intelligence literacy', *Education and Information Technologies*, Vol. 31, No. 7, pp 1879–1907. <https://doi.org/10.1007/s10639-025-13880-3>
- Zhang, S., Zhao, X., Zhou, T. and Kim, J. H. (2024) 'Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior', *International Journal of Educational Technology in Higher Education*, Vol. 21, No. 1. <https://doi.org/10.1186/s41239-024-00467-0>
- Zimmerman, B. J. (2013) 'Theories of self-regulated learning and academic achievement: An overview and analysis', in Zimmerman, B. J. and Schunk, D. H. (eds.) *Self-regulated learning and academic achievement: Theoretical perspectives*, 2nd edition, New York: Routledge, pp. 1–35. <https://doi.org/10.4324/9781410601032-1>

THE ASSESSMENT–FEEDBACK LITERACY NEXUS: A LONGITUDINAL MIXED-METHODS STUDY OF RESPONSIBLE ASSESSMENT PRACTICE IN PRESERVICE TEACHER EDUCATION

Ronald Quileste

Xavier University-Ateneo de Cagayan
Philippines

✉ rquileste@xu.edu.ph

ABSTRACT

Assessment literacy and feedback literacy are often treated as separate competencies in teacher education, although both underpin responsible assessment practice. This two-wave, multi-cohort mixed-methods secondary analysis examined their parallel development among 250 preservice teachers at a private university in the southern Philippines. Five successive cohorts from 2020–2024 were matched from Assessment of Learning 1 (A1) to Assessment of Learning 2 (A2), typically completed in consecutive semesters. Quantitative analyses included paired course-phase comparisons, within-domain regression-based prediction, and programme comparisons. Qualitative interviews and written reflections from the ALiPP corpus were reanalysed thematically and integrated through joint displays. Standardized A2 assessment-literacy performance and feedback-literacy scores exceeded A1 scores, while A1 performance strongly predicted A2 outcomes within each domain. Programme-level patterns indicated needs for targeted support in smaller and specialized programmes. Qualitative findings showed shifts from procedural, compliance-oriented experiences toward rubric use, feedback dialogue, fairness, and professional responsibility. Because A1 and A2 assessment-literacy examinations covered sequential but partly different content and were not psychometrically equated, differences indicate standardized course-based performance rather than growth on an invariant scale. The Assessment-Feedback Literacy Nexus integrates structural and dialogic dimensions of responsible assessment practice as an explanatory framework.

KEYWORDS

Assessment literacy, Feedback literacy, Preservice teachers, Professional responsibility, Teacher education

HOW TO CITE

Quileste R. (2026) 'The Assessment–Feedback Literacy Nexus: a Longitudinal Mixed-Methods Study of Responsible Assessment Practice in Preservice Teacher Education', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 210–221. <http://dx.doi.org/10.7160/eriesj.2026.190303>

Article history

Received

October 24, 2025

Received in revised form

July 15, 2026

Accepted

August 27, 2026

Available on-line

September 30, 2026

Highlights

- Matched two-wave data distinguish within-student course-phase differences from between-cohort context.
- A2 standardized assessment performance and feedback-literacy scores were higher than A1 values.
- AFLN links assessment structure with feedback dialogue and professional responsibility.
- Programme patterns and qualitative evidence point to targeted literacy support.

INTRODUCTION

Assessment is central to teachers' professional work because it shapes how learning is planned, monitored, interpreted, communicated, and supported. For preservice teachers, assessment literacy is therefore not limited to constructing tests, computing scores, or preparing rubrics. It involves the capacity to design valid and fair assessment tasks, interpret evidence responsibly, communicate results clearly, and use assessment

information to improve learning (Popham, 2009; Brookhart, 2011; DePascale, 2018). This broader view is especially important in teacher education, where future teachers must learn to make assessment decisions that are technically sound, ethically defensible, and responsive to learners' needs.

Despite this importance, assessment literacy is often taught and studied as a procedural competence. In many teacher education contexts, students learn tables of specifications,

item analysis, validity, reliability, rubrics, and grading systems as separate course requirements. These competencies are necessary, but they may not be sufficient if preservice teachers are not also prepared to use assessment evidence in ways that support learner agency, dialogue, and improvement. Assessment can become compliance-oriented when it is disconnected from feedback and professional judgment. This concern is consistent with studies showing that assessment literacy remains a continuing challenge in teacher education and that more integrative approaches are needed to connect assessment knowledge with classroom decision-making and responsibility (Khalid, Latif and Yusof, 2021; Girgla et al., 2021; Gan and Lam, 2024).

Feedback literacy extends this discussion by focusing on how learners and teachers understand, interpret, and use feedback. Carless and Boud (2018) define student feedback literacy as the understandings, capacities, and dispositions needed to make sense of feedback and use it for improvement. From the teacher's perspective, feedback literacy involves designing feedback processes that learners can understand, discuss, and act on (Boud and Dawson, 2023; Dawson et al., 2023). Feedback is therefore not simply a comment added after assessment. It is part of how assessment evidence becomes meaningful, usable, and educationally responsible.

However, assessment literacy and feedback literacy are still commonly treated as separate areas of scholarship. Assessment literacy research tends to emphasise evidence, criteria, validity, reliability, and judgment, while feedback literacy research often emphasizes uptake, dialogue, agency, and evaluative judgment. These emphases are complementary, yet the relationship between the two literacies remains underexamined. This separation creates a conceptual gap: preservice teachers may learn how to produce assessment evidence without fully learning how that evidence should be communicated, interpreted, and used to support learning. Conversely, feedback may be discussed as dialogue without sufficient attention to the quality and fairness of the assessment evidence on which it is based.

The need to connect assessment and feedback is supported by emerging work on evaluative judgment and feedback processes. Deneen and Hoo (2021) argued that assessment literacy is connected with self-evaluation and peer feedback, suggesting that assessment competence develops through interaction with evaluative processes. Panadero and Lipnevich (2022) emphasised that feedback should be understood as a system of processes rather than as a single message. Nieminen and Carless (2022) further argued that feedback literacy must be examined critically because feedback is shaped by agency, power, context, and access. These perspectives suggest that assessment and feedback should be viewed as interdependent practices within teacher education.

This interdependence is especially important when assessment is understood as a matter of professional responsibility. Responsible assessment requires attention not only to technical quality but also to fairness, transparency, consequences, and ethical use of evidence. Messick (1995) emphasised that validity includes the meaning and consequences of assessment interpretation and use. The *Standards for Educational and*

Psychological Testing similarly foregrounds validity, fairness, and appropriate interpretation as central to responsible assessment practice (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014). In teacher education, this means that preservice teachers must learn to ask not only whether an assessment tool works, but also whether it supports learners fairly, communicates expectations clearly, and informs better teaching decisions.

The same concern is visible in the Philippine teacher education context. The Commission on Higher Education requires teacher education programmes to prepare graduates who can demonstrate professional and pedagogical competencies across elementary education, secondary education, early childhood education, and special needs education (Commission on Higher Education, 2012, 2017a, 2017b, 2017c, 2017d). The Philippine Professional Standards for Teachers also emphasises assessment, reporting, learner diversity, and the use of evidence to support learning (Department of Education, 2017). These expectations make assessment and feedback literacy essential to preservice teacher preparation. Yet they also create the risk that assessment becomes framed mainly as compliance with outcomes, rubrics, and documentation rather than as a dialogic and ethical practice.

To address this gap, this study proposes the Assessment–Feedback Literacy Nexus (AFLN) as an integrative framework for preservice teacher education. AFLN positions assessment literacy as the structural foundation of responsible assessment practice and feedback literacy as the dialogic process through which assessment evidence becomes meaningful for learning. Assessment literacy includes knowledge and practice related to learning outcomes, test construction, validity, reliability, rubrics, scoring, grading, and interpretation. Feedback literacy includes understanding criteria, interpreting comments, engaging in dialogue, using feedback for revision, and supporting learner agency. Their connection appears when preservice teachers learn to link assessment evidence with feedback action. For example, a rubric is not only a scoring tool; it is also a feedback tool that communicates expectations and guides improvement. Item analysis is not only a statistical procedure; it can inform instructional decisions and feedback to learners.

Figure 1 summarizes the proposed relationship among assessment literacy, feedback literacy, contextual factors, and responsible assessment practice.

AFLN does not assume that one literacy simply causes the other. Rather, it proposes that assessment and feedback literacy develop through repeated cycles of design, interpretation, communication, revision, and reflection. In this sense, AFLN is both conceptual and pedagogical. Conceptually, it explains why assessment and feedback should be studied together. Pedagogically, it suggests that teacher education courses should provide repeated opportunities for preservice teachers to design assessment tools, receive and use feedback, revise outputs, discuss criteria, and apply assessment practices in authentic or simulated teaching contexts.

This framework is also informed by prior local work at a private university in the southern Philippines. The ALiPP framework

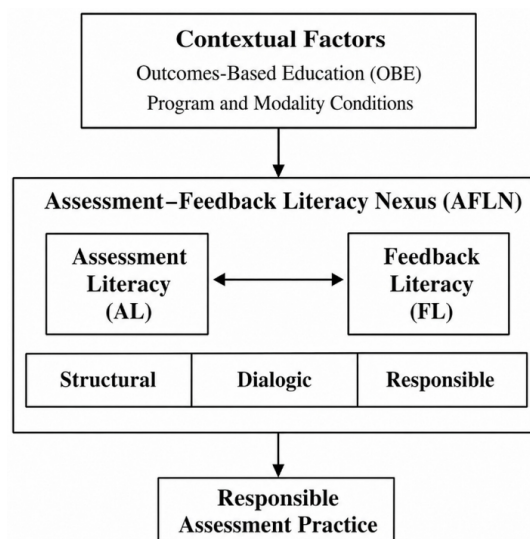


Figure 1: Assessment-Feedback Literacy Nexus framework, 2020-2024 (source: own elaboration)

examined assessment literacy by integrating students' assessment performance and perceptions of assessment learning (Author, 2026). The ASPIRE framework explored how inquiry-based teacher education can support assessment literacy development through active engagement, reflection, and evidence-informed practice (Author, Author and Author, 2025). These frameworks provide useful contextual grounding, but they are not treated as the sole foundation of AFLN. Instead, they are positioned as context-specific precursors. AFLN extends this work by explicitly integrating feedback literacy with assessment literacy and by interpreting both as dimensions of professional responsibility.

The present study reanalyses historical longitudinal mixed-methods data from 250 preservice teachers at a private university in the southern Philippines across five successive cohorts from 2020 to 2024. It examines matched A1-A2 course-phase differences, within-domain longitudinal continuity, programme-level patterns, and qualitative evidence linking assessment, feedback, fairness, and professional responsibility. The study addresses the following questions:

1. How did matched preservice teachers' standardized course-based assessment-literacy performance and overall feedback-literacy scores differ between A1 and A2 across the sequential course sequence?
2. To what extent did A1 scores predict A2 scores within each literacy domain, and what programme-level patterns were observed across assessment literacy and feedback literacy?
3. How do qualitative experiences across the two course phases and successive cohorts explain connections among assessment, feedback, fairness, and professional responsibility?

By addressing these questions, the study contributes to teacher education research by offering a more integrated account of how preservice teachers learn to design, interpret, communicate, and use assessment evidence responsibly. It also responds to the need for research that positions assessment and feedback not as separate technical competencies, but as connected practices of professional responsibility in teacher education.

MATERIALS AND METHODS

Research Design

This study used a two-wave, multi-cohort convergent parallel mixed-methods secondary-analysis design to examine assessment literacy and feedback literacy among preservice teachers at a private university in the southern Philippines. The 2020-2024 window represents five successive cohorts, not one cohort observed continuously for four years. Within each cohort, the same students were matched from Assessment of Learning 1 (A1) to Assessment of Learning 2 (A2), which were normally taken in consecutive semesters; the typical interval between the two measurement phases was therefore approximately one academic semester. Quantitative within-student analyses compared matched A1 and A2 records, whereas cohort year and instructional modality were treated as between-cohort contextual descriptors. The qualitative strand reanalysed archived interview and reflection materials concerning assessment design, feedback use, fairness, rubric development, and professional responsibility. The two strands were analysed separately and integrated during interpretation through joint displays and meta-inferences (Creswell and Creswell, 2018). Qualitative references to A1/A2 or earlier/ later phases are reported as phase- or cohort-level patterns and are not interpreted as individual four-year trajectories unless the same participant could be linked across phases.

Relationship to Prior Studies and Originality

The present article is a secondary reanalysis and extension of the ALiPP source dataset (Quileste, 2026). The 250 matched A1/A2 assessment-literacy records and the original interview/ reflection corpus overlap with ALiPP; accordingly, AL-only descriptive, A1-to-A2 prediction, and programme-level results reported here are identified as previously reported contextual anchors rather than as new findings. The FL_A1 and FL_A2 variables and their longitudinal analyses were not reported in the ALiPP publication and constitute the new quantitative component of the present study. The principal novel contribution is the integration of those feedback-literacy indicators with

the assessment-literacy evidence and the development of AFLN as a responsibility-oriented explanatory framework. ASPIRE (Quileste, Estacio and Pagara, 2025) was a separate mixed-methods study involving 29 preservice teachers under a separate institutional ethics approval; no ASPIRE participant-level data, quantitative results, interviews, or reflections were pooled into or reanalysed in the present dataset. ASPIRE is used only as conceptual and contextual grounding.

Research setting

The study was conducted in a teacher education unit of a private university in the southern Philippines. The research context involved two sequential undergraduate professional education courses: Assessment of Learning 1 (A1) and Assessment of Learning 2 (A2). A1 introduced foundational competencies such as outcomes alignment, table of specifications development, test construction, validity, reliability, and interpretation of assessment results. A2 extended these competencies to applied assessment tasks, including rubric development, performance assessment, grading systems, item analysis, and feedback practices. A1 was normally completed in one semester and A2 in the succeeding semester. The study covered five successive cohorts from 2020 through 2024, including cohorts exposed to fully online delivery during the pandemic and later cohorts returning to face-to-face learning. These cohort differences are treated as contextual conditions rather than as four-year within-person follow-up.

Participants and Sampling

The quantitative sample comprised 250 preservice teachers who completed both A1 and A2 and whose records could be matched across the two course phases. They came from four teacher education programmes: Bachelor of Elementary Education (BEED), Bachelor of Secondary Education (BSED), Bachelor of Special Needs Education (BSNED), and Bachelor of Early Childhood Education (BECED). The programme distribution was BECED $n = 7$, BEED $n = 107$, BSED $n = 88$, and BSNED $n = 48$. Each participant contributed one matched A1-A2 pair within a single cohort year; no participant was treated as contributing observations across multiple cohort years.

Purposive sampling was used because the analysis required complete paired longitudinal records. Inclusion required enrolment in both sequential assessment courses, a matched student identifier across A1 and A2, complete assessment-literacy scores, corresponding FL_A1 and FL_A2 overall scores, and membership in one of the 2020-2024 cohorts. Incomplete or unmatched records were excluded from the quantitative analysis.

The secondary qualitative reanalysis drew on the ALiPP source corpus of semi-structured interviews and written reflections. The interview sample comprised 20 participants, four from each cohort, purposively selected to represent variation in programme and gender. Interviews lasted approximately 30-45 minutes and were conducted via Zoom for the 2020-2021 cohorts and face-to-face for the 2022-2024 cohorts. Written reflections were collected after A1 and A2 as course requirements and were used to triangulate the interview themes.

In the de-identified secondary archive, the reflection materials had been consolidated during the source-study analysis rather than retained as a reliably countable set of individual documents; therefore, this revision does not infer an exact reflection-document count. No standalone course artifacts are treated as a separate qualitative dataset in the present analysis.

Data Sources and Instruments

The study used four data sources: standardized course-based assessment-literacy scores, overall feedback-literacy scores, qualitative interviews and written reflections, and programme/cohort contextual variables.

Assessment literacy was represented by comprehensive examination scores from A1 and A2. Each course used a 60-item multiple-choice comprehensive examination administered as a final assessment. Raw scores ranged from 0 to 60 and were standardized to a 0-100 scale for analysis. The A1 examination sampled foundational assessment competencies, including test construction, validity principles, and table of specifications development; the A2 examination sampled more advanced and applied competencies, including rubric design, item analysis, and KR-20 calculations (Quileste, 2026). The two examinations were separately developed for their respective courses and did not use a psychometric equating or common-item linking procedure. Accordingly, the shared 0-100 transformation facilitates descriptive comparison but does not establish measurement invariance or psychometric equivalence. A1-A2 differences are therefore interpreted as differences in standardized course-based assessment performance across sequential phases, not as gain on an invariant assessment-literacy scale.

Feedback literacy was represented by FL_A1 and FL_A2 overall scores from an archived, locally adapted 20-item Feedback Literacy Scale used at the two course phases. Items were rated on a five-point agreement scale from 1 (strongly disagree) to 5 (strongly agree), and the secondary dataset retained an overall phase score expressed on the same 1-5 metric, with higher values indicating stronger feedback literacy. Archived examples included items addressing evaluation of the usefulness of feedback and recognition of emotional barriers to responding to feedback. The local instrument was conceptually informed by the feedback-literacy literature (Carless and Boud, 2018; Dawson et al., 2023) but was not a direct administration of the later 24-item Feedback Literacy Behaviour Scale reported by Dawson et al. (2023). FL_A1 corresponded to the first-course phase and FL_A2 to the second-course phase. Because only overall phase scores, rather than complete item- or subscale-level responses, were retained in the secondary dataset, the present analysis uses total feedback-literacy scores and makes no claims about specific subdimensions such as emotion management or appraisal.

Qualitative data consisted of semi-structured interview transcripts and post-A1/post-A2 written reflections from the ALiPP source corpus. These materials addressed students' experiences with assessment design, rubric development, feedback, peer review, instructor guidance, online and face-to-face learning conditions, fairness, and application to future teaching.

Programme, cohort year, instructional modality, and gender were retained as descriptive contextual variables. Programme comparisons were examined directly; cohort year and modality were used to contextualise qualitative and descriptive patterns rather than as repeated observations from the same individuals across years.

Validity and Reliability

Several procedures were used to document measurement quality and to delimit the claims that can be made from the secondary dataset. For assessment literacy, both 60-item course examinations were faculty-developed and aligned with the intended outcomes and tables of specifications of A1 and A2. The source study reported KR-20 coefficients above .75 for both examinations (Quileste, 2026). This supports acceptable internal consistency within each course examination, but it does not establish equivalence between the two examinations. Because the forms sampled partly different content and were not psychometrically equated, the study does not interpret the A1-A2 assessment-literacy difference as invariant-scale growth.

For feedback literacy, archived source-study documentation described a locally adapted 20-item, five-point scale reviewed by five faculty experts, with a reported content-validity ratio above .80 and an overall Cronbach's alpha of .85. These values are reported as legacy validation information from the source documentation. Phase-specific alpha coefficients for FL_A1 and FL_A2 could not be recomputed because item-level responses were not retained in the de-identified secondary dataset. This constraint is reported explicitly as a limitation, and the analysis is restricted to the overall FL_A1 and FL_A2 scores.

The quantitative dataset was screened for duplicate records, missing values, inconsistent identifiers, and out-of-range values. Only complete matched cases were retained. Descriptive statistics and assumption checks were reviewed before inferential analysis. Shapiro-Wilk tests were used as an initial normality check, but interpretation also considered distributional shape, sample size, and the robustness of the planned analyses.

The qualitative approach is described in this revision as codebook-oriented thematic analysis rather than reflexive thematic analysis. The six-phase sequence of familiarisation, initial coding, theme construction, theme review, theme definition, and reporting provided the procedural structure (Braun and Clarke, 2006). In the source analysis, the primary researcher coded the corpus and a second researcher independently reviewed 20% of the material; the archived analysis reported 85% agreement (Cohen's kappa = .82), with discrepancies discussed to refine the shared coding framework. Member-checking with five interview participants was also used in the source study. In the present secondary reanalysis, interview and reflection codes were compared within a common thematic matrix so that source type, course phase, cohort context, and divergent cases could be considered together. The agreement check is reported as a quality-control feature of the codebook-oriented approach, not as a criterion for reflexive thematic analysis.

Mixed-methods validity was addressed through integration after the separate analyses. Quantitative patterns were compared with qualitative themes through joint displays, with attention to convergence, complementarity, and divergence. Because the revised quantitative analysis does not report an ambiguously specified cross-domain correlation or a pooled repeated-observation regression, AFLN is supported through the convergence of parallel two-wave score patterns, within-domain continuity, programme comparisons, and qualitative evidence rather than through a claim of direct causal coupling between assessment literacy and feedback literacy.

Data Preparation and Screening

Before analysis, student records were matched across A1 and A2 using student identifiers. Duplicate entries were reviewed and removed when they represented repeated identifiers rather than distinct participants. The final analytical dataset contained 250 matched students, each with one AL_A1, AL_A2, FL_A1, and FL_A2 record. This structure yields 250 paired cases, not 500 independent participants or observations for ordinary least-squares modelling.

AL_A1 and AL_A2 represented standardized course-based assessment performance at the first and second course phases; FL_A1 and FL_A2 represented overall feedback-literacy scores at the corresponding phases. For feedback literacy, the within-student phase difference can be interpreted on the same 1-5 score metric. For assessment literacy, no construct-level gain score is interpreted because A1 and A2 were separately developed, non-equated course examinations. Because feedback-literacy scores were aggregated on a five-point scale, they were treated as approximately continuous for parametric analysis. This decision was based on the use of overall scale scores rather than single Likert-type items and the sample size; the limitation is acknowledged in the Discussion.

Quantitative Data Analysis

Quantitative analyses were conducted in four steps. First, descriptive statistics were computed for AL_A1, AL_A2, FL_A1, and FL_A2 to summarize the score patterns across the two course phases.

Second, paired-samples t-tests were used to compare matched A1 and A2 scores. For feedback literacy, the same five-point total-score metric was used at both phases, so the test evaluates a within-student phase difference on a common score metric. For assessment literacy, the paired test is retained only as a comparison of standardized course-phase performance; because the A1 and A2 examinations were not psychometrically equated, the resulting difference and Cohen's *d* are not interpreted as evidence of growth on an invariant assessment-literacy measure.

Third, separate within-domain regression models examined longitudinal continuity across the two course phases. AL_A2 was regressed on AL_A1, and FL_A2 was regressed on FL_A1. These models test whether first-course performance predicts second-course performance within the same literacy domain; they do not test cross-domain causality.

Fourth, programme differences were examined separately for AL_A1, AL_A2, FL_A1, and FL_A2 using Welch's ANOVA because of the markedly unequal subgroup sizes. Post hoc comparisons were conducted where significant omnibus effects were found. No pooled 500-observation ordinary least-squares regression is reported in the revised analysis because each student contributes repeated A1/A2 observations and those observations are not independent.

Qualitative Data Analysis

Qualitative data were analysed using a codebook-oriented thematic approach structured around Braun and Clarke's six phases (Braun and Clarke, 2006). The primary researcher conducted the full reanalysis of the de-identified interview/reflection corpus using a codebook focused on students' experiences of assessment and feedback across the two course phases. Initial codes included conceptual overload, difficulty applying assessment concepts, feedback gaps, statistical anxiety, modality challenges, instructor modelling, LMS resources, peer collaboration, digital tool use, practicum integration, fairness and validity awareness, rubrics as learning tools, and feedback as dialogue. A second researcher independently reviewed 20% of the source material in the original ALiPP analysis; discrepancies were discussed to refine coding decisions, and the archived analysis reported Cohen's kappa = .82. Interview and reflection evidence was then compared across source type, course phase, and cohort context before themes were finalized. Because the design comprises successive cohorts, statements about A1/A2 or pandemic/post-pandemic patterns are treated as corpus-level or cohort-level patterns rather than as evidence that the same interview participants were followed continuously for four years.

Mixed-methods Integration

Integration occurred after the separate quantitative and qualitative analyses. Joint displays were used to compare statistical patterns with qualitative themes. Integration was organised around convergence, complementarity, and divergence.

Convergence was identified when quantitative and qualitative findings supported the same interpretation. Complementarity was identified when one strand expanded or explained the other. Divergence was identified when one strand complicated the other. Meta-inferences were then developed by weighing both strands together.

This procedure allowed the study to examine whether quantitative course-phase patterns converged with students' accounts of fairness, feedback dialogue, assessment responsibility, and professional meaning-making. Integration was interpretive: it did not assume that parallel quantitative patterns demonstrate a direct causal relation between the two literacies.

Researcher Positionality and Ethical Considerations

The study involved an insider research context because the researcher was familiar with the assessment courses and teacher education programme. This positionality provided contextual knowledge of course design and assessment practices, but it also created potential power, social-desirability, and interpretive biases. These risks are therefore treated explicitly rather than assumed to be eliminated.

Participant-generated responses in the source materials were collected through course-based Google Forms that included electronic consent language authorizing the use of responses for research purposes and assuring the confidentiality of personal identifiers. Historical assessment records originated from routine instructional processes. For the present secondary analysis, identifiers used for record matching were removed before the analytical dataset was prepared, quantitative findings are reported in aggregate, and qualitative excerpts use participant codes. No new participant recruitment or data collection was undertaken for the present study, and the analysis did not affect students' grades or academic standing. The researcher's insider/instructor role is acknowledged as a limitation.

The present manuscript reports the consent and data-governance procedures directly documented for the records used in this secondary analysis. The separate ASPIRE study (Quileste, Estacio and Pagara, 2025), cited only as conceptual and contextual grounding, received institutional ethics approval under XU REC Package No. SOE-202400222; that approval is not represented as covering the ALiPP/AFLN dataset, and no ASPIRE participant data are included in the present analysis. Archival course records are reported only in aggregate, qualitative materials are pseudonymised, and subgroup reporting is intentionally cautious where small programmes could increase re-identification risk.

RESULTS

The results present the quantitative and qualitative findings for the matched two-wave dataset. The quantitative strand is organised around A1-A2 course-phase score differences, within-domain longitudinal prediction, and programme comparisons. The qualitative strand presents the secondary thematic reanalysis of interviews and written reflections, followed by mixed-methods integration. Cohort year and modality are used only as contextual descriptors in the qualitative interpretation.

Quantitative Results

A1-A2 Score Differences within Matched Students

Table 1 presents the matched A1 and A2 scores for standardized assessment-literacy course performance and overall feedback-literacy scores.

Measure	<i>N</i>	A1 Mean (SD)	A2 Mean (SD)	<i>t</i>	<i>df</i>	<i>p</i>	Cohen's <i>d</i>
Assessment literacy score	250	49.2 (7.23)	52.0 (6.92)	-21.3	249	< .001	1.35
Feedback literacy score	250	3.4 (0.70)	4.1 (0.60)	-15.2	249	< .001	0.96

Table 1: A1 and A2 scores for matched students, 2020-2024 (source: secondary analysis; AL statistics previously reported in Author, 2026)

Standardized assessment-literacy course performance was higher at A2 ($M = 52.0$, $SD = 6.92$) than at A1 ($M = 49.2$, $SD = 7.23$), $t(249) = -21.3$, $p < .001$, $d = 1.35$. Because the two examinations sampled partly different content and were not equated, this result is interpreted as a statistically detectable course-phase score difference, not as psychometrically invariant assessment-literacy growth. Overall feedback-literacy scores,

which were recorded on the same five-point metric at both phases, were also higher at A2 ($M = 4.1$, $SD = 0.60$) than at A1 ($M = 3.4$, $SD = 0.70$), $t(249) = -15.2$, $p < .001$, $d = 0.96$.

Within-domain Longitudinal Prediction

Separate regressions examined whether A1 scores predicted A2 scores within each literacy domain. Table 2 presents the results.

Model	Predictor	Estimate	SE	t	p	R ²
AL model	Intercept	6.78	0.85	7.97	<.001	.921
AL model	A1 assessment literacy	0.92	0.02	53.68	<.001	.921
FL model	Intercept	0.92	0.12	7.67	<.001	.884
FL model	A1 feedback literacy	0.88	0.03	29.33	<.001	.884

Table 2: Within-domain A1 scores predicting A2 scores, 2020-2024 (source: secondary analysis; AL model previously reported in Author, 2026)

A1 standardized assessment-literacy performance strongly predicted A2 standardized assessment-literacy performance, explaining 92.1% of variance in the later course score ($R^2 = .921$). A1 overall feedback-literacy score likewise strongly predicted A2 overall feedback-literacy score ($R^2 = .884$). These results indicate strong within-domain continuity across the two sequential course phases. For assessment literacy, the result should be

understood as prediction across related but non-equated course assessments rather than as stability on an identical test form.

Programme Differences

Programme differences were examined separately at A1 and A2 for the four teacher education programmes. Table 3 presents the descriptive statistics.

Programme	N	A1 AL Mean (SD)	A2 AL Mean (SD)	A1 FL Mean (SD)	A2 FL Mean (SD)
BECED	7	39.6 (6.52)	43.5 (6.94)	3.1 (0.60)	3.6 (0.50)
BEED	107	51.5 (5.68)	54.1 (5.54)	3.6 (0.70)	4.2 (0.60)
BSED	88	50.7 (6.44)	53.1 (6.05)	3.5 (0.70)	4.3 (0.60)
BSNED	48	42.9 (7.23)	46.4 (7.48)	3.2 (0.80)	3.8 (0.70)

Table 3: Assessment- and feedback-literacy scores by programme, 2020-2024 (source: secondary analysis; AL statistics previously reported in Author, 2026)

Welch's ANOVA showed significant programme differences for A1 assessment-literacy performance, $F(3, 27.5) = 22.8$, $p < .001$; A2 assessment-literacy performance, $F(3, 27.3) = 16.9$, $p < .001$; A1 feedback literacy, $F(3, 28.1) = 19.4$, $p < .001$; and A2 feedback literacy, $F(3, 27.9) = 15.2$, $p < .001$. Post hoc comparisons indicated that BEED and BSED students generally had higher observed scores than BECED and BSNED students. These are unadjusted programme comparisons and should not be interpreted as evidence of programme quality or causal programme effects. In particular, the BECED subgroup was small ($n = 7$), so the results are used to identify possible support needs rather than to rank programmes.

Qualitative Results

The qualitative findings derive from a secondary thematic reanalysis of the ALiPP interview/reflection corpus. They explain how students described assessment and feedback across course phases and cohort contexts. Because the interview sample was not a single panel followed continuously from 2020 to 2024, temporal statements below refer to phase- or cohort-level patterns in the corpus rather than to four-year within-person change.

Challenges and Supports in Literacy Co-development

A detailed thematic summary of students' difficulties and sources of support is presented in Appendix A.

A1-phase and online-cohort materials frequently described assessment as technically demanding. Concepts such as validity, reliability, tables of specifications, item analysis, and KR-20 were often described as difficult, and students reported uncertainty about whether assessment tools and rubrics were valid or usable in practice. Feedback gaps, statistical anxiety, and limited opportunities for immediate clarification were recurring challenges, particularly in the 2020-2021 online context.

Across the corpus, students identified LMS resources, templates, instructor modelling, peer collaboration, digital tools, and practicum experiences as supports for interpreting assessment tasks and using feedback. A2-phase and later face-to-face cohort materials more often emphasized rubric use, dialogue, revision, and professional application. These patterns are interpreted as contextual and phase-level tendencies rather than as proof that every individual participant followed the same developmental sequence.

Perceived Relevance and Responsible Application

Table 4 presents themes related to students' understanding of assessment and feedback as future teaching responsibilities.

The qualitative corpus showed a phase- and cohort-level contrast between procedural/compliance-oriented accounts and accounts emphasizing professional responsibility. A1-

Theme	Description	Illustrative quote	Pattern	AFLN interpretation
Developing real skills	Shift from completing requirements to taking responsibility for assessment work	"Assessment is a skill I'll need every day." (S3, 2022, F2F)	Evident from 2022	AL and FL became connected to professional competence
Understanding fairness in testing	Recognition that unclear or poorly designed assessment can affect learners	"One wrong test item confused my kids." (S9, 2024, F2F)	Evident from 2021	Feedback experiences deepened fairness awareness
Using rubrics as helpful guides	Rubrics were understood as learning and communication tools, not only grading tools	"Students liked knowing how they were graded." (S2, 2023, F2F)	Strong after 2022	Rubrics joined structural clarity with feedback dialogue
Feedback through conversation	Feedback was increasingly viewed as a two-way process	"I learned to ask good questions." (S7, 2024, F2F)	Strong in F2F cohorts	Dialogic feedback supported reflective assessment practice
Limits of online exams	Online assessments were seen as limited in capturing thinking and applied competence	"Rubrics showed more about how I think." (S4, 2021, Zoom)	Strong in 2020-2021	Modality affected how assessment and feedback were experienced

Table 4: Responsible application themes across phases and cohorts, 2020-2024 (source: secondary thematic reanalysis)

phase materials more often emphasized difficulty, uncertainty, and completion of course requirements, whereas A2-phase and later-cohort materials more often framed assessment as a skill needed for teaching. This pattern is not treated as a four-year within-person trajectory; rather, it is evidence that different course phases and instructional contexts were associated with different emphases in participants' accounts.

Fairness emerged as an important ethical theme. Students recognised that assessment quality was not only a matter of technical correctness but also of learner protection and professional accountability. Rubrics were increasingly

understood as tools for communicating expectations and supporting learning. Feedback was also described as a conversation that helped students revise, clarify, and make better pedagogical decisions.

Mixed-methods Integration

The quantitative and qualitative results were integrated to examine how course-phase patterns in assessment literacy and feedback literacy corresponded with students' accounts of assessment practice. Table 5 presents the joint display of quantitative patterns, qualitative themes, and AFLN interpretations.

Quantitative pattern	Qualitative theme	Integrated AFLN interpretation
A2 standardized AL performance exceeded A1 (paired standardized course-score difference, $d = 1.35$); FL total scores were higher at A2 than A1 ($d = 0.96$).	Practicum integration, rubric use, and development of real skills	Both domains showed parallel A1-A2 phase patterns, while the non-equated AL examinations require cautious interpretation of the AL difference.
A1 AL strongly predicted A2 AL ($R^2 = .921$); A1 FL strongly predicted A2 FL ($R^2 = .884$).	LMS resources, instructor modelling, and peer collaboration	Strong within-domain continuity is consistent with cumulative course learning, but it does not demonstrate cross-domain causality.
Programme differences were significant across AL and FL phase scores.	Feedback gaps and uneven learning support	Programme patterns identify possible differences in learning conditions and support needs; they are not interpreted as causal programme effects.
The quantitative revision does not report the previously ambiguous cross-domain correlation or the 500-observation pooled OLS model.	Students nevertheless described assessment and feedback as connected through rubrics, dialogue, revision, and fairness.	AFLN is supported as an explanatory mixed-methods framework through convergence of evidence, not as a validated causal model.

Table 5: Integrated AFLN findings, 2020-2024 (source: secondary analysis and thematic reanalysis)

The integrated findings are consistent with AFLN as an explanatory framework, but they do not establish direct cross-domain causality. Quantitatively, A2 scores were higher than A1 scores in both domains, first-phase scores strongly predicted second-phase scores within each domain, and programme differences were evident. Qualitatively, participants described how feedback, rubrics, peer collaboration, instructor modelling, and practicum experiences shaped their understanding of assessment as a fair, dialogic, and professionally responsible practice. The convergence of these strands supports an integrated interpretation of the two literacies without requiring an ambiguously specified cross-domain correlation.

The integration suggests that assessment literacy can be interpreted as the structural dimension of responsible assessment practice, while feedback literacy provides a dialogic and reflective dimension through which assessment evidence is discussed, revised, and used. This is an interpretive mixed-methods inference: the quantitative analyses establish parallel phase patterns and within-domain continuity, while the qualitative findings explain how students connected assessment and feedback in practice.

DISCUSSION

The present study examined assessment literacy and feedback literacy within a two-wave, multi-cohort teacher

education context. Quantitative findings showed higher A2 than A1 scores in both domains and strong A1-to-A2 prediction within each domain, while qualitative findings explained how different course phases and cohort contexts were associated with changing emphases on fairness, rubric use, peer collaboration, instructor feedback, and professional application. Because the A1 and A2 assessment-literacy examinations were not psychometrically equated, the AL phase difference is interpreted as a standardized course-performance difference rather than as invariant-scale growth. Taken together, the findings support treating assessment literacy and feedback literacy as complementary dimensions of responsible assessment practice, without claiming direct causal co-development.

Connected Development of Assessment and Feedback Literacy

The quantitative findings show parallel course-phase patterns rather than a demonstrated cross-domain statistical effect. A2 assessment-literacy course performance and overall feedback-literacy scores were higher than A1 values, and first-phase scores strongly predicted second-phase scores within each domain. Assessment literacy has traditionally been conceptualized as knowledge of assessment design, interpretation, and decision-making (Popham, 2009; Brookhart, 2011; DePascale, 2018), while feedback literacy emphasizes the capacity to interpret, seek, and use feedback (Carless and Boud, 2018; Boud and Dawson, 2023). The present analysis does not report a cross-domain correlation whose score pairing cannot be unambiguously reconstructed from the de-identified secondary archive. The claim of connection therefore rests on parallel quantitative patterns plus qualitative and theoretical integration, not on an asserted direct AL-FL coefficient.

This interpretation is strengthened by the qualitative findings. Students consistently described how rubric clarification, instructor modelling, peer review, and opportunities to revise assessment outputs transformed assessment from a technically demanding requirement into a more understandable and professionally meaningful practice. These experiences illustrate that assessment competence is reinforced when learners are actively engaged in feedback processes rather than merely receiving assessment outcomes. Such findings are consistent with the growing literature describing feedback as an active process of meaning-making rather than a one-way transmission of comments (Carless and Boud, 2018; Winstone, Mathlin and Nash, 2019; Boud and Dawson, 2023).

The strong within-domain longitudinal continuity suggests that first-course performance was closely associated with second-course performance. For assessment literacy, this continuity spans related but non-equated examinations; for feedback literacy, it spans repeated total scores on the same local score metric. These findings are compatible with cumulative learning across sequential courses, but they should not be read as evidence that one literacy causes change in the other. This interpretation aligns with Deneen and Hoo (2021), who emphasize participation in evaluative activities, and with Panadero and Lipnevich (2022), who view assessment and feedback as interconnected educational processes.

AFLN and Responsible Assessment Practice

The findings provide mixed-methods evidence consistent with the Assessment-Feedback Literacy Nexus (AFLN) proposed in this study. AFLN extends assessment-literacy and feedback-literacy scholarship by positioning the two literacies within a single responsibility-oriented framework. Assessment literacy provides a structural foundation for generating and interpreting defensible evidence, while feedback literacy provides a dialogic process through which that evidence can be discussed, revised, and used for learning. The framework is explanatory: it does not assert a statistically demonstrated bidirectional causal mechanism.

This perspective responds to a gap in the literature concerning how assessment and feedback can be understood together in preservice teacher education. Rather than claiming that the present data prove that the two literacies develop as a single latent process, AFLN proposes that responsible assessment practice requires both sound assessment design and meaningful feedback engagement. The present quantitative patterns and qualitative accounts are treated as convergent evidence for that integrative proposition.

The qualitative findings illustrate this interaction. Students increasingly viewed rubrics as communication tools rather than grading instruments, feedback as dialogue rather than correction, and assessment as a professional responsibility rather than a procedural requirement. These observations support broader arguments that responsible assessment extends beyond measurement to include fairness, transparency, learner participation, and ethical use of evidence (Messick, 1995; Shepard, 2000; McArthur, 2016). Consequently, AFLN contributes not only a conceptual integration of two literacy traditions but also a professional perspective that aligns assessment practice with educational responsibility.

Program Differences and Targeted Support

Programme differences observed in the quantitative analyses indicate that assessment and feedback literacy did not develop uniformly across teacher education programmes. While differences should not be interpreted as indicators of programme quality, they suggest that preservice teachers encounter different opportunities, curricular emphases, and learning conditions throughout their preparation.

The qualitative findings provide possible explanations for these differences. Students repeatedly identified instructor modelling, peer collaboration, access to exemplars, structured feedback, and practicum experiences as important supports for literacy development. Programmes providing more frequent opportunities for collaborative assessment design and iterative feedback may therefore create stronger conditions for literacy development than programs with fewer such opportunities.

These findings have practical implications for teacher education. Rather than adopting a uniform approach to assessment instruction, institutions may consider differentiated support based on programme characteristics. Additional scaffolded practice, structured feedback cycles, rubric calibration activities, and authentic assessment experiences may be particularly beneficial in programmes where students have fewer opportunities to engage in assessment design

and interpretation. Such recommendations are consistent with recent work advocating sustained assessment literacy development throughout teacher preparation rather than within isolated assessment courses (Gan and Lam, 2024; Soicher, Baker and Thomas, 2024).

Implications for Teacher Education

The study has several implications for teacher education curriculum and practice. First, assessment courses should integrate assessment literacy and feedback literacy rather than treating them as separate instructional topics. Preservice teachers should repeatedly design assessment tools, receive structured feedback, revise their work, and reflect on how assessment evidence informs future instructional decisions.

Second, teacher educators should position feedback as a central component of assessment instruction rather than as an activity occurring after assessment has been completed. Structured peer review, instructor modelling, rubric discussion, and opportunities for revision appear to support both technical competence and professional judgment. Such approaches encourage students to understand assessment as an iterative process rather than a sequence of isolated tasks.

Third, the findings reinforce the importance of professional responsibility within assessment education. Students increasingly recognized that assessment decisions influence learner confidence, participation, and educational opportunity. Consequently, teacher education should continue to emphasise validity, fairness, transparency, and learner agency as integral components of assessment competence rather than as supplementary ethical considerations.

Finally, AFLN offers a useful conceptual lens for future curriculum development. Rather than organizing assessment courses around disconnected technical skills, teacher education programmes may benefit from structuring learning experiences that explicitly connect assessment design, feedback processes, reflection, and responsible professional practice.

The Contribution of the AFLN Framework

The findings are consistent with AFLN as an integrative framework for understanding assessment and feedback literacy development in teacher education. Its contribution is not to claim that assessment literacy and feedback literacy follow a fixed causal sequence. Rather, AFLN explains how assessment and feedback become connected in the formation of responsible teaching practice.

Assessment literacy provides the structural dimension of the framework. It helps preservice teachers design valid tasks, align assessment with outcomes, construct rubrics, interpret scores, and make defensible judgments. Feedback literacy provides the dialogic dimension. It helps preservice teachers understand how assessment information is communicated, interpreted, questioned, revised, and used for learning. AFLN brings these dimensions together by positioning assessment and feedback as interdependent practices of professional responsibility.

This contribution extends prior local work while making the overlap explicit. ALiPP (Author, 2026) is the source of the 250 matched assessment-literacy records and the interview/reflection corpus; AL-only results reproduced in this article

are therefore contextual anchors rather than novel findings. ASPIRE (Author, Author and Author, 2025) was a separate 29-participant study and contributes conceptual context only; none of its participant-level data are included here. The new contribution of the present article is the FL_A1/FL_A2 component and its integration with the ALiPP evidence to formulate AFLN as a responsibility-oriented explanatory framework grounded additionally in wider scholarship on feedback literacy, fairness, validity, dialogue, and professional responsibility.

The present findings also clarify the appropriate scope of AFLN. At this stage, AFLN should not be presented as a fully validated causal model. It is better understood as a responsibility-oriented explanatory framework supported by parallel two-wave score patterns, strong within-domain continuity, programme-level differences, and qualitative evidence of dialogic and ethical assessment practice. This positioning preserves the theoretical usefulness of AFLN while remaining consistent with the limits of the secondary data.

Limitations and Future Research

Several limitations should be considered when interpreting the findings. First, the study was conducted within a single teacher education institution, limiting generalisability. Second, programme sizes were markedly unequal, especially BECED ($n = 7$), so programme comparisons are descriptive of this context and should not be used to rank programme quality. Third, A1 and A2 assessment-literacy examinations were separately developed 60-item course assessments with acceptable internal consistency ($KR-20 > .75$ in the source study) but were not psychometrically equated; consequently, their standardized score difference cannot be interpreted as growth on an invariant scale. Fourth, feedback literacy was available only as overall FL_A1 and FL_A2 scores. Archived documentation reported content review and overall $\alpha = .85$, but item-level data were unavailable for recomputing phase-specific reliability or testing measurement invariance. Fifth, the de-identified secondary archive did not permit unambiguous reconstruction of the previously reported cross-domain correlation, so that coefficient is not reported in this revision. Sixth, the qualitative archive contained 20 semi-structured interviews and post-A1/post-A2 reflection materials, but the consolidated archive did not preserve a defensible count of separate reflection documents; phase- and cohort-level patterns are therefore reported cautiously and are not treated as four-year within-person trajectories. Finally, the researcher occupied an insider/instructor position, creating potential power and social-desirability effects despite documented research-use consent for participant-generated responses, de-identification, pseudonymisation, and aggregate reporting.

Future research should examine AFLN across multiple institutions using prospectively designed measurement that places assessment literacy and feedback literacy on well-documented, repeated metrics. At least three measurement occasions would permit stronger modelling of developmental trajectories, and item-level data would allow phase-specific reliability and measurement-invariance testing. Future studies

should also pre-specify mixed-effects or other repeated-measures models where observations are nested within students, document qualitative sample sizes and source counts prospectively, and test cross-domain relationships using clearly defined score pairings. Such work could establish whether the conceptual connections proposed by AFLN hold beyond the present secondary-analysis context.

CONCLUSION

This study examined assessment literacy and feedback literacy among 250 matched preservice teachers across two sequential course phases embedded within five successive cohorts from 2020 to 2024. A2 standardized assessment-literacy course performance and A2 overall feedback-literacy scores were higher than A1 values, and first-course scores strongly predicted second-course scores within each domain. The assessment-literacy difference is interpreted cautiously because the A1 and A2 examinations were not psychometrically equated, and the revised analysis does not claim a direct cross-domain statistical association or causal co-development.

The qualitative reanalysis showed phase- and cohort-level contrasts between compliance-oriented accounts and accounts emphasizing professional responsibility. Rubric use, peer

collaboration, instructor modelling, practicum application, and feedback dialogue helped explain how students understood assessment as involving fairness, clarity, transparency, revision, and learner agency. These qualitative patterns contextualise the quantitative findings without being treated as a four-year within-person trajectory.

The study contributes the Assessment-Feedback Literacy Nexus as an integrative explanatory framework for teacher education. AFLN positions assessment literacy as a structural dimension of responsible assessment practice and feedback literacy as a dialogic dimension through which assessment evidence becomes meaningful for learning. Its contribution is conceptual and integrative rather than causal: the present secondary data support convergence across quantitative and qualitative evidence but do not constitute full validation of a causal model.

For teacher education programmes, the findings point to the need for assessment courses that include repeated feedback cycles, rubric calibration, guided revision, peer review, and practicum-based application. Future studies may extend the AFLN framework across institutions, programmes, and practicum contexts to examine how assessment and feedback literacy transfer into actual teaching practice.

REFERENCES

- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014) *Standards for educational and psychological testing*, Washington, DC: American Educational Research Association.
- Boud, D. and Dawson, P. (2023) 'What feedback literate teachers do: An empirically-derived competency framework', *Assessment & Evaluation in Higher Education*, Vol. 48, No. 2, pp. 158–171. <https://doi.org/10.1080/02602938.2021.1910928>
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, Vol. 3, No. 2, pp. 77–101. <https://doi.org/10.1080/2159676x.2019.1704846>
- Braun, V. and Clarke, V. (2021) 'To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales', *Qualitative Research in Sport, Exercise and Health*, Vol. 13, No. 2, pp. 201–216. <https://doi.org/10.1080/2159676x.2019.1704846>
- Brookhart, S. M. (2011) 'Educational assessment knowledge and skills for teachers', *Educational Measurement: Issues and Practice*, Vol. 30, No. 1, pp. 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Carless, D. and Boud, D. (2018) 'The development of student feedback literacy: Enabling uptake of feedback', *Assessment & Evaluation in Higher Education*, Vol. 43, No. 8, pp. 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Commission on Higher Education (2012) *CHED Memorandum Order No. 46, s. 2012: Policy-standard to enhance quality assurance (QA) in Philippine higher education through an outcomes-based and typology-based QA*, Quezon City: Commission on Higher Education.
- Commission on Higher Education (2017a) *CHED Memorandum Order No. 74, s. 2017: Policies, standards and guidelines for Bachelor of Elementary Education*, Quezon City: Commission on Higher Education.
- Commission on Higher Education (2017b) *CHED Memorandum Order No. 75, s. 2017: Policies, standards and guidelines for Bachelor of Secondary Education*, Quezon City: Commission on Higher Education.
- Commission on Higher Education (2017c) *CHED Memorandum Order No. 76, s. 2017: Policies, standards and guidelines for Bachelor of Early Childhood Education*, Quezon City: Commission on Higher Education.
- Commission on Higher Education (2017d) *CHED Memorandum Order No. 77, s. 2017: Policies, standards and guidelines for Bachelor of Special Needs Education*, Quezon City: Commission on Higher Education.
- Creswell, J. W. and Creswell, J. D. (2018) *Research design: Qualitative, quantitative, and mixed methods approaches*, 5th edition, Thousand Oaks, CA: SAGE Publications.
- Dawson, P., Yan, Z., Lipnevich, A., Tai, J., Boud, D. and Mahoney, P. (2023) 'Measuring what learners do in feedback: The feedback literacy behaviour scale', *Assessment & Evaluation in Higher Education*, Vol. 49, No. 3, pp. 348–362. <https://doi.org/10.1080/02602938.2023.2240983>
- Deneen, C. C. and Hoo, H.-T. (2021) 'Connecting teacher and student assessment literacy with self-evaluation and peer feedback', *Assessment & Evaluation in Higher Education*, Vol. 48, No. 2, pp. 214–226. <https://doi.org/10.1080/02602938.2021.1967284>
- DePascale, C., Sharp, A., Ryan, K. and Betebenner, D. (2018) *Building a conceptual framework for assessment literacy*, Dover, NH: National Center for the Improvement of Educational Assessment.

- Department of Education (2017) *National adoption and implementation of the Philippine Professional Standards for Teachers*, Department Order No. 42, s. 2017, Pasig City: Department of Education.
- Gan, L. and Lam, R. (2024) ‘Unveiling classroom assessment literacy: Does teachers’ self-directed development play out?’, *Education Sciences*, Vol. 14, No. 9, p. 961. <https://doi.org/10.3390/educsci14090961>
- Girgla, A., Good, L., Krstic, S., McGinley, B., Richardson, S., Sneidze-Gregory, S. and Star, J. (2021) *Developing a teachers’ assessment literacy and design competence framework*, Camberwell: Australian Council for Educational Research.
- Khalid, N. H. M., Latif, A. A. and Yusof, I. J. (2021) ‘Research focusing on assessment literacy: A systematic literature review’, *Estudios de Economía Aplicada*, Vol. 39, No. 10, pp. 1–25. <https://doi.org/10.25115/eea.v39i10.5816>
- McArthur, J. (2016) ‘Assessment for social justice: The role of assessment in achieving social justice’, *Assessment & Evaluation in Higher Education*, Vol. 41, No. 7, pp. 967–981. <https://doi.org/10.1080/02602938.2015.1053429>
- Messick, S. (1995) ‘Validity of psychological assessment: Validation of inferences from persons’ responses and performances as scientific inquiry into score meaning’, *American Psychologist*, Vol. 50, No. 9, pp. 741–749. <https://doi.org/10.1037/0003-066x.50.9.741>
- Nieminen, J. H. and Carless, D. (2022) ‘Feedback literacy: A critical review of an emerging concept’, *Higher Education*, Vol. 85, No. 6, pp. 1381–1400. <https://doi.org/10.1007/s10734-022-00895-9>
- Panadero, E. and Lipnevich, A. A. (2022) ‘A review of feedback models and typologies: Towards an integrative model of feedback elements’, *Educational Research Review*, Vol. 35, p. 100416. <https://doi.org/10.1016/j.edurev.2021.100416>
- Popham, W. J. (2009) ‘Assessment literacy for teachers: Faddish or fundamental?’, *Theory Into Practice*, Vol. 48, No. 1, pp. 4–11. <https://doi.org/10.1080/00405840802577536>
- Quileste, R. M. (2026) ‘Assessment literacy in teacher education: Integrating performance and perceptions via the ALiPP framework’, *International Journal of Research in Education and Science*, Vol. 12, No. 1, pp. 158–176. <https://doi.org/10.46328/ijres.5244>
- Quileste, R., Estacio, M. L. P. and Pagara, C. R. (2025) ‘Exploring the ASPIRE framework: Enhancing assessment literacy through inquiry-based teacher education’, *International Journal of Education and Teaching Zone*, Vol. 4, No. 3, pp. 243–258. <https://doi.org/10.57092/ijetz.v4i3.447>
- Shepard, L. A. (2000) ‘The role of assessment in a learning culture’, *Educational Researcher*, Vol. 29, No. 7, pp. 4–14. <https://doi.org/10.3102/0013189x029007004>
- Soicher, R. N., Baker, A. R. and Thomas, R. C. (2024) ‘A Mixed-Methods Research Design to Advance inclusive and equitable teaching’, *Innovative Higher Education*, Vol. 49, No. 6, pp. 1105–1125. <https://doi.org/10.1007/s10755-024-09741-5>
- Winstone, N. E., Mathlin, G. and Nash, R. A. (2019) ‘Building feedback literacy: Students’ perceptions of the Developing Engagement with Feedback Toolkit (DEFT)’, *Frontiers in Education*, Vol. 4, p. 39. <https://doi.org/10.3389/feduc.2019.00039>

APPENDIX

APPENDIX A

Theme	Description	Illustrative quote	Pattern	AFLN interpretation
Conceptual overload	Difficulty managing technical assessment concepts such as TOS, validity, reliability, and KR-20	“CHED CMO topics like reliability... were overwhelming.” (S2, 2020, Zoom)	Strong in 2020-2021	Technical AL demands created uncertainty and feedback-related anxiety
Difficulty applying concepts	Students could prepare assessment tools but were unsure whether these worked in practice	“I could make a TOS but didn’t know if it worked.” (S4, 2021, Zoom)	High in online cohorts	Limited practice weakened the assessment-feedback connection
Feedback gaps	Students needed more specific feedback on outputs such as rubrics and assessment tools	“I didn’t know if [my rubric] was valid or not.” (S7, 2023, F2F)	Persistent	Feedback gaps limited revision and professional judgment
Statistical anxiety	Computation and interpretation tasks caused anxiety	“KR-20 felt like a different subject.” (S9, 2020, Zoom)	High in A1	Feedback and guided modeling helped reduce assessment anxiety
Modality challenges	Online learning limited immediate clarification and modeling	“Hard to know if I was doing things right on screen.” (S5, 2021, Zoom)	Strong in 2020-2021	Modality shaped access to feedback dialogue
LMS resources	Templates and samples helped students understand assessment expectations	“Rubric samples in our Drive helped a lot.” (S3, 2021, Zoom)	Across years	Structured resources supported the AL-FL bridge
Instructor modeling	Demonstration and guided revision helped students understand criteria	“We revised rubrics in class with sir.” (S6, 2023, F2F)	Strong after 2022	Modeling supported dialogic and responsible assessment practice
Peer collaboration	Peer review helped students notice gaps in their work	“Groupmates pointed out what I missed.” (S1, 2024, F2F)	Strong in F2F cohorts	Peer feedback strengthened interpretation and revision
Digital tools	Excel and Jamovi reduced fear of assessment computations	“Jamovi... wasn’t scary anymore.” (S8, 2023, F2F)	Strong after 2022	Technology supported assessment interpretation
Practicum integration	Real or simulated teaching helped students see the use of assessment tools	“My self-made rubric actually worked!” (S10, 2024, F2F)	Strong after 2022	Practice connected assessment structure with feedback use

Table A1: Challenges and supports across course phases and cohort contexts, 2020-2024 (source: secondary thematic reanalysis)

INSTITUTIONAL RESPONSIBILITY AND CONNECTIVITY EFFICIENCY IN PUBLIC SCHOOLS DURING THE COVID-19 PANDEMIC: A COMPARATIVE ANALYSIS OF THE DIGITAL DIVIDE IN TUNJA, COLOMBIA (2020–2021)

ABSTRACT

The COVID-19 pandemic highlighted deep structural inequalities in educational connectivity across low- and middle-income countries. This study evaluated the relationship between socioeconomic vulnerability and internet access in 14 public schools in Tunja, Colombia, during the most critical stage of social isolation in 2020 and one year later. An institutional census design was combined with Pearson correlations, paired-samples t-tests, and clustered ordinary least squares regression to analyse school-level connectivity patterns. The mean school-level proportion of students without internet access decreased from 56.5% in 2020 to 32.5% in 2021, corresponding to a 24.0-percentage-point increase in average school-level connectivity (95% CI: 10.7–37.2; $p = 0.0018$). A strong inverse association between the socioeconomic index and school-level connectivity was observed in both 2020 ($r = -0.804$, $p = 0.00053$) and 2021 ($r = -0.794$, $p = 0.00069$). Despite government efforts to distribute laptops and SIM cards, some of the most vulnerable institutions, particularly rural and peri-urban schools, showed limited improvement or even deterioration in connectivity. The findings suggest that expanding connectivity is necessary but may be insufficient to achieve educational equity. Digitally inclusive policies should therefore be territorially sensitive and pedagogically grounded if they are to address structural educational inequalities.

KEYWORDS

COVID-19, education, efficiency, technology adoptions

HOW TO CITE

Uscategui-Gómez A., Igua Muñoz B. D., Campos Medina J. C., Igua Muñoz J. S., Igua Muñoz J. (2026) 'Institutional Responsibility and Connectivity Efficiency in Public Schools During the COVID-19 Pandemic: a Comparative Analysis of the Digital Divide in Tunja, Colombia (2020–2021)', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 222–236. <http://dx.doi.org/10.7160/eriesj.2026.190304>

Alejandra Uscategui-Gómez¹
Byron David Igua Muñoz²
Juan Carlos Campos Medina^{4, 5}
Joan Sebastian Igua Muñoz^{3, 4}
Jonathan Salomón Igua-Muñoz^{6, 7}✉

¹Maestría Interdisciplinaria en Ciudades Sostenibles, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

²Universidad Santo Tomás, Facultad de Ingeniería Electrónica, Maestría en Ingeniería, Tunja, Colombia

³Programa Todos a Aprender PTA/FI 3.0- Secretaría de Educación, Sogamoso, Colombia

⁴Universidad Pedagógica y Tecnológica de Colombia, Tunja, Colombia

⁵ENS Santiago de Tunja, Colombia

⁶Instituto Potosino de Investigación Científica y Tecnológica, A.C., San Luis Potosí, Departamento de Ciencias Ambientales, México

⁷LPAS: Laboratorio de Paisajes Antrópicos Sostenibles. San Luis Potosí, México

✉ jonathan.igua@ipicyt.edu.mx

Article history

Received

August 17, 2025

Received in revised form

May 13, 2026

Accepted

September 7, 2026

Available on-line

September 30, 2026

Highlights

- Average school-level internet connectivity increased by 24.0 percentage points between 2020 and 2021 across public schools in Tunja.
- Socioeconomic vulnerability was strongly and inversely associated with school connectivity in both 2020 and 2021.
- Connectivity gains remained uneven, with rural and highly vulnerable schools continuing to experience substantial digital-access gaps.
- Providing laptops and SIM cards alone did not consistently translate into greater connectivity, highlighting the need for territorially sensitive digital-inclusion policies.

INTRODUCTION

The COVID-19 pandemic and the lockdown measures adopted by governments exposed widespread inequalities across multiple sectors of society, with education among

the most severely affected. Worldwide, schools were closed for an average of nearly 95 school days between March 2020 and February 2021, leaving more than 1.6 billion students without access to in-person education and exacerbating

learning gaps, particularly among vulnerable populations lacking adequate connectivity or digital resources. Before the pandemic, approximately 59% of the global population had internet access. By 2021, this figure had risen to around 61%, equivalent to almost 4.8 billion users worldwide, reflecting the rapid expansion of global connectivity from a penetration rate of only 17% in 2005 to more than 53% in 2019 (Galeno, 2020; Wang and Lin, 2024). However, this growth was uneven across socioeconomic and geographic contexts (Zapletal et al., 2023; Haelermans et al., 2022; UN, 2020).

One of the clearest manifestations of this digital inequality is found in Latin America, where, during the COVID-19 pandemic, less than 42% of the urban population and only 14% of the rural population had internet access (LEE, 2020). This contrasts with countries in the Global North, including the United States, where internet connectivity reached 77% in urban areas and 73% in rural areas, and the European Union, where the corresponding figures were 82% and 68%. These differences reflected substantial disparities in educational response capacity during lockdown. In Latin America, the digital divide has been linked to differences in the access, coverage, and affordability of public information and communication technology (ICT) services (Becerra and Mastrini, 2017). At the same time, the health emergency accelerated digitalisation processes and the expansion of connectivity in response to the educational demands imposed by lockdown (Olson et al., 2025; Clercq et al., 2023; Busuttil and Farrugia, 2020; Morata-Sanz, 2020).

In response, Latin American governments implemented policies intended to guarantee connectivity as an essential service and sustain academic, work, and community activities remotely. Argentina, Brazil, Chile, Colombia, and Mexico, among others, adopted measures such as declaring ICT services essential, prohibiting disconnections for non-payment, freezing tariffs, and offering low-cost plans to vulnerable populations. Measures were also introduced to manage data traffic and consumption, strengthen infrastructure and radio-spectrum management, and support telecommunications operators in maintaining network availability and protecting vulnerable users. Although virtual classes, social media platforms, television, and radio became important resources for teachers, the international literature also documented inadequate or improvised infrastructure, limited teacher preparation, and restricted household access to internet services and devices, all of which affected learning quality in vulnerable settings (ENACOM, 2020; Califano and Becerra, 2024; Joshi et al., 2023).

In Colombia, limitations in broadband access have long reflected broader social, economic, and territorial inequalities (Montes, 2019). According to the Banco Mundial (2020), substantial income inequality also persisted during this period. MinTIC subsequently reported a 140% increase in 4G adoption between July 2017 and November 2019, an expansion that contributed to the technological capacity available when educational services moved online (MinTIC, 2024; Pardo-Figueroa and Obando-Fernández, 2021). Nevertheless, marked regional disparities remained: approximately half of households in the Andean and Pacific regions had fixed internet access, whereas rates in the Amazon and Orinoquía regions

remained below 20% (Kairuz, 2020). These inequalities reinforce the need to examine how connectivity gaps intersect with educational equity and how public action can strengthen technological infrastructure and ICT access.

Tunja, the capital of Colombia's Boyacá Department, provides a relevant case for examining these issues. Recognised by UNESCO as a Learning City, Tunja has a comparatively robust public education system, with approximately 18.3% of its 200,000 inhabitants enrolled in basic, middle, and secondary education institutions (SIMAT, 2020). However, it remained unclear how strongly institutional socioeconomic vulnerability was associated with connectivity across public schools during the crisis, how connectivity changed between the peak of lockdown and the following year, and whether resource provision corresponded to improvements in the most vulnerable institutions. In line with the concepts of efficiency and accountability (Reuge et al., 2021), this study treats efficiency as the alignment between technological resources allocated through public policy and observed institutional connectivity outcomes, while accountability concerns the prioritisation of institutions facing greater socioeconomic disadvantage. Accordingly, the study aims to: (i) characterise the technological and connectivity conditions of public schools in Tunja during full social isolation (2020) and one year later (2021); (ii) analyse the association between institutional socioeconomic index and internet access during the study period; and (iii) describe the distribution of equipment and connectivity measures implemented by the local administration in relation to observed changes in the digital divide among schools under its jurisdiction.

This paper is structured as follows: the next section describes the data collection and analytical methods; subsequently, the results are presented; this is followed by a discussion of the findings in dialogue with the literature; and it concludes with recommendations aimed at advancing digital inclusion in vulnerable educational contexts.

MATERIALS AND METHODS

Study Area

The study was conducted in Tunja, capital of the department of Boyacá, Colombia. Founded in 1539, Tunja is recognised as a UNESCO Learning City and as the birthplace of Colombia's public education system, established on 17 May 1822 through an initiative led by Vice President Francisco de Paula Santander and Secretary of the Interior José Manuel Restrepo (Colboy, 2020). Since then, Tunja has maintained a longstanding educational tradition and has been recognised for strong performance in national standardised assessments (Moreno-Villamil, 2023; Rivera, 2013).

The public primary and secondary education network of Tunja consists of 14 schools managed by the local administration, serving approximately 23,000 students, representing 64.0% of the city's total enrolment of 35,971 students. This study is an institutional census. All 14 public schools under the jurisdiction of the municipal education authority were included in the census in both 2020 and 2021. The results should not be generalised to private institutions or other municipalities without comparable administrative and socio-economic contexts.

Tunja was certified as an autonomous educational entity by Resolution No. 2755 of 3 December 2002, issued by the Ministry of National Education. The city has maintained nationally recognised education-quality initiatives and has performed strongly in the Saber 11 assessments among public institutions. The local administration provides broadband

internet service to 37 official school sites under its authority, located in central urban (15), peri-urban (21), and rural (1) areas (Figure 1, Table 2). Nevertheless, access to and use of digital technologies remain more limited among vulnerable populations, particularly in rural and peripheral areas in the southern part of the city.

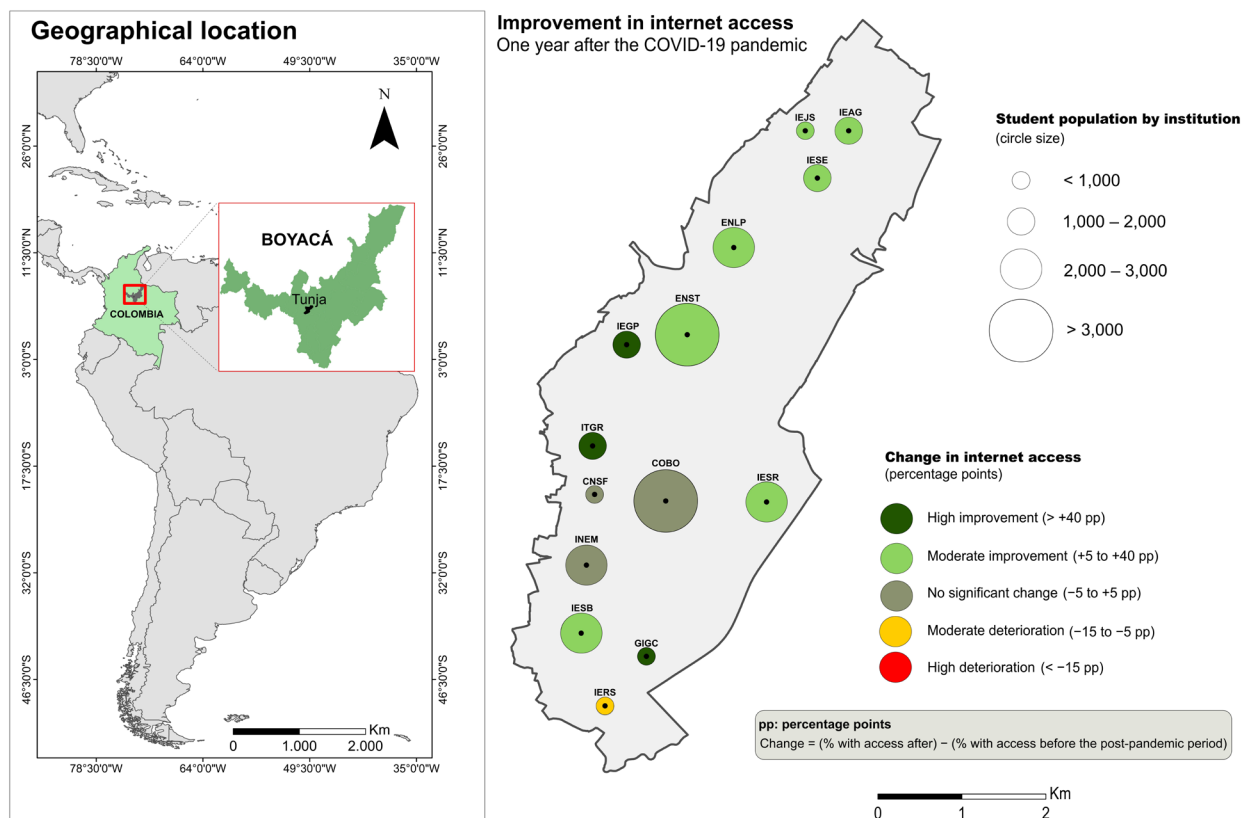


Figure 1: Study area. Left panel: location of Boyacá within Colombia and South America, with the urban boundary of Tunja shown in black. Right panel: location of schools, each represented by a circle. Circle size is proportional to the average student population over the study period (2020–2021). Circle colour indicates the change in the proportion of students with internet access (percentage points, pp), calculated as the percentage with access in 2021 minus the percentage with access in 2020. Population size categories and change ranges are detailed in the legend. Table 2 lists the reference codes for the institutions and associated complementary information. Reference system: WGS84 / UTM Zone 18N (EPSG: 32618).

Information collection

The research was conducted in accordance with Colombia's ethical protocols for information management and human participant research (Resolution 8430 of 1993, Ministry of Health). The dataset analysed in this study corresponds to administrative information collected by the Municipal Secretariat of Education of Tunja within its statutory framework to evaluate and consolidate municipal educational data (Law 115 of 1994, Article 151; Resolution No. 2755 of 2002, Ministry of National Education). This study constitutes a secondary analysis of pre-existing institutional records rather than primary research involving the active recruitment of human participants. Table S2 corresponds to a blank template of the internal classroom-level tracking instrument used by each principal to compile teachers' reports within the institution; it was not submitted to the Municipal Secretariat of Education or to the research team. Consequently,

although the template included a field for the teacher's name for internal administrative use by the principal, this identifying information never left the institution and was not accessible to, or processed by, the research team. The methods employed correspond to documentary and secondary analysis techniques, categorized as low-risk under Article 11(a) of Resolution 8430 of 1993. Consequently, given the secondary, anonymized, and aggregated nature of the data, individual informed consent and formal ethics committee review were not required. The instruments were supervised and validated through review by the professional team of the Educational Quality Division of the Municipal Secretariat of Education. Formal communication with public educational institutions was carried out under Resolution 00237 of 2020, and Circular No. 090 of 2021 enabled the reapplication of the instruments to compare data across both periods.

Data collection operated through a two-level procedure. At the classroom level, teachers completed a structured

survey covering student internet access, available devices, and pedagogical communication tools used during remote instruction (Table S2). Each teacher reported exclusively on their own assigned group of students, with mutually exclusive, non-overlapping student rosters across teachers within a school, preventing double counting at the classroom level. Because each student was enrolled in only one classroom group per year, and each teacher's roster was restricted to their own assigned group, no student could be reported by more than one teacher, even across different campuses of the same institution. Percentages of communication-tool usage reported in Figure 4 correspond to the total number of students reached through each channel, relative to total enrolment for each year (Table 1), consistent with the denominator used in Table 3. At the institutional level, school principals consolidated these non-overlapping group-level reports into the institutional-level form (Table S1), and the resulting totals were subsequently reconciled against official per-grade enrolment recorded in SIMAT (Sistema de Matriculas Estudiantil), Colombia's official student enrolment system, providing an additional safeguard against double counting. Both instruments were designed and administered via Google Forms and together comprised more than 30 response fields organised across six thematic sections. Before completion, the instruments and their objectives were shared with teachers and school administrators to ensure accurate and consistent responses across all campuses. Although a formal pilot study was not possible in the context of the pandemic emergency, expert review, institutional authorisation, and socialisation with participants served as the primary mechanisms for ensuring the validity and reliability of the collected data. The same procedure was applied identically in 2020 and 2021.

For the purposes of this study, internet connectivity was defined as any form of functional internet access available for educational activities, including broadband and mobile data connections. This definition encompasses the two modalities reported separately in the surveys and underpins the aggregated connectivity indicators used throughout the analyses. All 14 public schools participated in both study periods (2020 and 2021), achieving full institutional coverage of the target population.

Information systematization

After data collection, the information was tabulated and analysed using R version 4.4.2 and Microsoft Excel 2016. Data reported by each institution were subjected to internal consistency checks, including verification of totals and

proportions, and were cross-referenced with official enrolment figures recorded in SIMAT for each year. Minor discrepancies of less than 1% were verified directly with the corresponding institution before the database was finalised.

For historical equipment inventories from 2020 and selected administrative records of technological provision, not available (NA) values occurred for Colegio de Boyacá and Colegio Nuestra Señora de Fátima (Table 4) because consolidated reports were unavailable or administrative records were still being updated during data collection.

Missing data management

Because the study was an institutional census ($n = 14$), no missing values were recorded for the primary study variables (connectivity, number of students, and socioeconomic stratification), all of which were cross-checked against the SIMAT database. NA values occurred only in secondary variables related to technological inventories and equipment-provision records for two institutions, as described above. These cases were excluded from the relevant descriptive summaries; no statistical imputation was performed.

Information analysis

Descriptive statistics were computed for connectivity and socioeconomic status. The unit of analysis was the public school ($n = 14$). Unless otherwise indicated, the unweighted mean of the 14 school-level connectivity percentages (Table 2) was adopted as the primary estimand of this study, consistent with the institutional census design and unit of analysis, and is reported consistently across the text, tables, and figures. These values differ from the enrolment-weighted proportions derived from the raw totals in Table 1 (45.4% in 2020 and 69.7% in 2021), which arise because the weighted calculation gives proportionally greater influence to schools with larger enrolments in the resulting average; these values are reported here once, for transparency, and are not used elsewhere in the manuscript. The Shapiro–Wilk test was used to assess the distribution of within-school changes in connectivity between 2020 and 2021 ($W = 0.958$, $p = 0.695$) before applying the paired-samples t-test. Pearson correlation coefficients were used to examine the association between average school-level connectivity and the socioeconomic index. The index was calculated as the percentage of enrolled students, from pre-school to Year 11, belonging to the most vulnerable socioeconomic categories (strata 0 and 1) according to the Colombian socioeconomic classification system (Bonilla et al., 2014), using the following formula:

$$\text{Socioeconomic index} = \frac{\text{No. of students in strata 0 and 1}}{\text{Total enrolled students}} \cdot 100$$

Note: All index values were derived exclusively from SIMAT records, without additional transformations. The strata considered correspond exclusively to strata 0 (very low) and strata 1 (low). For example: the Institución Educativa Libertador Simón Bolívar reported 1,170 students in strata 0 and 1 out of a total of 1,334 enrolled students, representing a Socioeconomic Index of 87.7%.

Pearson's product-moment correlation coefficient (r) was used to quantify the association between school-level connectivity and the socioeconomic index. The coefficient of determination ($R^2 = r^2$) was reported as the proportion of variance in connectivity statistically associated with the socioeconomic index. Because the same 14 institutions were observed in 2020 and 2021, between-year comparisons were conducted

using a paired-samples *t*-test on school-level connectivity percentages, thereby accounting for the paired institutional structure. In addition, an ordinary least squares (OLS) model was fitted with year as a binary predictor (2020 = reference; 2021 = 1). Standard errors were clustered at the school level to account for repeated observations within institutions.

Statistical analysis

Analyses were aligned with the study objectives. To characterise changes in connectivity between 2020 and 2021 (Objective I), a paired-samples *t*-test was performed on school-level percentages. To examine the association between institutional socioeconomic vulnerability and connectivity (Objective II), Pearson correlation coefficients (*r*) were calculated for each year at school level; and (Objective III), concerning the distribution of equipment and connectivity measures implemented by the local administration, was addressed descriptively; no inferential analysis was performed to test whether resource allocation predicted connectivity change or was proportional to institutional vulnerability. Finally, an OLS regression model with year as a binary predictor (2020 = reference; 2021 = 1) was used as a supplementary analysis of average change over time, with school-clustered standard errors to account for repeated observations. Cluster-robust standard errors were estimated using the CR2 bias-reduced linearisation correction with Satterthwaite degrees of freedom (Pustejovsky and Tipton, 2018), which is recommended when the number of clusters is small. Because the dataset comprises the full population of public schools under municipal jurisdiction rather than

a probability sample, inferential statistics are interpreted as model-based evidence regarding the observed pattern rather than as conventional finite-population sampling inference.

RESULTS

This section presents the results of the comparative analysis of connectivity and the socioeconomic characterization of public schools in Tunja. The analysis included all 14 public schools under municipal jurisdiction in both 2020 and 2021, corresponding to a 100% institutional coverage rate in both periods. All results are based on school-level aggregated data (*n* = 14), consistent with the census design of the study.

For descriptive purposes, schools were grouped according to the proportion of students in vulnerable socioeconomic strata and changes in internet coverage across the two study years. The comparison of school-level connectivity in 2020 and 2021 provides insight into how access changed during the COVID-19 emergency. These patterns should be interpreted cautiously because location and other unmeasured institutional and territorial factors may also have influenced connectivity outcomes.

At the school level, the mean proportion of students without internet access was 56.5% in 2020, while the mean proportion with access was 43.5% (Figure 2). By 2021, mean school-level connectivity had increased to 67.5%, corresponding to an average within-school gain of 24.0 percentage points. In terms of raw student counts, the number reported as having internet access increased from 10,655 to 16,394, a difference of 5,739 students (Table 1).

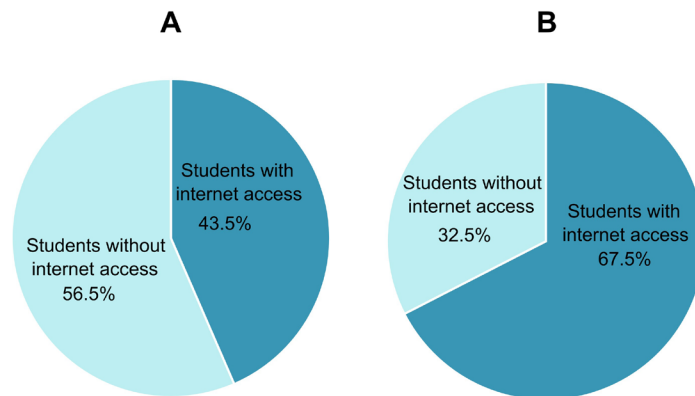


Figure 2: Student population surveyed. A. Students with/without internet access in 2020, B. Students with/without internet access in 2021. Percentages correspond to the unweighted mean of school-level connectivity percentages (see Methods).

Despite the restrictions of the pandemic, the student population recorded in the Student Enrolment System (SIMAT) remained

stable, from 23,474 students in 2020 to 23,509 in 2021, which represents a marginal variation of 0.1% (Vargas, 2021).

Variable	2020	2021
	Number of students	Number of students
Students with internet access	10,655	16,394
Students without internet access	12,819	7,115
TOTAL	23,474	23,509

Table 1: Students with broadband internet access vs. without connectivity

School Typology

The descriptive comparison identified three broad ranges of school connectivity for interpretative purposes. These ranges were defined post hoc based on the observed distribution of connectivity values and were used solely to facilitate descriptive comparison, rather than through a formal clustering or classification procedure. Schools with connectivity above 90% tended to have lower socioeconomic index values and were generally located in central urban areas. Schools with intermediate connectivity (approximately 50–80%) generally showed socioeconomic index values of roughly 25–40%, whereas schools with connectivity below 50% were more commonly located in peri-urban or rural areas and tended to serve more socioeconomically vulnerable populations (Figure 1, Table 2). The socioeconomic-index ranges were likewise used descriptively to characterise the observed patterns and do not represent statistically derived classification thresholds. In 2020, connectivity was concentrated in the north and centre of the city. CNSF, COBO, and IESE reported coverage above 80% and were located in areas with comparatively better urban infrastructure. By contrast, several schools in western and southeastern areas, including IESR, ITGR, and IERS, reported internet access below 20%, illustrating the substantial territorial variation in connectivity at the start of the remote-learning period (Figure 1, Table 2).

By 2021, connectivity had improved in many schools, although substantial differences remained. COBO and CNSF reported fewer than 3% of students without access, while several schools with intermediate baseline connectivity showed marked gains. ENLP increased from 59.6% in 2020 to 95.0% in 2021, and IEJS and ENST exceeded 70% connectivity. These changes

occurred during the same period in which devices, SIM cards, and network improvements were introduced (Table 4), although the present analyses do not establish that these interventions caused the observed changes.

Among the most vulnerable peripheral and rural schools, improvements were more limited and heterogeneous. Although ITGR increased from 6.5% to 75.0% connectivity, IESR reached 48.5%, while other highly vulnerable institutions remained at substantially lower levels. Overall, connectivity became more widespread by 2021, but low-connectivity areas persisted, particularly in the southwestern part of the municipality (Figure 1).

A strong inverse association was observed between the Socioeconomic Index and average school connectivity in both 2020 (Pearson's $r = -0.804$, $t(12) = -4.69$, $p = 0.00053$) and 2021 (Pearson's $r = -0.794$, $t(12) = -4.53$, $p = 0.00069$), corresponding to R^2 values of approximately 65% and 63%, respectively (Figure 3). The similarity of the coefficients across the two years indicates a persistent association between institutional socioeconomic vulnerability and connectivity coverage in this setting. A paired comparison showed that average school-level connectivity was higher in 2021 than in 2020 ($t(13) = 3.91$, $p = 0.0018$), with an average gain of 24.0 percentage points (95% CI [10.7, 37.2]). An OLS regression with year as a binary predictor produced the same estimated mean difference ($B = 24.0$ percentage points; cluster-robust SE = 6.12; CR2 bias-reduced linearisation with Satterthwaite degrees of freedom; $t(13) = 3.91$, $p = 0.0018$; 95% CI [10.8, 37.2]; $n = 28$ observations, 14 clusters). Given the small number of clusters, this regression is presented as supplementary evidence corroborating the paired-samples analysis rather than as an independent causal estimate.

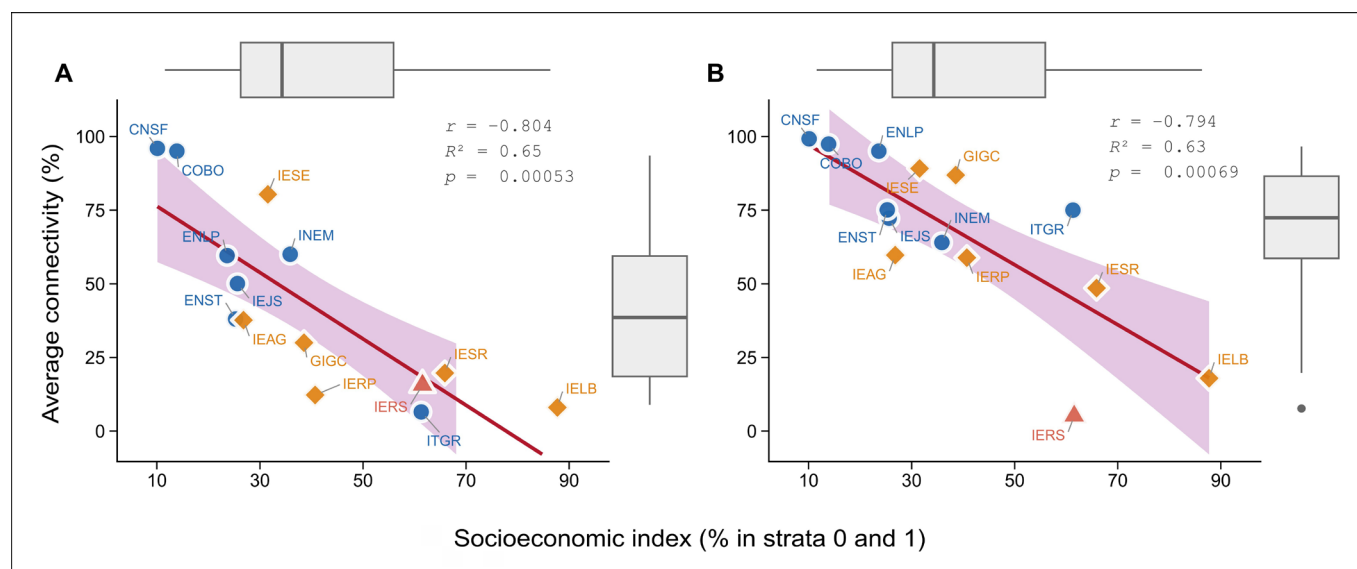


Figure 3: Pearson correlation between the Socioeconomic Index and average internet connectivity (%) across the 14 public schools of Tunja, Colombia, in (A) 2020 and (B) 2021. Each point represents one school, identified by its acronym (see Table 2) and classified by spatial location: urban (blue circles), peri-urban (amber diamonds), and rural (terracotta triangles). The red line represents the linear regression fit; the shaded band indicates the 95% confidence interval. Marginal boxplots display the univariate distribution of each variable. Statistical annotations report Pearson's correlation coefficient (r), coefficient of determination (R^2), and the associated p -value.

Public Schools	Code	Number of school sites	Spatial classification	Socioeconomic Average Percentage (Σ Stratum 0–1)	2020				2021			
					Students with internet access	Percentage %	Students without internet access	Percentage %	Students with internet access	Percentage %	Students without internet access	Percentage %
Colegio Nuestra Señora de Fátima	CNSF	1	Urban	10.1%	357	96.0%	15	4.0%	384	99.2%	3	0.8%
Colegio de Boyacá	COBO	3	Urban	13.9%	2,939	95.0%	154	5.0%	2,999	97.4%	79	2.6%
I.E. San Jerónimo Emiliani	IESE	3	Peri-urban	31.5%	1,191	80.4%	291	19.6%	1,482	89.2%	159	10.8%
INEM Carlos Arturo Torres	INEM	3	Urban	35.9%	1,432	60.0%	954	40.0%	2,386	64.0%	888	36.0%
Escuela Normal Superior Leonor Álvarez Pinzón	ENLP	1	Urban	23.6%	1,390	59.6%	943	40.4%	2,333	95.0%	117	5.0%
Institución Educativa Julius Sieber	IEJS	1	Urban	25.6%	361	50.1%	360	49.9%	721	72.3%	195	27.7%
Escuela Normal Superior Santiago de Tunja	ENST	2	Urban	25.3%	1,280	38.0%	2,085	62.0%	3,365	75.1%	856	24.9%
Institución Educativa Antonio José Sandoval Gómez	IEAG	3	Peri-urban	26.8%	400	37.7%	662	62.3%	1,062	59.7%	445	40.3%
Gimnasio Gran Colombiano	GIGC	1	Peri-urban	38.6%	266	30.0%	621	70.0%	887	86.9%	114	13.1%
Institución Educativa Silvino Rodríguez	IESR	7	Peri-urban	65.9%	540	19.7%	2,201	80.3%	2,741	48.5%	1,376	51.5%
Institución Educativa Rural del Sur	IERP	1	Rural	61.5%	147	15.6%	793	84.4%	940	5.2%	901	94.8%
Institución Educativa Rojas Pinilla	IERP	2	Peri-urban	40.7%	140	12.3%	1,003	87.7%	1,143	58.9%	441	41.1%
Institución Educativa Libertador Simón Bolívar	IELB	5	Peri-urban	87.7%	107	8.0%	1,227	92.0%	1,334	17.9%	1,157	82.1%
Instituto Técnico Gonzalo Suarez Rendón	ITGR	4	Urban	61.3%	105	6.5%	1,510	93.5%	1,615	75.0%	384	25.0%
TOTAL	--	37	--	--	10,655	43.5%	12,819	56.5%	23,474	67.5%	7,115	32.5%

Note: In the TOTAL row, the 'Students with internet access' and 'Students without internet access' columns represent the sum of raw student counts across the 14 schools, whereas the adjacent 'Percentage %' columns represent the unweighted mean of the corresponding school-level percentages. Therefore, these percentages do not correspond to the ratio of the total number of students with internet access to total enrolment.

Table 2: Socioeconomic index and connectivity of Public Schools

The device-access categories reported in Table 3 correspond directly to Section 3 of the classroom-level instrument (Table S2): laptop with/without internet, tablet, mobile phone with/without data, and ‘none of the above’, which teachers reported as percentages of their assigned student groups. These categories were mutually exclusive, with each student classified under a single predominant device-access condition. This exclusivity is supported by the fact that the six category percentages sum to approximately 100% in

both years (100.00% in 2020 and 99.99% in 2021, Table 3), with the minor deviation from 100% attributable to rounding. Access to internet-enabled devices increased between the two study years. The percentage of students reported as having laptops with internet access rose from 22.90% in 2020 to 26.86% in 2021, while the percentage with laptops without internet access increased from 9.09% to 16.34%. The proportion reported as using mobile phones with data plans increased from 24.40% to 26.59% (Table 3).

Year	Technological Means	N° students	%
2020	Laptops with internet	5,376	22.90%
	Laptops without internet	2,134	9.09%
	Tablet	444	1.89%
	Mobile phones with data	5,728	24.40%
	Mobile phones without data	6,398	27.26%
	None of the above	3,394	14.46%
2021	Laptops with internet	6,315	26.86%
	Laptops without internet	3,842	16.34%
	Tablet	673	2.86%
	Mobile phones with data	6,250	26.59%
	Mobile phones without data	4,082	17.36%
	None of the above	2,347	9.98%

Note: Each variable is measured over the total number of students for each year, equivalent to 23,474 students in 2020 and 23,509 students in 2021 according to SIMAT, with data provided by the Public Schools through the survey responses.

Table 3: Digital and audiovisual devices.

The proportion of students reported as having mobile phones without data plans decreased by 9.9 percentage points, while the proportion without any listed device decreased by approximately 4.5 percentage points. The SIM-card distribution programme reached 27.9% of enrolled students (Table 4). Nevertheless, rural areas continued to face substantial connectivity constraints related to infrastructure, device availability, and socioeconomic disadvantage.

Dynamics of pedagogical interaction in virtual environments

Among the communication channels used for remote teaching and learning, email reached 92.9% of enrolled students in 2021 compared with 62.0% in 2020. Videoconferencing showed the largest absolute increase, rising from 26.2% to 85.7%. Use of institutional virtual platforms increased by 26.5 percentage points, and social media by 23.1 percentage points (Figure 4), indicating broader use of digital communication channels during the second study year.

Frequently reported tools included WhatsApp, Google Meet, Google Classroom, Skype, and Zoom, which supported synchronous and asynchronous interaction. By contrast, platforms such as Facebook, Twitter, and Instagram were reported less frequently for structured educational communication.

Institutional strategies and provision of resources

Local and national authorities expanded technological provision during the study period. In total, 2,838 laptops were reported as delivered to the institutions, of which 1,073 were financed with local government funds and 1,765 were procured through the national programme ‘Computadores para Educar’

(Table 4). Reported laptop provision was distributed across schools with differing levels of socioeconomic vulnerability, including several rural and peri-urban institutions. However, this distribution is descriptive, and the present analysis does not test whether allocation was statistically proportional to institutional vulnerability.

Additionally, 6,567 SIM cards with mobile data were distributed as a measure to mitigate connectivity gaps, complemented by the installation of Wi-Fi zones and broadband expansion. These actions sought to ensure internet access in classrooms. Nevertheless, persistent limitations were identified in implementation, due to logistical delays, technical infrastructure constraints, and irregular coverage in certain areas. Such factors require priority attention to achieve continuous improvement in the provision of educational connectivity.

Finally, platforms promoted by the Ministry of Education of the national government, such as Colombia Aprende, offered more than 80,000 open-access digital resources for students, teachers, and families. The use of platforms such as Edmodo was also recorded, which facilitated the management of assignments, assessments, and content in a private and secure manner, promoting more structured and functional virtual learning environments (MEN, 2020).

DISCUSSION

The findings of this study suggest that the inequalities observed during the pandemic may have reflected pre-existing territorial and socioeconomic asymmetries in Tunja, rather than being entirely caused by the pandemic itself. However, as this study lacks a pre-pandemic baseline, it is not possible to empirically determine whether these inequalities predated

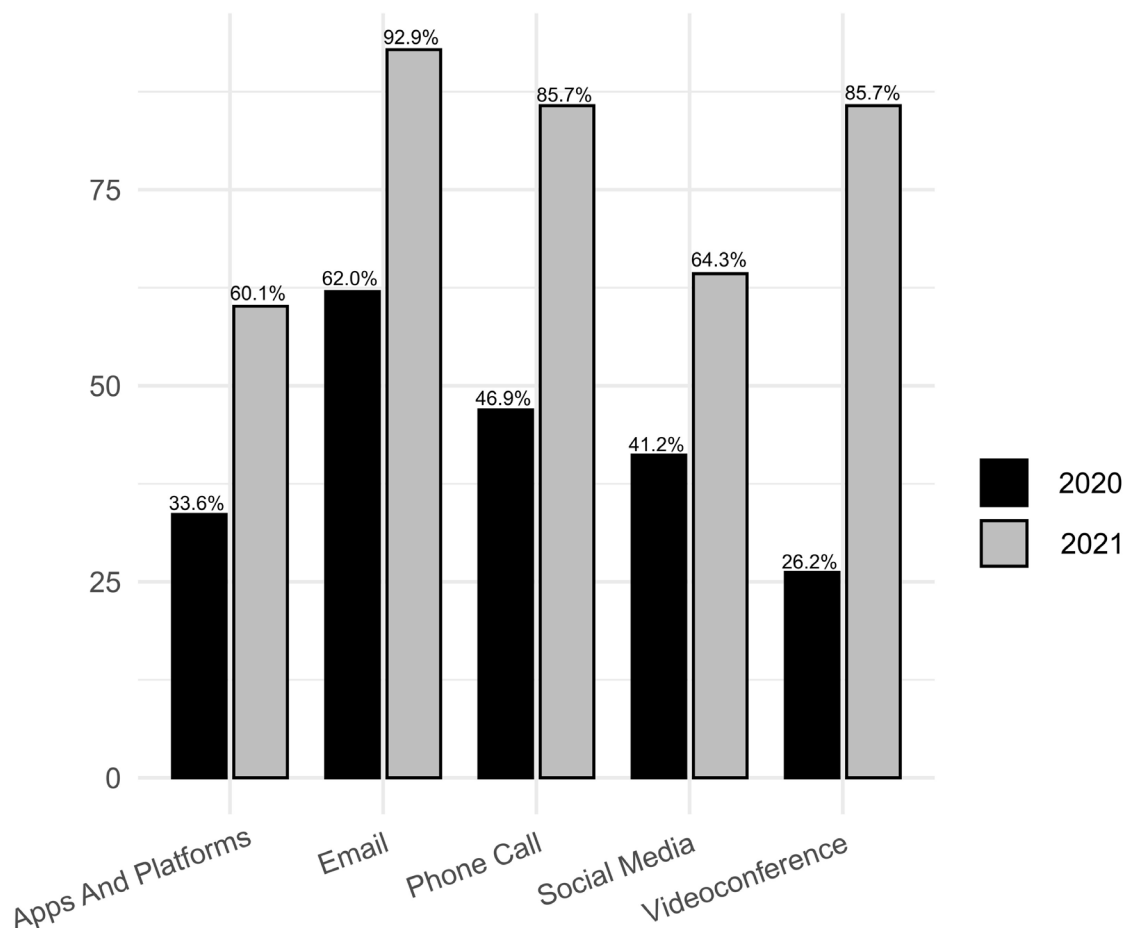


Figure 4: Percentage of virtual interaction media used by teachers in public schools. Note: Percentages represent the proportion of students reached through each communication channel relative to total enrolment for each year (23,474 in 2020; 23,509 in 2021).

Public Schools of Tunja	Number of laptops		Local government support		
	Inventory 2020	Inventory 2021	SIM Card Delivery	Laptops Delivery	TOTAL
Institución Educativa San Jerónimo Emiliani	NA	90	400	90	490
Instituto Técnico Gonzalo Suarez Rendón	120	160	700	75	775
Institución Educativa INEM Carlos Arturo Torres	57	406	500	349	849
Institución Educativa Julius Sieber	15	35	400	20	420
Institución Educativa Libertador Simón Bolívar	44	179	700	150	850
I.E. Rural del Sur	25	745	800	720	1,520
Normal Superior Leonor Álvarez Pinzón	65	309	500	244	744
Institución Educativa Silvino Rodríguez	30	55	767	75	842
Institución Educativa Rojas Pinilla	40	170	400	130	530
Colegio Gran Colombiano	21	186	500	165	665
Escuela Normal S. Santiago de Tunja	28	338	500	310	810
Institución Educativa Antonio José Sandoval Gómez	127	277	400	150	550
Colegio de Boyacá	NA	360	NA	360	360
Colegio Nuestra Señora de Fátima	NA	NA	NA	NA	NA
TOTAL	572	3,310	6,567	2,838	9,405

Table 4: Supply of laptops and delivery of SIM cards to Public Schools in 2020 and 2021.

the pandemic or were created or exacerbated by it. This interpretation is consistent with the growing international literature on the educational impacts of the pandemic, which documents how pre-existing socioeconomic and territorial

disparities shape unequal access to educational opportunities during periods of school disruption (Alejo, 2024; Wöbbekind et al., 2025). Although overall connectivity improved by 24.0 percentage points, this improvement was not proportional to

the magnitude of need in the most vulnerable sectors. This pattern is consistent with the inverse correlation observed between the socioeconomic index and average connectivity. While urban schools such as COBO and CNSF achieved over 97% connectivity in 2021, peripheral schools with low socioeconomic profiles, such as IELB and IERS, remained below 20%. According to Bracco et al. (2025), such disparities are not an isolated concern: accumulated learning losses in Latin America during the pandemic may have intergenerational implications for the formation of human capital among the most vulnerable populations.

For descriptive interpretation, schools were grouped into the three post hoc connectivity ranges described above, based on observed connectivity, socioeconomic index, and location. Two cases were particularly notable. IERS, the only rural school in the census, received the largest combined package of laptops and SIM cards reported in Table 4 (1,520 resources), yet its connectivity declined from 15.6% to 5.2%. IELB, which served the most socioeconomically vulnerable student population in the census (87.7% in strata 0 and 1), received 850 combined resources while connectivity increased from 8.0% to 17.9%. These contrasting patterns illustrate that greater resource provision did not invariably coincide with larger connectivity gains; however, the descriptive design does not establish the causal effect of resource allocation.

According to van Dijk (2005), the digital divide extends beyond the provision of devices and connectivity and also encompasses differences in the skills, conditions, and opportunities required to use digital resources effectively. The Tunja data are consistent with this broader perspective because technological provision did not uniformly coincide with improvements in connectivity. For example, IERS, the rural school with high socioeconomic vulnerability, declined from 15.6% to 5.2% connectivity despite receiving substantial technological resources. This pattern suggests that resource provision alone may not overcome structural territorial disadvantages. In settings characterised by unstable infrastructure, poor signal coverage, limited technical support, or constrained household capacity, devices may not translate directly into sustained educational access (Kennedy et al., 2022; Katz et al., 2025).

Mobile telephony played an important role in maintaining communication between schools and students. The SIM-card programme reached 27.9% of enrolled students. However, connectivity gains remained limited in some peri-urban and rural institutions, where intermittent signal coverage, limited access to suitable smartphones, and socioeconomic disadvantage may have constrained the usefulness of mobile-data provision. Across Latin America, 87% of the population lives within 4G coverage, yet mobile-internet penetration is substantially lower and fewer than half of households have fixed broadband access (WEF, 2021; World Bank, 2024). This distinction between nominal coverage and functional access is relevant to the persistent gaps observed in vulnerable educational settings in Tunja.

The marked expansion of digital communication channels between 2020 and 2021 is consistent with a substantial reconfiguration of teacher-student interaction during the pandemic. Videoconferencing showed the greatest

relative growth, suggesting an effort to reproduce elements of synchronous, real-time classroom interaction in remote settings. This is understandable given the limited availability of consolidated alternative models in basic education during the emergency period (Hodges et al., 2020; Bozkurt and Sharma, 2020). The simultaneous increase in email, telephone calls, and social media may also have functioned as a compensatory response for students unable to participate consistently in synchronous sessions. Muñoz-Najar et al. (2021) describe a comparable ‘staggered channel portfolio’, in which the choice of communication medium is constrained by the infrastructure available in students’ homes. These interpretations are consistent with the observed communication patterns but were not tested directly in the present dataset.

One notable development coinciding with the pandemic in public institutions in Tunja was the expansion of the digital educational ecosystem. Virtual platforms such as Colombia Aprende, Edmodo, Google Classroom and Google Meet, which had previously been underutilised, were widely adopted, ensuring the continuity of educational services. This coincides with a global trend of sustained growth in the interest of students, teachers and families in ICT-mediated learning resources (Radha et al., 2020; Arnett, 2021). Tunja is a UNESCO Learning City and, according to the Competitive Cities Index 2021 (CCI, 2021), the third Colombian city in terms of fixed broadband penetration and digital coverage. It has an institutional foundation upon which to consolidate this transformation, provided that efforts in vulnerable settings are strengthened in line with the objectives of Colombia’s National Digital Strategy 2023–2026. This strategy identifies reducing regional disparities and strengthening digital competencies in the education sector as national priorities (MinTIC, 2023).

From a normative perspective, the findings of this research make a valuable contribution to contemporary debates on the right to education in societies where there is a digital divide. The case study of Tunja showed that ensuring educational access in the digital era requires more than just the formal availability of technology. The convergence of pandemics, climate crises and geopolitical instability poses a constant and escalating threat to the educational continuity of the most vulnerable populations. The World Economic Forum’s 2024 Global Risks Report warns that, if commercial and geopolitical interests prevail over the public interest, the digital divide between high- and low-income countries will deepen. This will leave the most vulnerable communities isolated from digital advances impacting education, the economy, and health (WEF, 2024). Furthermore, the 2024 Report on the Sustainable Development Goals identifies the ongoing impact of the pandemic, escalating conflicts, geopolitical tensions and climate chaos as the main obstacles to educational progress worldwide (UN DESA, 2024). Thus, efficient educational inclusion hinges on the interplay of material infrastructure, institutional capacity, pedagogical mediation, and social conditions that facilitate meaningful participation in digital learning environments.

Recent research in Latin America argues that educational digital-inclusion policies should combine affordable infrastructure with contextualised services and mechanisms that support informed community participation, particularly among vulnerable groups

(Plaza de la Hoz, Espinosa Zárate and Camilli Trujillo, 2024). From this perspective, efficiency in resource allocation and accountability toward institutions facing greater structural disadvantage are complementary dimensions of educational governance (Reuge et al., 2021; Zapletal et al., 2023; Gottschalk and Weise, 2023). The present findings show that higher connectivity alone did not eliminate the socioeconomic gradient in access across schools. Reducing digital inequality therefore requires territorially sensitive and pedagogically grounded policies that can translate technological access into meaningful educational opportunities (Ramírez-Correa et al., 2025; Reuge et al., 2021; Mirazchiyski, 2025).

The findings also raise questions about Tunja's preparedness for future disruptions. In 2024, climate-related crises disrupted the education of 242 million children across 85 countries and territories, with low- and middle-income countries particularly affected (UNICEF, 2021; World Bank, 2024). Climate disasters, health emergencies, and conflicts can occur simultaneously, generating prolonged and compound crises that require educational systems to plan proactively (UNESCO, 2024; INEE, 2024). In this context, six areas warrant attention in Tunja: (i) universal, high-quality connectivity; (ii) sustained technological maintenance and territorial technical support; (iii) continuous digital-literacy programmes for teachers, students, and families; (iv) hybrid curricula designed in advance; (v) institutional monitoring capable of identifying inequalities at disaggregated territorial scales; and (vi) emerging inequalities in access to artificial-intelligence tools and capabilities. These priorities should be considered as policy implications rather than outcomes directly tested in the present study.

Several limitations should be considered. Although the census design achieved 100% institutional coverage, the analysis included only 14 school-level units, limiting precision and the stability of inferential estimates. The use of paired analyses and CR2-adjusted cluster-robust standard errors addresses the repeated-measures structure but does not eliminate the limitations associated with a small number of institutions. Connectivity was used as a proxy for educational access; however, it does not capture effective participation, quality of engagement, or academic achievement. The analysis also did not disaggregate outcomes by gender, despite the relevance of gender to digital vulnerability in Latin America (Ancheta-Arrabal et al., 2021). Finally, the two-year observation window captures change during a specific emergency period but cannot establish the sustainability of the observed gains. Longer

follow-up and replication in municipalities with different levels of institutional capacity are therefore warranted.

CONCLUSIONS

Against the backdrop of the abrupt transition to remote education during the COVID-19 pandemic, this study examined the inverse association between socioeconomic vulnerability and internet access across 14 public schools in Tunja in 2020 and 2021. The findings indicate that institutions serving more socioeconomically vulnerable student populations consistently had lower connectivity levels. Although connectivity improved substantially on average between the two years, marked differences persisted across schools, and greater reported technological provision did not invariably coincide with greater connectivity gains. These findings are consistent with regional evidence that structural interventions must extend beyond device provision if digital inequalities are to be reduced sustainably (Bracco et al., 2025).

Future research should combine connectivity data and local socioeconomic characteristics with standardised learning outcomes produced annually by the Colombian Institute for the Evaluation of Education (ICFES) to determine whether gains in digital access correspond to reductions in inequalities in academic achievement. Extending the observation period beyond the pandemic, incorporating gender-disaggregated analyses, and replicating the methodology in municipalities with lower institutional capacity would strengthen understanding of these patterns across contexts of greater structural disadvantage. Qualitative research capturing the experiences of teachers, students, and families would also provide the contextual depth needed to determine how improvements in connectivity translate into meaningful educational opportunities.

ACKNOWLEDGMENT

We extend our gratitude to the Municipal Administration of Tunja (2020–2023), to Dr. José Moreno Villamil, Secretary of Education during this period; to Julieta Bernal Moreno, from the ICT Quality Division of the Municipal Secretariat of Education, for her support during the data collection and consolidation process; to the staff of the Secretariat of Education; and to the principals of the schools, teachers, and students who supported this research. Additionally, we acknowledge the valuable assistance of ChatGPT-4 in improving the English translation and readability, thereby facilitating understanding for the readers of this manuscript.

REFERENCES

- Alejo, A., Jenkins, R. and Yao, H. (2024) 'Learning losses during the COVID-19 pandemic: Understanding and addressing increased learning disparities', *Future in Educational Research*, Vol. 2, No. 1, pp. 16–29. <https://doi.org/10.1002/fer3.21>
- Ancheta-Arrabal, A., Pulido-Montes, C. and Carvajal-Mardones, V. (2021) 'Gender Digital Divide and Education in Latin America: A Literature Review', *Education Sciences*, Vol. 11, No. 12, p. 804. <https://doi.org/10.3390/educsci11120804>
- Arnett, T. (2021) *Breaking the mold: How a global pandemic unlocks innovation in K–12 instruction*, Lexington, MA: Clayton Christensen Institute for Disruptive Innovation. Available at: <https://www.christenseninstitute.org/publication/online-learning-survey/>
- Banco Mundial (2020) *Índice de Gini — Colombia* [dataset]. Available at: <https://datos.bancomundial.org/indicador/SI.POV.GINI> (Accessed: 1 August 2025).

- Becerra, M. and Mastrini, G. (2017) *La concentración infocomunicacional en América Latina (2000–2015): Nuevos medios y tecnologías, menos actores*, Bernal: Universidad Nacional de Quilmes and Observacom.
- Bonilla, J., López, D. and Sepúlveda, C. E. (2014) 'Estratificación socioeconómica en Colombia: Contexto general y diagnóstico', in Sepúlveda-Rico, C. E., López-Camacho, D. and Gallego-Acevedo, J. M. (eds.) *Los límites de la estratificación: En busca de alternativas*, Bogotá, D.C.: Editorial Universidad del Rosario, pp. 23–39. <https://doi.org/10.7476/9789587385373.0004>
- Bozkurt, A. and Sharma, R. C. (2020) 'Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic', *Asian Journal of Distance Education*, Vol. 15, No. 1. <https://doi.org/10.5281/zenodo.3778083>
- Bracco, J., Ciaschi, M., Gasparini, L., Marchionni, M. and Neidhöfer, G. (2025) 'The Impact of COVID-19 on Education in Latin America: Long-Run Implications for Poverty and Inequality', *Review of Income and Wealth*, Vol. 71, No. 1, p. e12687. <https://doi.org/10.1111/roiw.12687>
- Busuttill, L. and Farrugia, R. C. (2020) 'Teachers' Response to the Sudden Shift to Online Learning during COVID-19 Pandemic: Implications for Policy and Practice', *Malta Review of Educational Research*, Vol. 14, No. 2, pp. 211–241.
- Califano, B. and Becerra, M. (2024) 'Policy interventions to address digital inequalities in Latin America in the face of the pandemic', in Yates, S. and Carmi, E. (eds.) *Digital inclusion: International policy and research*, Cham: Palgrave Macmillan, pp. 39–61. https://doi.org/10.1007/978-3-031-28930-9_3
- Consejo Privado de Competitividad and Universidad del Rosario (2021) *Índice de Competitividad de Ciudades 2021*, Bogotá, D.C.: Consejo Privado de Competitividad and SCORE–Universidad del Rosario.
- Colegio de Boyacá (2020) *Historia*, [Online]. Available at: <https://www.colboy.edu.co/historia/> [Accessed 18 October 2021].
- de Clercq, M., D'Haese, M. and Buysse, J. (2023) 'Economic growth and broadband access: The European urban-rural digital divide', *Telecommunications Policy*, Vol. 47, No. 6, 102579. <https://doi.org/10.1016/j.telpol.2023.102579>
- Departamento Nacional de Planeación (2024) *Estrategia Nacional Digital de Colombia 2023–2026*, Bogotá D.C.: Departamento Nacional de Planeación, [Online]. Available at: https://www.mintic.gov.co/portal/715/articles-334120_recurso_1.pdf [Accessed 16 September 2026].
- Ente Nacional de Comunicaciones (2020) 'El Gobierno nacional congela los precios de telefonía fija y móvil, internet y de la TV paga', *ENACOM*, 18 May, [Online]. Available at: https://www.enacom.gob.ar/institucional/el-gobierno-nacional-congela-las-tarifas-de-telefonía-fija-y-movil--internet-y-de-la-tv-paga_n2365 [Accessed 19 July 2024].
- GADRRRES — Global Alliance for Disaster Risk Reduction and Resilience in the Education Sector (2025) *Global Status of School Safety: Key Findings from the 2024 Comprehensive School Safety Policy Survey*, [Online]. Available at: <https://www.unesco.org/en/articles/global-policy-brief-reveals-gaps-school-safety-impacting-330-million-children> [Accessed 1 March 2026].
- Galeano, S. (2020) 'El número de usuarios de internet en el mundo crece un 7% y alcanza los 4.540 millones (2020)', *Marketing4eCommerce*, [Online]. Available at: <https://marketing4ecommerce.net/usuarios-internet-mundo/> [Accessed 10 May 2020].
- Gonzales, S. (2024) 'AI literacy and the new digital divide: A global call for action', *UNESCO Global AI Ethics and Governance Observatory*, [Online]. Available at: <https://www.unesco.org/ethics-ai/en/articles/ai-literacy-and-new-digital-divide-global-call-action> [Accessed 1 August 2025].
- Gottschalk, F. and Weise, C. (2023) 'Digital equity and inclusion in education: An overview of practice and policy in OECD countries', *OECD Education Working Papers*, No. 299, Paris: OECD Publishing. <https://doi.org/10.1787/7cb15030-en>
- Haelermans, C., Korthals, R., Jacobs, M., de Leeuw, S., Vermeulen, S., van Vugt, L., Aarts, B., Prokic-Breuer, T., van der Velden, R., van Wetten, S. and de Wolf, I. (2022) 'Sharp increase in inequality in education in times of the COVID-19-pandemic', *PLOS ONE*, Vol. 17, No. 2, e0261114. <https://doi.org/10.1371/journal.pone.0261114>
- Hodges, C., Moore, S., Lockee, B., Trust, T. and Bond, M.A. (2020) 'The difference between emergency remote teaching and online learning', *EDUCAUSE Review*, 27 March, [Online]. Available at: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning> [Accessed 1 August 2025].
- INEE — Inter-agency Network for Education in Emergencies (2024) *Minimum Standards for Education: Preparedness, Response, Recovery*, 3rd edn, New York: INEE, [Online]. Available at: <https://inee.org/minimum-standards> [Accessed 1 August 2025].
- Joshi, D.R., Khanal, J. and Dhakal, R.H. (2023) 'From resistance to resilience: Teachers' adaptation process to mediating digital devices in pre-COVID-19, during COVID-19, and post-COVID-19 classrooms in Nepal', *Education Sciences*, Vol. 13, No. 5, p. 509. <https://doi.org/10.3390/educsci13050509>
- Kairuz, J. (2020) 'La falta de conectividad está aumentando las brechas sociales del país', *El Nuevo Día*. Available at: <http://www.elnuevodia.com.co/nuevodia/especiales/sucesos/448477-la-falta-conectividad-esta-aumentando-las-brechas-sociales-del-pais> (Accessed: 10 May 2020).
- Katz, V.S., Jordan, A.B. and Ognyanova, K. (2025) 'Digital inequalities and U.S. undergraduate outcomes over the first two years of the COVID-19 pandemic', *PLOS ONE*, Vol. 20, No. 3, e0319000. <https://doi.org/10.1371/journal.pone.0319000>
- Kennedy, A.I., Mejía-Rodríguez, A.M. and Strello, A. (2022) 'Inequality in remote learning quality during COVID-19: Student perspectives and mitigating factors', *Large-scale Assessments in Education*, Vol. 10, p. 29. <https://doi.org/10.1186/s40536-022-00143-7>
- Laboratorio de Economía de la Educación de la Pontificia Universidad Javeriana (2020) *Informe 01: ¿Están preparados los colegios públicos del país para implementar lecciones virtuales ante una eventual suspensión de clases a causa del coronavirus?* Available at: <https://lee.javeriana.edu.co/documents/5581483/7093554/INFORME-1-INTERNET-COMPUTADORES-LEE-JAVERIANA.pdf>.
- Ministerio de Educación Nacional de Colombia (2020) *Gobierno nacional adelanta acciones para avanzar en la conectividad de los estudiantes a las plataformas educativas que hacen parte de la estrategia Aprender Digital*. Available at: <https://www.mineducacion.gov.co/1759/w3-article-396260.html> (Accessed: 24 October 2021).
- Ministerio de Educación Nacional de Colombia (n.d.) *Sistema de Matrícula Estudiantil de Educación Básica y Media (SIMAT)*, [Online]. Available at: <https://www.sistemamatriculadas.gov.co/simat/app> [Accessed 14 July 2020].

- Ministerio de Tecnologías de la Información y las Comunicaciones (n.d.) *Colombia TIC: Estadísticas del sector TIC*, [Online database]. Available at: <https://colombiatic.mintic.gov.co/> [Accessed 1 September 2024].
- Mirazchiyski, P.V. (2025) 'Contemporary gaps in research on digital divide in education: A literature review', *Universal Access in the Information Society*, Vol. 24, No. 2, pp. 991–1008. <https://doi.org/10.1007/s10209-024-01166-3>
- Montes Cadavid, J.C. (2019) 'Colombia, ¿país sin conectividad?', *La República*, 28 September, [Online]. Available at: <https://www.larepublica.co/analisis/juan-carlos-montes-cadavid-2551471/colombia-pais-sin-conectividad-2914902> [Accessed 10 May 2020].
- Morata Sanz, J.I. (2020) 'Uso de TIC en orientación educativa en tiempos de COVID-19', *Revista de Orientación Educativa AOSMA*, Special Issue No. 28, pp. 88–91. Available at: <https://dialnet.unirioja.es/servlet/articulo?codigo=7381638> [Accessed 16 September 2026].
- Moreno-Villamil, J.A. (2023) 'Plan de Alimentación Escolar y su incidencia en la calidad educativa de Tunja, Boyacá', *Praxis & Saber*, Vol. 14, No. 39, p. e15973. <https://doi.org/10.19053/22160159.v14.n39.2023.15973>
- Muñoz-Najar, A., Gilberto, A., Hasan, A., Cobo, C., Azevedo, J.P. and Akmal, M. (2021) *Remote learning during COVID-19: Lessons from today, principles for tomorrow*, Washington, D.C.: World Bank Group. <https://doi.org/10.1596/36665>
- Olson, K., Phillips, A., Smyth, J.D. and Stenger, R. (2025) 'The urban–rural digital divide in internet access and online activities during the COVID-19 pandemic', *Rural Sociology*, Vol. 90, No. 3, p. e70012. <https://doi.org/10.1111/ruso.70012>
- Pardo-Figueroa, H.M. and Obando-Fernández, J.A. (2021) *Análisis del proceso de implementación de redes 4G en Colombia dentro del plan de transición a nuevas tecnologías*, Bachelor's thesis, Bogotá D.C.: Universidad Nacional Abierta y a Distancia, [Online]. Available at: <https://repository.unad.edu.co/handle/10596/42461> [Accessed 16 September 2026].
- Plaza de la Hoz, J., Espinosa Zárate, Z. and Camilli Trujillo, C. (2024) 'Digitalisation and poverty in Latin America: A theoretical review with a focus on education', *Humanities and Social Sciences Communications*, Vol. 11, p. 1194. <https://doi.org/10.1057/s41599-024-03692-0>
- Pustejovsky, J.E. and Tipton, E. (2018) 'Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models', *Journal of Business & Economic Statistics*, Vol. 36, No. 4, pp. 672–683. <https://doi.org/10.1080/07350015.2016.1247004>
- Radha, R., Mahalakshmi, K., Sathish Kumar, V. and Saravanakumar, A.R. (2020) 'E-Learning during Lockdown of COVID-19 Pandemic: A Global Perspective', *International Journal of Control and Automation*, Vol. 13, No. 4, pp. 1088–1099.
- Ramírez-Correa, P. E., Mariano, A. M. and Santos, M.R. (2025) 'Digital and Sustainable Education and Social Inclusion: A Bibliometric Review with the Consolidated Meta-Analytical Approach', *Sustainability*, Vol. 17, No. 13, p. 5677. <https://doi.org/10.3390/su17135677>
- Reuge, N., Jenkins, R., Brossard, M., Soobrayan, B., Mizunoya, S., Ackers, J., Jones, L. and Taulo, W. G. (2021) 'Education response to COVID 19 pandemic, a special issue proposed by UNICEF: Editorial review', *International Journal of Educational Development*, Vol. 87, p. 102485. <https://doi.org/10.1016/j.ijedudev.2021.102485>
- Rivera Reyes, F. Á. (2013) 'El origen de la educación en la ciudad de Tunja: La historia de un derecho', *Revista Principia Iuris*, Vol. 19, No. 19, pp. 215–224.
- United Nations (2020) 'Uno de cada tres niños en el mundo no puede acceder a clases a distancia si su escuela cierra', *Noticias ONU*, [Online]. Available at: <https://news.un.org/es/story/2020/08/1479572> [Accessed 1 August 2025].
- United Nations Department of Economic and Social Affairs (2024) *The Sustainable Development Goals Report 2024*, New York: United Nations. Available at: <https://unstats.un.org/sdgs/report/2024/> (Accessed: 1 August 2025).
- UNICEF (2021) *The Climate Crisis Is a Child Rights Crisis: Introducing the Children's Climate Risk Index*, New York: UNICEF, [Online]. Available at: <https://www.unicef.org/reports/climate-crisis-child-rights-crisis> [Accessed 1 August 2025].
- van Dijk, J. A. G. M. (2005) *The Deepening Divide: Inequality in the Information Society*, Thousand Oaks, CA: SAGE Publications.
- Vargas, F. (2021) *Sistema de matrícula estudiantil: Tunja, septiembre de 2021*, Tunja: Alcaldía Mayor de Tunja.
- Wang, J. and Lin, P.-C. (2024) 'Spatial Spillover of the Global Internet Penetration Rate and the Digital Gender Divide', *SAGE Open*, Vol. 14, No. 2, p. 21582440241255448. <https://doi.org/10.1177/21582440241255448>
- WEF — World Economic Forum (2021) 'Less than 50% of Latin America has fixed broadband. Here are 3 ways to boost the region's digital access', *World Economic Forum*, [Online]. Available at: <https://www.weforum.org/stories/2021/07/latin-america-caribbean-digital-access/> [Accessed 1 August 2025].
- World Economic Forum (2024) *The global risks report 2024*, 19th edition, Geneva: World Economic Forum, [Online]. Available at: <https://www.weforum.org/publications/global-risks-report-2024/> [Accessed 1 August 2025].
- Wöbbekind, L., Weihe, J., Bóte-Vericad, J.-J. and Mandl, T. (2025) 'Experiences With Online Teaching During COVID-19 in Latin America: Digital Divide, Digital Literacy, and Lessons Learned', *Education for Information*, Vol. 41, No. 4, pp. 302–317. <https://doi.org/10.1177/01678329251340465>
- World Bank (2024) *World development report 2024: The middle-income trap*, Washington, DC: World Bank. <https://doi.org/10.1596/978-1-4648-2078-6>
- Zapletal, F., Hudec, M., Švaňa, M., Chytilová, L., Hlaváček, K., Lokaj, A., Urbanek, A., Glova, J., Samartinho, J. P., Rodriguez, C. M. C. and Guðnason, S. (2023) 'The Complex Evaluation of the Impact of COVID-19 Pandemic at Universities: A Soft Computing Approach', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 16, No. 3, pp. 231–244. <https://doi.org/10.7160/eriesj.2023.160307>

APPENDIX

MAYOR'S OFFICE OF TUNJA – MUNICIPAL EDUCATION OFFICE

Institutional Survey on ICT Capacity, Student Connectivity, and Educational Strategies During the COVID-19 Pandemic

SECTION 1 – GENERAL INFORMATION

Name of Educational Institution:

Municipal jurisdiction? ☐ Yes ☐ No

Number of campuses:

Number of students:

Address and neighbourhood of the institution and campuses:

SECTION 2 – TEACHER ICT SKILLS

Percentage of teachers with management of ICT skills:

ICT tools used by teachers (mark all that apply):

☐ Facebook

☐ Email

☐ Storage on cloud

☐ Virtual platforms

☐ Phone calls

☐ Physical guides

☐ None

SECTION 3 – STUDENT INTERNET CONNECTIVITY

Percentage of students with internet coverage — Broadband:

Cell phone with internet — Data
mobiles %

Percentage of students without internet access:

SECTION 4 – TECHNOLOGICAL TOOLS AVAILABLE TO STUDENTS

Percentage of technological tools available to students (mark and indicate %):

☐ Storage on cloud _____ %

☐ Virtual platforms _____ %

☐ WhatsApp _____ %

☐ Phone calls _____ %

☐ Email _____ %

☐ Physical guides _____ %

☐ None _____ %

SECTION 5 – INSTITUTIONAL COMMUNICATION

Communication channel:

Circular:

SECTION 6 – CONNECTIVITY PROBLEMS AND STRATEGIES

Identify the main problems of student internet access:

Initiatives proposed to solve these problems:

Strategies that can be used to assess and monitor connectivity improvements:

Note: This form was used to collect information on connectivity, access to ICT tools, and institutional strategies during the COVID-19 pandemic in public schools of the municipality of Tunja.

Table S1: Blank template of the institutional data collection instrument administered to public school principals in the municipality of Tunja regarding connectivity, ICT tools, and institutional strategies during the COVID-19 pandemic.

MAYOR'S OFFICE OF TUNJA – MUNICIPAL EDUCATION OFFICE
Teacher Survey on Student Connectivity, Digital Device Availability, and Pedagogical Communication Strategies During the COVID-19 Pandemic

SECTION 1 – GENERAL INFORMATION

Teacher's Name: _____

Educational Institution: _____

Does the school have internet access? ☐ Yes ☐ No

Location: ☐ Urban ☐ Peri-urban
☐ Rural

SECTION 2 – STUDENT INTERNET CONNECTIVITY

Number of students WITH internet access: _____ students

Number of students WITHOUT internet access: _____ students

with internet: _____ % Mobile data _____ %

without internet: _____ % None of the above _____ %

SECTION 3 – DIGITAL DEVICES AVAILABLE TO STUDENTS

Laptop with internet _____ % Mobile phone with data _____ %

Laptop without internet _____ % Mobile phone without data _____ %

Tablet _____ % None of the above _____ %

SECTION 4 – PEDAGOGICAL COMMUNICATION TOOLS USED
(Mark and indicate % of students reached)

☐ Email _____ %

☐ Virtual learning platforms (e.g., Google Classroom, Moodle) _____ %

☐ Videoconferencing (e.g., Zoom, Google Meet) _____ %

☐ Social media (e.g., WhatsApp, Facebook) _____ %

☐ Telephone calls _____ %

SECTION 5 – MAIN PROBLEMS RELATED TO CONNECTIVITY
(Briefly describe)

SECTION 6 – RESOURCE PROVISION

Number of laptops delivered: _____ units

Number of SIM cards with data provided: _____ units

Additional connectivity actions: _____

Note: This form was used to assess students' connectivity, availability of digital devices, and use of pedagogical communication tools in public schools of the municipality of Tunja during the COVID-19 pandemic.

Table S2: Blank template of the classroom-level administrative tracking instrument internally managed by school principals to evaluate classroom teachers' student connectivity metrics, digital device availability, and pedagogical communication tools.

UNDERSTANDING LMS ACCEPTANCE IN VOCATIONAL SECONDARY EDUCATION: EXTENDING THE TECHNOLOGY ACCEPTANCE MODEL USING SEM-GSCA

Lifa Farida Panduwinata✉
Prisilia Joyceline Atmojo
Febrika Yogie Hermanto

Universitas Negeri Surabaya, Indonesia

✉ ifapanduwinata@unesa.ac.id

ABSTRACT

This study examines vocational secondary school students' acceptance of Learning Management Systems (LMS) by extending the Technology Acceptance Model (TAM) with external factors, including computer self-efficacy (CSE), subjective norm (SN), perceived enjoyment (PE), perceived accessibility (PA), and perceived playfulness (PP). A quantitative cross-sectional survey was conducted with 252 students at one vocational secondary school in Surabaya, Indonesia, and the data were analyzed using SEM-GSCA. The findings showed that PA was the only external factor significantly associated with PEOU, whereas the other external variables showed no significant associations with PU or PEOU. Within the TAM framework, PEOU was significantly associated with PU and ATU, whereas PU emerged as the strongest predictor of both ATU and BI. BI also significantly predicted actual LMS use. In this sample, LMS acceptance was more strongly associated with accessibility and functional utility than with motivational or hedonic factors. The study extends the application of TAM to vocational secondary education and suggests that schools and LMS developers should prioritize accessibility, usability, and pedagogical relevance to support sustainable LMS adoption.

KEYWORDS

Learning Management System, SEM-GSCA, Technology Acceptance Model, Vocational Secondary Education, Post-pandemic Education

HOW TO CITE

Panduwinata L. F., Atmojo P. J., Hermanto F. Y. (2026) 'Understanding LMS Acceptance in Vocational Secondary Education: Extending the Technology Acceptance Model Using SEM-GSCA', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 237–253. <http://dx.doi.org/10.7160/eriesj.2026.190305>

Article history

Received

August 13, 2025

Received in revised form

May 12, 2026

Accepted

August 21, 2026

Available on-line

September 30, 2026

Highlights

- LMS acceptance in vocational education is shaped mainly by usefulness and accessibility.
- Accessible LMS platforms support more efficient access to vocational learning activities.
- Perceived usefulness strengthens students' attitudes and continued LMS engagement.
- Responsible LMS implementation requires pedagogically meaningful digital integration.
- Sustainable LMS adoption supports learning continuity and competency-based education.

INTRODUCTION

The integration of digital technology into education has transformed teaching and learning practices across educational levels. Learning Management Systems (LMS) are among the most widely adopted educational technologies and facilitate online learning activities through web-based platforms that support content delivery, communication, assessment, and learning management (Snoussi, 2019; Al-Fraihat et al., 2020). LMS platforms enable students and teachers to engage in asynchronous and remote learning activities, thereby increasing flexibility and accessibility in educational processes.

The implementation of LMS has been associated with several educational benefits, including improved learning efficiency, increased learner autonomy, and enhanced student engagement (Unlu and Suel, 2014). Previous studies have also reported that students tend to demonstrate more positive learning attitudes when using e-learning platforms due to their perceived usefulness, convenience, and accessibility (Davis, 1989; Mohammadi, 2015). Consequently, LMS adoption has become an important area of research, particularly within the context of technology-enhanced learning environments. Despite the growing body of literature on LMS acceptance, previous findings remain inconsistent regarding the factors

influencing students' acceptance and continued use of LMS platforms. Several studies have reported significant associations between external variables such as self-efficacy, accessibility, enjoyment, and social influence and students' perceptions of LMS usefulness and ease of use (Abdullah and Ward, 2016; Salloum et al., 2019). However, other studies have reported weak or insignificant relationships between these external factors and core Technology Acceptance Model (TAM) constructs (Surjono, Muhtadi and Trilisiana, 2019; Llamas, Vilela-Malabanan and Dinawanao, 2025; Rogers, Mercado and Decano, 2025). These inconsistencies suggest that the relevance of external factors may depend on contextual characteristics, including educational level, learning environment, and patterns of technology usage.

Most prior studies on LMS acceptance have focused primarily on university students, whereas comparatively less attention has been given to vocational secondary school students (Benson et al., 2020; Ahmad et al., 2023). This limitation is important because vocational students have distinct educational characteristics, including practice-oriented learning approaches, task-based instructional activities, and different patterns of digital technology utilization (Sollner et al., 2018; Im, 2021). In vocational education contexts, LMS use may be shaped more strongly by functional and task-related considerations than by psychological or affective factors (Ahmad et al., 2023). Therefore, further investigation is needed to determine whether commonly used external TAM variables remain relevant in vocational secondary education settings. Unlike previous TAM studies that primarily emphasize university students and generalized e-learning environments, this study examines whether external TAM variables remain applicable in vocational secondary education, where learning activities are more practice-oriented and task-driven.

To address this gap, the present study examines LMS acceptance among students at one vocational secondary school in Surabaya, Indonesia using an extended Technology Acceptance Model (TAM). Specifically, this study investigates the influence of external factors, namely computer self-efficacy (CSE), subjective norm (SN), perceived enjoyment (PE), perceived accessibility (PA), and perceived playfulness (PP), on students' perceptions of LMS use (Abdullah and Ward, 2016; Salloum et al., 2019). By examining how students evaluate LMS accessibility, usefulness, and continued use, this study provides insight into factors associated with digital learning acceptance and their implications for efficient and responsible technology integration in vocational education. This study is particularly relevant to efficiency and responsibility in vocational learning because LMS implementation should not only facilitate access to learning activities but also ensure that digital platforms meaningfully support students' academic tasks, learning continuity, and competency-oriented educational needs (Snoussi, 2019; Al-Fraihat et al., 2020). Thus, the study contributes to the contextual application of TAM while offering practical implications for strengthening sustainable LMS implementation in vocational secondary education.

LITERATURE REVIEW

Technology Acceptance Model (TAM) and LMS Acceptance

The TAM, developed by Davis (1989), is one of the most widely used frameworks for explaining technology acceptance behavior. TAM emphasizes two primary determinants of technology adoption: PU and PEOU. PU refers to users' beliefs that technology improves performance, while PEOU reflects the extent to which technology is perceived as easy to use (Davis, 1989).

TAM has been extensively applied in e-learning and learning management system (LMS) research (Lee, Yoon and Lee, 2009; Abdullah and Ward, 2016; Salloum et al., 2019). LMS platforms facilitate learning activities such as content delivery, assessment, communication, and class management through synchronous and asynchronous systems (Al-Adwan et al., 2021). Previous studies generally confirm that TAM effectively explains LMS acceptance in educational settings (Al-Gahtani, 2016; Hanif, Jamal and Imran, 2018; Alshammari, 2020; Gurban and Almogren, 2022). However, the influence of external factors on PU and PEOU remains inconsistent across contexts (Llamas, Vilela-Malabanan and Dinawanao, 2025; Rogers, Mercado and Decano, 2025), indicating that LMS acceptance depends on specific educational environments, including vocational education.

Therefore, this study extends TAM by incorporating five external variables frequently used in e-learning acceptance research: CSE, SN, PE, PA, and PP (Abdullah and Ward, 2016; Salloum et al., 2019) (see Figure 1). The inclusion of these variables enables a more comprehensive understanding of LMS acceptance among vocational secondary school students.

Computer Self-efficacy (CSE)

CSE reflects individuals' confidence in their ability to use digital technologies effectively to accomplish specific tasks (Downey and Kher, 2015). In LMS environments, students with higher self-efficacy tend to perceive learning systems as easier to operate and more beneficial for completing academic activities (Alshammari, 2020; Gurban and Almogren, 2022; Rokhim et al., 2024). Prior studies consistently report positive relationships between CSE, PU, and PEOU in e-learning contexts (Salloum et al., 2019; Arief et al., 2021; Espinosa, Ruiz and Mercado, 2021). Accordingly, the following hypotheses are proposed:

- H1: CSE positively influences PU in LMS use
- H2: CSE positively influences PEOU in LMS use

Subjective Norm (SN)

SN refers to the perceived social influence that encourages individuals to perform or avoid a particular behavior based on the expectations or opinions of significant others, such as teachers, peers, and the surrounding social environment (Ajzen, 1991; Venkatesh and Davis, 2000; Jimenez et al., 2021). In educational technology contexts, SN reflects the extent to which students perceive that important individuals believe they should use an LMS.

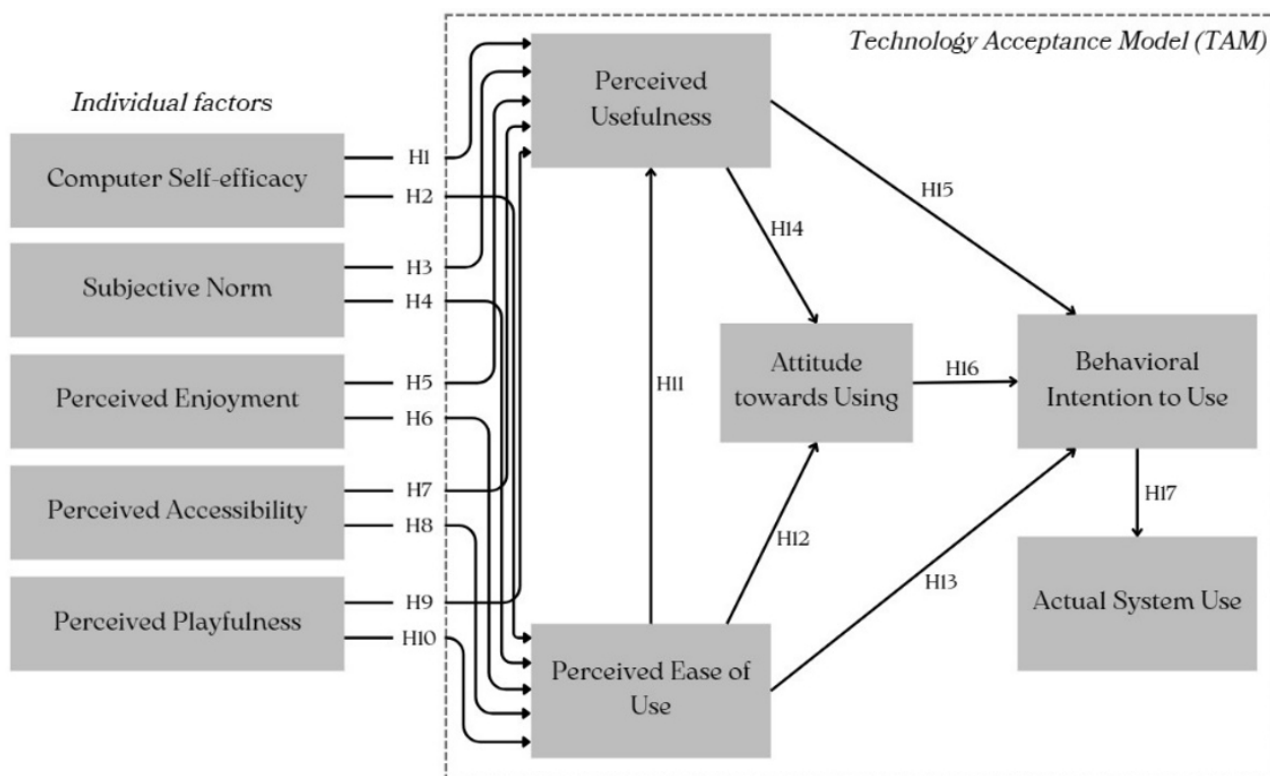


Figure 1: Conceptual LMS Acceptance Model

Prior studies suggest that social influence can shape students' perceptions regarding both the PU and PEOU of e-learning technologies. When students receive positive recommendations or encouragement from peers and instructors, they are more likely to perceive LMS platforms as beneficial and easier to operate (Al-Gahtani, 2016; Khan et al., 2020). Empirical evidence also indicates that subjective norm positively affects PU and PEOU in e-learning environments (Ching-Ter, Hajiyeve and Su, 2017; Revyathi and Tselios, 2019). In vocational education settings, where collaborative learning and teacher guidance remain important, social influence may play a significant role in shaping students' acceptance of LMS use. Accordingly, the following hypotheses are proposed:

- H3: SN positively influences PU in LMS use
- H4: SN positively influences PEOU in LMS use

Perceived Enjoyment (PE)

PE refers to the extent to which individuals perceive the use of a technology system as enjoyable and intrinsically satisfying, regardless of its functional or performance-related outcomes (Venkatesh and Davis, 2000). In e-learning environments, PE reflects students' positive emotional experiences when interacting with an LMS (Cahyanti et al., 2025; Rohana, Nugraha and Koblianska, 2025). Technologies that provide enjoyable and engaging learning experiences are more likely to encourage students' acceptance and continued use (Sorensen, 2019).

Previous studies have consistently reported that PE positively influences students' perceptions of both the PU and PEOU of e-learning systems. Students who experience enjoyment while using LMS platforms tend to perceive the systems as more accessible, easier to operate, and beneficial for supporting their learning activities (Al-Gahtani, 2016; Ching-Ter,

Hajiyeve and Su, 2017). Empirical findings also indicate that PE significantly affects PEOU and PU in various e-learning contexts (Al-Gahtani, 2016; Hanif, Jamal and Imran, 2018; Salloum et al., 2019; Gurban and Almogren, 2022). In vocational education settings, enjoyable learning experiences further enhance students' engagement with technology-based learning activities, particularly when LMS platforms support interactive and practice-oriented instructional processes. Therefore, the following hypotheses are proposed:

- H5: PE positively influences PU in LMS use.
- H6: PE positively influences PEOU in LMS use.

Perceived Accessibility (PA)

PA refers to the extent to which users perceive that information and technological resources can be accessed easily, efficiently, and without significant barriers to support their learning activities (Al-Debei, 2014). In the context of e-learning, accessibility reflects students' ability to access LMS platforms anytime and anywhere using various digital devices and internet connections. High levels of accessibility allow students to obtain learning materials, submit assignments, and participate in online learning activities more conveniently, thereby enhancing their overall learning experience.

Previous studies have demonstrated that accessibility is an important factor influencing students' acceptance of e-learning systems. When LMS platforms can be accessed quickly and reliably, students are more likely to perceive the systems as more useful and easier to use (Almaiah, Jalil and Man, 2016; Revyathi and Tselios, 2019; Salloum et al., 2019). Accessibility also supports learning flexibility, which is particularly important in digital learning environments where students frequently rely on mobile devices and remote access

to participate in educational activities (Revythi and Tselios, 2019). Empirical evidence further indicates that PA positively affects both PU and PEOU in LMS adoption contexts (Almaiah, Jalil and Man, 2016; Revythi and Tselios, 2019; Salloum et al., 2019). In vocational education settings, where students require continuous access to instructional materials and practice-oriented resources, accessibility becomes a critical factor shaping LMS acceptance. Therefore, the following hypotheses are proposed:

- H7: PA positively influences PU in LMS use.
- H8: PA positively influences PEOU in LMS use.

Perceived Playfulness (PP)

PP refers to the extent to which users perceive technology use as enjoyable, engaging, and intrinsically stimulating during the interaction process (Hung, Tsai and Chou, 2016). In e-learning environments, playfulness reflects students' positive experiential responses when using an LMS, particularly when the platform enables interactive, creative, and flexible learning activities. Technologies perceived as playful are more likely to encourage students' active participation and sustained engagement in digital learning processes.

Previous studies have reported that PP significantly influences students' perceptions of e-learning usability and effectiveness. Students who experience enjoyable and interactive learning activities through LMS platforms tend to perceive the systems as easier to use and more beneficial for supporting their learning performance (Salloum et al., 2019; Khoirunisa, Warsono and Mirza, BR, 2023). Empirical evidence also demonstrates that PP positively affects PEOU in LMS adoption contexts (Salloum et al., 2019; Rokhim et al., 2024). In addition, prior research indicates that students who perceive LMS platforms as engaging and enjoyable are more likely to recognize their usefulness for learning activities and skill development (Gurban and Almogren, 2022). In vocational education settings, where instructional activities often emphasize creativity, practice-oriented learning, and interactive tasks, PP may strengthen students' acceptance of LMS use. Therefore, the following hypotheses are proposed:

- H9: PP positively influences PU in LMS use.
- H10: PP positively influences PEOU in LMS use.

Perceived Ease of Use (PEOU)

PEOU refers to the extent to which individuals believe that using a technology system requires minimal effort and is easy to understand and operate (Davis, 1989). In the context of LMS, PEOU reflects students' perceptions regarding the simplicity, flexibility, and convenience of interacting with e-learning platforms during learning activities. Technologies perceived as easy to use are more likely to encourage positive user experiences and continued technology use.

Within the TAM, PEOU is considered a fundamental determinant influencing both perceived usefulness and users' behavioral responses toward technology adoption. Previous studies have consistently shown that students who perceive LMS platforms as easy to use tend to recognize greater usefulness in supporting their learning activities (Al-Gahtani, 2016; Almaiah, Jalil and Man,

2016; Salloum et al., 2019; Rokhim et al., 2024). PEOU reduces the cognitive effort required to operate the system, enabling students to focus more effectively on learning tasks and academic performance.

In addition, prior research indicates that PEOU is positively associated with students' attitudes toward LMS use and their BI to continue using e-learning technologies (Al-Gahtani, 2016; Almaiah, Jalil and Man, 2016; Ching-Ter, Hajiyeve and Su, 2017; Salloum et al., 2019; Alshammari, 2020; Gurban and Almogren, 2022; Rokhim et al., 2024). Students who perceive LMS platforms as simple and user-friendly are more likely to develop favorable attitudes and stronger intentions to continue using LMS platforms (Hanif, Jamal and Imran, 2018; Revythi and Tselios, 2019; Salloum et al., 2019; Gurban and Almogren, 2022). In vocational education contexts, where students often engage in practice-oriented and task-based learning activities, system simplicity and operational convenience become particularly important in shaping LMS acceptance. Therefore, the following hypotheses are proposed:

- H11: PEOU positively influences PU in LMS use.
- H12: PEOU positively influences ATU in LMS use.
- H13: PEOU positively influences BI in LMS use.

Perceived Usefulness (PU)

PU refers to the extent to which individuals believe that using a particular technology can enhance their performance and improve task accomplishment (Davis, 1989). In the context of LMS, PU reflects students' beliefs that LMS platforms effectively support learning activities, improve academic productivity, and facilitate access to educational resources. Technologies perceived as useful are more likely to encourage students' positive attitudes and continued intention to use the system.

Within the TAM, PU is considered one of the strongest predictors of technology acceptance behavior. Previous studies have consistently demonstrated that students who perceive LMS platforms as beneficial for learning tend to develop more favorable attitudes toward e-learning (Gurban and Almogren, 2022; Or, 2024). PU also positively influences students' BI to continue using LMS platforms in educational settings (Al-Gahtani, 2016; Hanif, Jamal and Imran, 2018; Octavia and Nugraha, 2024). Empirical evidence from various e-learning studies further confirms that students are more likely to engage in sustained LMS use when they perceive the technology as effective in supporting their academic activities and learning outcomes (Al-Gahtani, 2016; Ching-Ter, Hajiyeve and Su, 2017; Revythi and Tselios, 2019; Salloum et al., 2019; Gurban and Almogren, 2022; Rokhim et al., 2024).

In vocational education contexts, PU becomes particularly important because students frequently rely on LMS platforms to access instructional materials, complete assignments, and support practice-oriented learning activities. Therefore, students who perceive LMS platforms as beneficial for improving their learning performance are expected to develop more positive attitudes and stronger intentions toward LMS use. Accordingly, the following hypotheses are proposed:

- H14: PU positively influences ATU in LMS use.
- H15: PU positively influences BI in LMS use.

Attitude Towards Using (ATU)

ATU refers to an individual's overall positive or negative evaluation of using a particular technology system (Davis, 1993). In the context of LMS, ATU reflects students' affective responses and evaluations based on their experiences interacting with e-learning platforms during learning activities. More positive attitudes toward LMS use are generally associated with favorable perceptions of technology usefulness, convenience, and learning effectiveness.

Within the TAM, ATU is considered an important determinant of users' BI to adopt and continue using technology. Previous studies have consistently demonstrated that students who develop positive attitudes toward LMS platforms are more likely to report stronger intentions to continue using e-learning systems in the future (Al-Aulamie, 2013; Hanif, Jamal and Imran, 2018; Gurban and Almogren, 2022). Similar findings have also been reported in studies examining e-learning acceptance in higher education and digital learning environments, where positive attitudes significantly influence students' BI in using LMS platforms (Hanif, Jamal and Imran, 2018; Revythi and Tselios, 2019; Salloum et al., 2019; Gurban and Almogren, 2022).

In vocational education settings, students who perceive LMS platforms as supportive, engaging, and beneficial for learning activities develop more favorable attitudes toward technology use, which subsequently strengthens their intention to continue utilizing LMS platforms. Therefore, the following hypothesis is proposed:

- H16: ATU positively influences BI in LMS use.

Behavioral Intention to Use (BI)

BI refers to an individual's motivational tendency or willingness to perform a particular behavior, specifically the intention to adopt and continue using a technology system (Davis, 1989). In the context of LMS, behavioral intention reflects students' readiness and commitment to engage in sustained LMS use for learning activities. BI is considered a key predictor of actual technology usage behavior within the TAM.

Previous studies have consistently demonstrated that students with stronger BI toward e-learning systems are more likely to actively and continuously use LMS platforms in their academic activities (Alshammari, 2020; Gurban and Almogren, 2022). Students who intend to use LMS platforms regularly tend to integrate the technology into their learning routines, including accessing instructional materials, completing assignments, and participating in online discussions. Empirical evidence further confirms that BI significantly influences AU in various e-learning environments (Davis, 1989; Gurban and Almogren, 2022).

In vocational education contexts, students' intention to use LMS platforms becomes particularly important because digital learning systems increasingly support practice-oriented learning activities and flexible access to instructional resources. Therefore, students with stronger BI toward LMS use are expected to demonstrate higher levels of AU. Accordingly, the following hypothesis is proposed:

- H17: BI positively influences AU in LMS use.

MATERIALS AND METHODS

Participants

This study employed a quantitative cross-sectional survey design to examine students' acceptance of LMS use in vocational secondary education. A cross-sectional approach was considered appropriate because the study aimed to investigate the relationships among TAM constructs and external variables at a specific point in time within an authentic educational setting (Creswell and Creswell, 2018). The research was conducted at a vocational secondary school in Surabaya, East Java, Indonesia, within the Economics and Business expertise area and included the E-commerce and Marketing, Office Management and Business Services, and Accounting programs.

The selected school was considered suitable for this study because it had systematically implemented LMS-based learning across multiple subjects and vocational study programs. Students regularly engaged with digital learning technologies through both synchronous and asynchronous instructional activities, including accessing learning materials, participating in classroom discussions, and submitting assignments online. This educational context provided a relevant setting for examining LMS acceptance among vocational secondary school students. This vocational education setting was selected because vocational students typically participate in practice-oriented and technology-supported learning activities, which may shape their perceptions and acceptance of educational technologies differently from those of students in general secondary or higher education contexts (Sánchez-Prieto et al., 2019; Almaiah, Al-khasawneh and Althunibat, 2020).

The study focused specifically on students who actively used the LMS as part of their regular learning activities. To ensure that respondents had sufficient experience with LMS use, only students who reported using the LMS for at least 10 hours per week were included. Weekly LMS-use time was self-reported in the questionnaire. The 10-hour threshold was used operationally in this study to identify relatively experienced users and should not be interpreted as a validated cut-off. Because this inclusion criterion conditions the sample on relatively high LMS exposure, the findings, including the BI→AU relationship, apply primarily to active/high-exposure users and should not be generalized to students with limited LMS exposure.

Data were collected using an online questionnaire distributed through Google Forms. Participation in the study was voluntary, and respondents were informed that all responses would remain anonymous and confidential. The data were used solely for academic research purposes. Ethical considerations throughout the research process followed principles of voluntary participation, informed consent, and participant confidentiality in educational research (Creswell and Creswell, 2018).

Table 1 presents the demographic characteristics and LMS usage profile of the 252 vocational secondary school students who participated in this study. Female students constituted the majority of the sample (59.5%), while male students accounted for 40.5%. In terms of grade level, most respondents were enrolled in Grade XI (48.4%), followed by Grade X (30.6%) and Grade XII (21.0%).

Variables	Category	Frequency
Gender	Female	150
	Male	102
Grade Level	Grade X	77
	Grade XI	122
	Grade XII	53
Programs	E-commerce and Marketing	48
	Office Management and Business Services	134
	Accounting	70
The Use of LMS Platform	LMS Portal Developed by Schools	135
	Moodle	58
	Google Classroom	48
	Multiple LMS	11
Weekly LMS Usage Time	10–15 hours	103
	16–20 hours	106
	More than 20 hours	43

Table 1: Respondent Demographic and LMS Profile

Regarding academic specialization, the respondents predominantly came from the Office Management and Business Services program (53.2%), followed by Accounting (27.8%) and E-commerce and Marketing (19.0%). Concerning LMS usage, the majority reported using LMS platforms developed internally by the school (53.6%), while others reported Moodle (23.0%), Google Classroom (19.0%), or multiple LMS platforms (4.4%). Regarding LMS engagement intensity, most students reported using an LMS for 16–20 hours per week (42.1%), followed by 10–15 hours (40.9%), while 17.1% reported more than 20 hours weekly. The study therefore involved students with substantial self-reported exposure to digital learning. Because the AU construct was measured through the questionnaire, it represents self-reported LMS use rather than objective usage-log data. Analyses pooled students across platform types and did not estimate platform-specific effects; this heterogeneity should be considered when interpreting the findings.

Instruments

The research instrument consisted of a structured questionnaire adapted from previous studies on e-learning acceptance and the TAM, particularly the instrument developed by Salloum et al. (2019). The questionnaire measured ten constructs, namely PEOU, PU, ATU, BI, AU, CSE, SN, PE, PA, and PP; the items are presented in Appendix 1.

A total of 44 measurement items were used to assess the proposed constructs. All items were evaluated using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Prior to data collection, the questionnaire items were reviewed and adapted to ensure their suitability for the vocational secondary education context and LMS-based learning environments.

Data Analysis

Data were analyzed using Structural Equation Modeling–Generalized Structured Component Analysis (SEM-GSCA) with GSCA Pro version 1.2.1. SEM-GSCA was selected because it is appropriate for analyzing complex predictive models involving multiple latent constructs and moderate

sample sizes (Hwang and Takane, 2004). Compared with covariance-based SEM, GSCA provides greater flexibility regarding multivariate normality assumptions and enables component-based estimation of latent variables while using a component-based estimation framework (Hwang, Ho and Lee, 2010). In addition, GSCA is considered suitable for exploratory and prediction-oriented research models, particularly in studies examining behavioral and educational technology acceptance (Solimun and Fernandes, 2017).

The analysis procedure consisted of three stages. First, overall model fit was evaluated to assess the compatibility between the proposed research model and the observed data. Model fit evaluation in GSCA commonly involves several goodness-of-fit indices, including FIT, AFIT, GFI, and SRMR, which collectively indicate the adequacy of the overall model (Hwang and Takane, 2004). Second, the measurement model was assessed through convergent validity, discriminant validity, indicator loading significance, and internal consistency reliability to ensure the adequacy and consistency of the measurement indicators (Hair and Alamer, 2022). Convergent validity was evaluated using factor loadings and Average Variance Extracted (AVE), while reliability was assessed using Cronbach’s alpha and composite reliability values.

Third, the structural model was evaluated using path coefficients, Variance Inflation Factor (VIF), coefficient of determination (R^2), and effect size (f^2) to examine the relationships among constructs and the explanatory power of the proposed model. Path coefficient analysis was conducted to test the proposed hypotheses, while VIF values were used to identify potential multicollinearity issues among predictor constructs (Hair and Alamer, 2022). Furthermore, R^2 and f^2 values were interpreted to assess the explanatory performance and substantive contribution of the exogenous variables on endogenous constructs within the extended Technology Acceptance Model (TAM).

RESULTS

Evaluation of Model Fit

The overall model fit was evaluated prior to examining the measurement and structural models. In SEM-GSCA, model

fit assessment is commonly conducted using the Goodness-of-Fit Index (GFI), Standardized Root Mean Square Residual (SRMR), and FIT indices to determine the extent to which the proposed model adequately represents the observed data (Hwang and Takane, 2004).

As presented in Table 2, the model demonstrated an acceptable level of fit. The GFI value (0.964) exceeded the recommended threshold of 0.95, while the SRMR value (0.074) remained below the maximum acceptable threshold of 0.08, indicating satisfactory model fit. Furthermore, the overall FIT value

of 0.815 suggests that the proposed model explained 81.5% of the total variance across indicators and latent constructs. The measurement model showed a high explanatory contribution (FITm = 0.927), whereas the structural model demonstrated a moderate explanatory level (FITs = 0.296). Overall, these findings indicate that the proposed SEM-GSCA model provided an acceptable basis for subsequent measurement- and structural-model evaluation. According to Hwang, Cho and Choo (2024), acceptable SEM-GSCA model fit is generally indicated by high GFI values and low SRMR values.

Fit Index	Obtained Value	Interpretation
FIT	0.815	Good
FITm	0.927	Excellent
FITs	0.296	Moderate
GFI	0.964	Good fit
SRMR	0.074	Acceptable fit

Table 2: SEM-GSCA Model Fit Indices

Evaluation of Measurement Model

The measurement model was assessed to evaluate the validity and reliability of the constructs used in this study. The assessment included indicator reliability, convergent validity, discriminant validity, and internal consistency reliability. Specifically, the analysis examined indicator loadings, Average Variance Extracted (AVE), Fornell–Larcker criterion, Heterotrait–Monotrait ratio (HTMT), composite reliability (CR), and Cronbach’s alpha values.

All indicators demonstrated acceptable standardized loadings ranging from 0.539 to 0.860. In addition, all bootstrap confidence intervals excluded zero, indicating that the indicators significantly represented their respective latent constructs. These findings support satisfactory indicator reliability and construct relevance.

Table 3 presents the indicator loading estimates, standard errors, and 95% confidence intervals for all indicators. The strongest loading was observed for PE4 (loading = 0.860), indicating that PE4 was the most representative indicator of the PE construct. Meanwhile, PEOU3 and SN4 showed relatively lower but still acceptable loading values, suggesting that these indicators contributed moderately to their respective constructs.

Convergent validity was subsequently evaluated using AVE values. As presented in Table 4, several constructs, namely SN, PA, PP, ATU, and AU, reported AVE values below the recommended threshold of 0.50. Nevertheless, this condition remains acceptable because all constructs demonstrated composite reliability values above 0.70, consistent with the recommendation of Fornell and Larcker (1981) that convergent validity is still considered adequate when composite reliability exceeds the acceptable level (see Table 4). Overall, composite reliability values ranged from 0.731 to 0.856, while Cronbach’s alpha values ranged from 0.700 to 0.846, indicating satisfactory internal consistency reliability across all constructs. Discriminant validity was first examined using the Fornell–Larcker criterion, as presented in Table 5. After recalculating the diagonal values as the square roots of the AVEs reported in Table 4, the criterion was not satisfied for several construct pairs, indicating notable overlap among some theoretically related variables. Because the Fornell–Larcker criterion has limited

sensitivity in variance- and component-based structural equation modelling, HTMT was also examined as a complementary criterion (Henseler, Ringle and Sarstedt, 2015). Accordingly, discriminant validity is interpreted cautiously rather than as unequivocally established for all construct pairs.

As presented in Table 6, most HTMT values were below the commonly recommended threshold of 0.90. However, the HTMT value between PU and ATU was 0.919 and therefore did not meet the stricter 0.90 criterion. This result indicates substantial overlap between perceived usefulness and attitude toward use, which is theoretically plausible within TAM but should nevertheless be considered when interpreting paths involving these constructs. The measurement model therefore showed broadly distinguishable constructs overall, with the PU–ATU distinction requiring particular caution.

Evaluation of Structural Model

The structural model was evaluated by examining multicollinearity, explanatory power, effect size, and the significance of the hypothesized relationships among constructs. The assessment included Variance Inflation Factor (VIF), coefficient of determination (R^2), effect size (f^2), and path coefficient estimates obtained through bootstrap resampling.

First, multicollinearity was assessed using the Variance Inflation Factor (VIF). As presented in Table 7, all VIF values ranged from 1.815 to 4.900, which remained below the recommended threshold of 5.00 (Hair, Ringle and Sarstedt, 2011; Hair et al., 2017). SEM is equivalent to carrying out covariance-based SEM (CB-SEM). These findings indicate that none of the reported VIF values exceeded the prespecified threshold of 5.00.

The explanatory power of the model was then evaluated using the coefficient of determination (R^2), as reported in Table 4. The results showed that the model explained 75.2% of the variance in ATU, 72.3% of the variance in PU, 63.9% of the variance in BI, 44.9% of the variance in PEOU, and 30.2% of the variance in AU. These findings suggest that the proposed model possesses moderate to substantial explanatory power, particularly in explaining students’ attitudes and perceived usefulness of LMS usage in vocational education contexts.

Construct	Item	Loadings		
		Estimate	SE	95% CI
Computer self-efficacy (CSE)	CSE1	0.715	0.039	[0.640, 0.793]
	CSE2	0.693	0.047	[0.584, 0.755]
	CSE3	0.714	0.051	[0.631, 0.816]
	CSE4	0.726	0.035	[0.654, 0.793]
	CSE5	0.697	0.039	[0.615, 0.788]
Subjective norm (SN)	SN1	0.625	0.060	[0.525, 0.743]
	SN2	0.583	0.076	[0.477, 0.787]
	SN3	0.783	0.073	[0.662, 0.917]
	SN4	0.543	0.058	[0.433, 0.635]
Perceived enjoyment (PE)	PE1	0.741	0.039	[0.666, 0.820]
	PE2	0.627	0.050	[0.542, 0.721]
	PE3	0.810	0.032	[0.747, 0.870]
	PE4	0.860	0.026	[0.812, 0.907]
	PE5	0.629	0.049	[0.528, 0.724]
Perceived accessibility (PA)	PA1	0.754	0.035	[0.688, 0.823]
	PA2	0.701	0.058	[0.582, 0.795]
	PA3	0.610	0.053	[0.505, 0.723]
	PA4	0.635	0.058	[0.519, 0.741]
	PA5	0.774	0.044	[0.693, 0.869]
Perceived playfulness (PP)	PP1	0.613	0.057	[0.465, 0.728]
	PP2	0.657	0.055	[0.560, 0.768]
	PP3	0.765	0.049	[0.668, 0.862]
	PP4	0.698	0.076	[0.550, 0.841]
	PP5	0.595	0.051	[0.501, 0.699]
Perceived ease of use (PEOU)	PEOU1	0.712	0.040	[0.644, 0.809]
	PEOU2	0.785	0.054	[0.688, 0.909]
	PEOU3	0.539	0.064	[0.413, 0.662]
	PEOU4	0.767	0.054	[0.661, 0.862]
Perceived usefulness (PU)	PU1	0.699	0.037	[0.632, 0.799]
	PU2	0.799	0.050	[0.714, 0.891]
	PU3	0.738	0.042	[0.668, 0.836]
	PU4	0.746	0.039	[0.678, 0.840]
Attitude towards using (ATU)	ATU1	0.679	0.041	[0.619, 0.774]
	ATU2	0.757	0.060	[0.662, 0.901]
	ATU3	0.675	0.045	[0.580, 0.766]
	ATU4	0.694	0.052	[0.603, 0.798]
Behavioral intention to use (BI)	BI1	0.709	0.045	[0.629, 0.788]
	BI2	0.716	0.042	[0.636, 0.797]
	BI3	0.722	0.036	[0.656, 0.791]
	BI4	0.734	0.052	[0.650, 0.841]
Actual system use (AU)	AU1	0.700	0.049	[0.597, 0.779]
	AU2	0.591	0.069	[0.417, 0.689]
	AU3	0.751	0.047	[0.677, 0.852]
	AU4	0.724	0.050	[0.624, 0.827]

Table 3: Measurement Item Factor Loadings

Construct	AVE	CR	Cronbach's Alpha	R ²
Computer self-efficacy (CSE)	0.503	0.835	0.828	
Subjective norm (SN)	0.410	0.731	0.700	
Perceived enjoyment (PE)	0.547	0.856	0.846	
Perceived accessibility (PA)	0.487	0.825	0.812	
Perceived playfulness (PP)	0.447	0.800	0.785	
Perceived ease of use (PEOU)	0.500	0.797	0.780	0.449
Perceived usefulness (PU)	0.557	0.834	0.822	0.723
Attitude towards using (ATU)	0.493	0.795	0.781	0.752
Behavioral intention to use (BI)	0.519	0.812	0.802	0.639
Actual system use (AU)	0.482	0.787	0.767	0.302

Note: AVE = Average Variance Extracted; CR = Composite Reliability. AVE values below 0.50 were considered acceptable because CR values exceeded 0.70.

Table 4: Construct Reliability, Validity, and R-square

	CSE	SN	PE	PA	PP	PU	PEOU	ATU	BI	AU
CSE	0.709									
SN	0.730	0.640								
PE	0.663	0.775	0.740							
PA	0.657	0.60	0.680	0.698						
PP	0.627	0.760	0.842	0.742	0.669					
PU	0.641	0.713	0.681	0.653	0.697	0.746				
PEOU	0.543	0.551	0.539	0.614	0.497	0.792	0.707			
ATU	0.703	0.717	0.749	0.682	0.731	0.881	0.762	0.702		
BI	0.589	0.585	0.672	0.691	0.691	0.768	0.660	0.786	0.720	
AU	0.538	0.543	0.619	0.437	0.539	0.534	0.410	0.595	0.561	0.694

Note: Diagonal values are the recalculated square roots of AVE; off-diagonal values are inter-construct correlations. Several construct pairs do not satisfy the Fornell–Larcker criterion.

Table 5: Fornell-Larcker Discriminant Validity Matrix

Construct	CSE	SN	PE	PA	PP	PU	PEOU	ATU	BI	AU
CSE	-	0.789	0.691	0.685	0.664	0.664	0.567	0.734	0.610	0.581
SN	0.789	-	0.845	0.656	0.832	0.760	0.591	0.779	0.634	0.609
PE	0.691	0.845	-	0.702	0.886	0.704	0.557	0.778	0.696	0.667
PA	0.685	0.656	0.702	-	0.795	0.683	0.649	0.716	0.716	0.476
PP	0.664	0.832	0.886	0.795	-	0.734	0.527	0.779	0.733	0.583
PU	0.664	0.760	0.704	0.683	0.734	-	0.821	0.919	0.796	0.574
PEOU	0.567	0.591	0.557	0.649	0.527	0.821	-	0.796	0.695	0.438
ATU	0.734	0.779	0.778	0.716	0.779	0.919	0.796	-	0.820	0.648
BI	0.610	0.634	0.696	0.716	0.733	0.796	0.695	0.820	-	0.601
AU	0.581	0.609	0.667	0.476	0.583	0.574	0.438	0.648	0.601	-

Note: Heterotrait–Monotrait ratio. Values below 0.90 are preferred; PU–ATU = 0.919 exceeds this stricter criterion and should be interpreted cautiously.

Table 6: HTMT Discriminant Validity Matrix

Relationship	VIF	Effect Size (f^2)
CSE → PEOU	2.196	0.039
CSE → PU	2.264	0.006
SN → PEOU	3.180	0.070
SN → PU	3.300	0.012
PE → PEOU	3.889	0.002
PE → PU	3.892	0.011
PA → PEOU	2.635	0.186
PA → PU	2.919	0.001
PP → PEOU	4.871	0.017
PP → PU	4.900	0.041
PEOU → PU	1.815	0.327
PEOU → ATU	2.464	0.044
PEOU → BI	2.635	0.006
PU → ATU	2.464	0.950
PU → BI	4.425	0.023
ATU → BI	4.025	0.123
BI → AU	N/A	0.459

Note: VIF was not applicable for BI → AU because AU was predicted by a single construct; therefore, multicollinearity assessment was not required for this relationship.

Table 7: Collinearity and Effect Size

Table 8 presents the structural path estimates, bootstrap standard errors, confidence intervals, and hypothesis testing results. The findings indicate that PA showed a significant positive association with PEOU ($\beta = 0.396$, 95% CI [0.150,

0.694]), whereas the paths from CSE, SN, PE, and PP to PEOU and PU were not statistically significant. Thus, PA was the only external factor in the model that showed a statistically significant direct association with PEOU.

Hypothesis	Relationship	β	SE	95% CI	Results
H1	CSE → PU	0.075	0.088	[-0.120, 0.242]	Not Supported
H2	CSE → PEOU	0.194	0.111	[-0.038, 0.385]	Not Supported
H3	SN → PU	0.108	0.117	[-0.105, 0.386]	Not Supported
H4	SN → PEOU	0.257	0.141	[-0.134, 0.455]	Not Supported
H5	PE → PU	0.102	0.106	[-0.104, 0.301]	Not Supported
H6	PE → PEOU	0.040	0.149	[-0.214, 0.334]	Not Supported
H7	PA → PU	0.020	0.082	[-0.143, 0.181]	Not Supported
H8	PA → PEOU	0.396	0.134	[0.150, 0.694]	Supported
H9	PP → PU	0.198	0.135	[-0.076, 0.450]	Not Supported
H10	PP → PEOU	-0.128	0.196	[-0.608, 0.173]	Not Supported
H11	PEOU → PU	0.496	0.088	[0.312, 0.681]	Supported
H12	PEOU → ATU	0.206	0.109	[0.034, 0.415]	Supported
H13	PEOU → BI	0.079	0.102	[-0.137, 0.284]	Not Supported
H14	PU → ATU	0.698	0.104	[0.418, 0.862]	Supported
H15	PU → BI	0.432	0.157	[0.111, 0.692]	Supported
H16	ATU → BI	0.331	0.147	[0.071, 0.624]	Supported
H17	BI → AU	0.550	0.056	[0.413, 0.645]	Supported

Table 8: Structural Path Hypothesis Testing

Furthermore, PEOU was significantly associated with PU ($\beta = 0.496$, 95% CI [0.312, 0.681]) and ATU ($\beta = 0.206$, 95% CI [0.034, 0.415]). However, the direct relationship between PEOU and BI was not statistically significant. This finding indicates that PEOU did not show a significant direct association with BI in the specified model; any indirect association through other TAM constructs was not tested in the reported analysis.

The results also showed that PU was significantly associated with both ATU ($\beta = 0.698$, 95% CI [0.418, 0.862]) and BI

($\beta = 0.432$, 95% CI [0.111, 0.692]). In addition, ATU was positively associated with BI ($\beta = 0.331$, 95% CI [0.071, 0.624]), and BI significantly predicted AU ($\beta = 0.550$, 95% CI [0.413, 0.645]). These findings are consistent with several central relationships proposed in the TAM, particularly the important role of PU and BI in predicting actual technology usage behavior.

The effect size analysis (f^2), presented in Table 7, further revealed that most external factors demonstrated relatively small effects on PU and PEOU. In contrast, several structural relationships

showed moderate to large effects, particularly the influence of PU on ATU ($f^2 = 0.950$), BI on AU ($f^2 = 0.459$), and PEOU on PU ($f^2 = 0.327$). These findings indicate that internal TAM constructs contributed more strongly to LMS acceptance behavior than the external variables included in the present study.

Interestingly, several external variables commonly found significant in prior TAM studies did not demonstrate significant effects in the present study. In this sample, practical utility and accessibility showed stronger direct relationships with the core TAM constructs than the social and affective external variables examined. Thus, the structural model results demonstrate that LMS acceptance among vocational secondary school students was more strongly associated with functional TAM constructs than with social influence or enjoyment-related factors.

DISCUSSION

Re-examining TAM Relationships in Vocational LMS Adoption

The findings support several core relationships in the TAM proposed by Davis (1989), particularly those among PEOU, PU, ATU, BI, and AU. PEOU was significantly associated with PU and ATU, indicating that students who perceive LMS platforms as easy to operate are more likely to recognize their academic value (Revythi and Tselios, 2019; Alshammari, 2020; Gurban and Almogren, 2022).

However, the direct relationship between PEOU and BI was not significant. In the present sample, usability was not a significant direct predictor of BI, which may reflect the participants' substantial prior exposure to digital technologies and online platforms (Revythi and Tselios, 2019; Hai et al., 2022; Kazu and Yalçin, 2022; Al-Adwan et al., 2023). This interpretation should be treated as context-specific rather than as evidence of a general post-pandemic change in the role of PEOU.

Conversely, PU emerged as the strongest predictor of ATU and BI, indicating that perceived usefulness was strongly associated with favorable attitudes and behavioral intention in this sample (Al-Gahtani, 2016; Almaiah, Jalil and Man, 2016; Salloum et al., 2019). The relationship between ATU and BI was also significant, indicating that more favorable attitudes were associated with stronger intentions to continue using LMS platforms (Hanif, Jamal and Imran, 2018; Muhaimin et al., 2019; Salloum et al., 2019). Furthermore, BI significantly predicted AU, supporting previous TAM studies that identify intention as the strongest predictor of actual LMS adoption behavior (Salloum et al., 2019; Alshammari, 2020; Gurban and Almogren, 2022).

Thus, the findings suggest that TAM remains highly relevant for explaining LMS adoption in vocational secondary education. Within this sample, PU showed a stronger direct association with BI than PEOU. This pattern suggests that perceived functional value may be particularly salient in this vocational school context.

Limited Direct Effects of Motivational External Factors

The results showed that the motivational external factors—CSE, SN, PE, and PP—did not have statistically significant direct

paths to PU or PEOU. Accordingly, these variables contributed limited direct explanatory value in the present model.

CSE was found to have no significant effect on either PU or PEOU. While earlier studies identified self-efficacy as an important predictor of e-learning acceptance (Al-Gahtani, 2016; Salloum et al., 2019), the present findings align with more recent studies showing weaker effects among technologically experienced students (Hanif, Jamal and Imran, 2018; Revythi and Tselios, 2019). One possible explanation is that students in this setting already had substantial familiarity with smartphones and digital platforms, reducing variation in basic technological confidence.

Similarly, SN did not significantly influence students' LMS perceptions. Unlike findings during the COVID-19 pandemic, where social influence strongly shaped online learning adoption (Salloum et al., 2019; Khan et al., 2020), One possible explanation is that LMS use was already embedded in formal learning routines, which may have reduced the additional contribution of perceived social influence.

PE and PP were also insignificant predictors, supporting studies suggesting that students may view LMS platforms primarily as functional academic tools rather than entertainment-oriented environments (Jiang et al., 2021; Rokhim et al., 2024). In vocational education, students may prioritize task completion, efficiency, and academic performance over emotional engagement. Moreover, the structured and assessment-oriented nature of institutional LMS platforms may limit pleasurable or playful learning experiences.

Taken together, these findings are consistent with an LMS environment that was relatively established among the participating students. The present cross-sectional design did not compare pandemic and post-pandemic cohorts and therefore cannot determine whether the role of motivational factors changed over time. Comparisons with earlier pandemic-period studies should consequently be interpreted as contextual contrasts rather than evidence of a temporal decline or shift in these factors.

The findings therefore suggest that the motivational TAM extensions examined here had limited direct explanatory contribution in this single-school, high-exposure sample. Functional considerations and accessibility were more strongly associated with LMS acceptance than the social and enjoyment-related external variables examined. These results support context-sensitive interpretation of technology-acceptance models in vocational secondary education rather than a general claim that motivational factors have become less important over time.

Accessibility as the Primary Driver of LMS Ease of Use

PA emerged as the only external factor that significantly influenced PEOU, although it did not significantly affect PU. This pattern suggests that accessibility was more closely related to perceived ease of use than to perceived educational value in the specified model. Students who reported greater accessibility also tended to report greater ease of use, whereas the direct PA–PU path was not significant. These findings are consistent with previous studies showing that PA positively affects PEOU

(Almaiah, Jalil and Man, 2016; Hanif, Jamal and Imran, 2018; Revyathi and Tselios, 2019; Salloom et al., 2019), while its relationship with PU remains weak or inconsistent (Park, 2009; Park, Nam and Cha, 2012; Solichin and Wijaya, 2021).

In vocational learning environments, students regularly use LMS platforms to access materials, submit assignments, and complete assessments within limited instructional time. Consequently, stable connectivity, device compatibility, and smooth system access may contribute to perceptions of PEOU. This finding aligns with TAM, which emphasizes that users' perceptions of system simplicity influence PEOU evaluations (Davis, 1989). The result can also be interpreted through Task–Technology Fit (TTF) theory, which argues that technology adoption increases when system features effectively support task completion (Goodhue et al., 1995). In this study, accessibility was associated with ease of use but did not show a significant direct path to perceived usefulness. This pattern is consistent with accessibility functioning as an enabling condition for routine academic activities. Nevertheless, accessibility alone does not necessarily enhance students' beliefs that the LMS improves learning quality or academic performance.

The findings imply that schools and LMS developers should prioritize digital infrastructure, responsive system design, and cross-device accessibility to reduce technological barriers. Nevertheless, technical accessibility alone is insufficient to strengthen PU. LMS platforms must also provide pedagogically meaningful learning experiences, collaborative activities, and authentic vocational tasks to ensure sustainable student acceptance.

LMS Use in an Institutionally Embedded Digital Learning Context

A notable finding of this study is the non-significant relationship between PEOU and BI, which contrasts with the original assumptions of TAM (Davis, 1989). In the present sample, PEOU was not a significant direct predictor of BI. This pattern is consistent with some studies conducted in digitally experienced learning contexts (Revyathi and Tselios, 2019; Hai et al., 2022; Al-Adwan et al., 2023), but the current cross-sectional data do not demonstrate that the role of ease of use has changed over time or specifically because of the post-pandemic context.

This finding can be explained by the characteristics of today's students, who are highly familiar with digital technologies and online platforms (Larionova et al., 2018; Kazu and Yalçin, 2022). Given their reported level of LMS exposure, participants may have perceived LMS operation as relatively routine. Where LMS use is institutionally required for accessing materials, submitting assignments, and completing assessments, continued use may occur even when the platform is not perceived as particularly easy or enjoyable (Hai et al., 2022; Al-Adwan et al., 2023).

The findings further indicate that BI was more strongly associated with PU and ATU than with PEOU. This suggests that perceived educational value and favorable attitudes may be more salient for BI than interface simplicity in this sample (Ni and Cheung, 2023; Barz et al., 2024). The findings are compatible with an institutionally embedded digital-learning

context in which pedagogical relevance may be important alongside technological convenience.

The findings may also be considered in relation to the broader normalization of digital learning in secondary education discussed in prior literature. Pandemic-era LMS use often occurred under emergency remote-learning conditions, whereas later studies have described more institutionally embedded forms of digital learning (Hermanto et al., 2021). The present results are compatible with this broader context, but they do not themselves establish a pandemic-to-post-pandemic shift because no temporal comparison was conducted.

From a practical perspective, these findings suggest that schools and LMS developers should move beyond focusing exclusively on usability improvements. Although user-friendly design remains important, educational institutions should prioritize the development of pedagogically meaningful digital learning experiences that enhance students' academic engagement and perceived learning benefits. Teachers should integrate interactive learning activities, collaborative projects, authentic vocational tasks, and reflective learning opportunities into LMS environments to strengthen students' intrinsic motivation and sustained behavioral intention. Simultaneously, LMS developers should design platforms that support adaptive learning, communication, feedback mechanisms, and vocationally relevant learning simulations to ensure that LMS usage is perceived as educationally valuable rather than merely administratively driven.

Practical and Pedagogical Implications

The findings highlight several important implications for LMS implementation in vocational secondary education. First, PU emerged as the strongest predictor of students' ATU and BI, indicating that students who perceived greater academic benefits also reported more favorable attitudes and stronger behavioral intentions. This finding reinforces prior studies emphasizing that sustainable LMS adoption depends largely on perceived instructional value and performance enhancement (Davis, 1989; Al-Adwan et al., 2023; Ni and Cheung, 2023).

Pedagogically, teachers play a central role in transforming LMS platforms from administrative tools into meaningful learning environments. When LMS use is limited primarily to material distribution and assessment activities, opportunities for deeper student engagement may be reduced. Therefore, teachers should integrate more interactive approaches, such as collaborative learning, project-based activities, reflective discussions, and authentic vocational simulations, to strengthen students' perceptions of usefulness and engagement (Teo et al., 2018; Al-Rahmi et al., 2020; Barz et al., 2024; Halomoan et al., 2026).

The findings regarding accessibility also generate practical implications for educational policymakers and school administrators. Because PA was significantly associated with PEOU, schools should ensure that students have stable internet access, compatible digital devices, and reliable technological infrastructure. Accessibility issues remain particularly important in vocational education contexts because students frequently rely on digital systems for both theoretical and practice-oriented learning activities. Schools should therefore invest in sustainable digital infrastructure development,

including high-speed internet connectivity, technical support services, and device availability to minimize technological barriers that may disrupt learning continuity. In addition, LMS developers should prioritize responsive interface design, mobile compatibility, simplified navigation systems, and lightweight platform architecture to accommodate diverse learning conditions and technological resources (Revythi and Tselios, 2019; Salloum et al., 2019).

Another important implication concerns the changing role of motivational external factors in post-pandemic learning environments. The non-significant paths from PE and PP suggest that entertainment-related perceptions had limited direct explanatory value in the present model, whereas functionality and accessibility were more prominent. Nevertheless, this does not imply that engagement-oriented features are unimportant. Rather, gamification and interactive elements should be integrated strategically to support meaningful learning experiences instead of functioning merely as entertainment components. LMS developers and educators should therefore consider incorporating adaptive feedback systems, achievement tracking, interactive multimedia content, and collaborative learning spaces that align with vocational learning objectives while maintaining students' engagement and attention (Hamari, Koivisto and Sarsa, 2014; Wang and Lin, 2021).

The findings suggest that institutional LMS implementation should be accompanied by attention to students' perceptions of educational value. In the model, BI was more strongly associated with PU and ATU than with PEOU. Therefore, schools should avoid treating LMS implementation merely as a compliance-oriented administrative policy. Instead, educational stakeholders should develop comprehensive digital learning ecosystems that cultivate students' autonomy, self-regulated learning skills, and active participation in online learning environments. Such an approach is particularly important in vocational education because students are expected to develop technological adaptability and lifelong digital competencies relevant to future workforce demands (OECD, 2021).

Finally, the findings have broader implications for efficient and responsible LMS implementation in vocational education. The present study directly measured technology-acceptance constructs rather than learning efficiency, learning continuity, competency attainment, academic performance, or vocational readiness. Accordingly, the empirical results support recommendations concerning accessibility, perceived usefulness, attitudes, intention, and self-reported use. Potential contributions of accessible and pedagogically meaningful LMS platforms to learning continuity or competency-oriented education should be understood as literature-informed implications rather than outcomes demonstrated by the current data (Snoussi, 2019; Al-Fraihat et al., 2020).

Limitations and Future Research

This study has several limitations that should be considered when interpreting the findings. First, the research was conducted in a single vocational secondary school in Surabaya, Indonesia, within the Business and Economics expertise area, specifically involving students from Office Management and Business

Services, Accounting, and E-commerce and Marketing programs. This specific context limits generalizability to other schools, regions, educational levels, vocational fields, and cultural settings. Second, the cross-sectional design captured students' LMS acceptance at a single point in time and therefore cannot establish temporal change or post-pandemic trends. Third, the study focused only on selected external TAM variables, namely CSE, SN, PE, PA, and PP, which may not fully represent the broader contextual factors influencing LMS acceptance. Fourth, eligibility required at least 10 hours of self-reported LMS use per week, so the sample represents relatively active/high-exposure users and may have restricted variability in the AU construct. Fifth, students reported using different LMS platforms, but platform-specific effects were not modelled, and AU was based on self-report rather than objective usage logs. Finally, several constructs showed AVE values below 0.50 and the discriminant-validity analyses indicated overlap among some constructs, particularly PU and ATU; paths involving these constructs should therefore be interpreted cautiously.

Future research should extend the present model by involving multiple vocational schools and broader geographical contexts to improve the external validity of the findings. Longitudinal studies are also recommended to examine changes in students' LMS acceptance and use patterns over time, particularly as digital learning becomes increasingly embedded in vocational education. In addition, future studies may employ mixed-methods approaches to capture students' experiences more deeply and incorporate additional variables, such as instructional quality, digital learning readiness, self-regulated learning, institutional support, LMS design quality, and differences among LMS platforms, to provide a more comprehensive understanding of sustainable LMS adoption in vocational secondary education.

CONCLUSION

This study found that, within the participating vocational school, LMS acceptance was more strongly associated with accessibility and perceived usefulness than with the motivational or hedonic external factors examined. By extending the TAM with selected external variables, the findings show that PU remains central in forming students' ATU and BI, while PA was the only external factor with a significant direct association with PEOU. These results highlight the salience of functional value and accessibility in this specific vocational school context.

The study contributes to the application of TAM by showing that CSE, SN, PE, and PP had limited direct explanatory contribution in this institutionally embedded vocational-school context. In this context, usefulness and accessibility showed more prominent associations with LMS acceptance than the motivational external factors examined. This finding highlights the need to interpret technology acceptance theories more contextually, particularly in vocational education where digital tools are expected to support practical, task-based, and performance-oriented learning.

Practically, the findings suggest that LMS implementation should address more than digital access alone. Schools,

teachers, and LMS developers should ensure that LMS platforms are reliable, pedagogically meaningful, and aligned with vocational learning needs. Responsible LMS integration should aim to reduce operational barriers and support useful learning activities while preserving students' engagement with academic tasks. The present study did not directly measure competency attainment or learning outcomes; therefore, any contribution of LMS implementation to competency-oriented education remains a broader pedagogical implication rather than an empirical outcome demonstrated here.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Universitas Negeri Surabaya for its academic support during the preparation of this manuscript. The authors

also acknowledge the support of the Indonesian Education Scholarship (BPI-PDDI), Center for Higher Education Funding and Assessment (PPAT), and the Indonesian Endowment Fund for Education (LPDP), Ministry of Higher Education, Science, and Technology of the Republic of Indonesia. The authors declare that this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors also declare that there are no conflicts of interest regarding the publication of this article.

In the preparation of this manuscript, ChatGPT was utilized solely to assist with language refinement, grammar checking, readability improvement, and academic writing support for publication purposes. All conceptualization, data analysis, interpretation of findings, and final conclusions remain entirely the responsibility of the authors.

REFERENCES

- Abdullah, F. and Ward, R. (2016) 'Developing a General Extended Technology Acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors', *Computers in Human Behavior*, Vol. 56, pp. 238–256. <https://doi.org/10.1016/j.chb.2015.11.036>
- Ahmad, N. A., Elias, N. F., Sahari, N. and Mohamed, H. (2023) 'Learning management system acceptance factors for technical and vocational education training (TVET) institutions', *TEM Journal*, Vol. 12, No. 2, pp. 1156–1165. <https://doi.org/10.18421/TEM122-61>
- Ajzen, I. (1991) 'The Theory of Planned Behavior', *Organizational Behavior and Human Decision Processes*, Vol. 50, No. 2, pp. 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Al-Adwan, A. S., Albelbisi, N. A., Hujran, O., Al-Rahmi, W. M. and Alkhalifah, A. (2021) 'Developing a Holistic Success Model for Sustainable E-Learning: A Structural Equation Modeling Approach Ahmad', *Sustainability (Switzerland)*, Vol. 13, No. 16, p. 9453. <https://doi.org/10.3390/su13169453>
- Al-Adwan, A. S., Li, N., Al-Adwan, A., Abbasi, G. A., Albelbisi, N. A. and Habibi, A. (2023) 'Extending the Technology Acceptance Model (TAM) to Predict University Students' Intentions to Use Metaverse-Based Learning Platforms', *Education and Information Technologies*, Vol. 28, No. 11, pp. 15381–15413. <https://doi.org/10.1007/s10639-023-11816-3>
- Al-Aulamie, A., Mansour, A., Daly, H. and Adjei, O. (2012) 'The effect of intrinsic motivation on learners' behavioural intention to use e-learning systems', *Proceedings of the 2012 International Conference on Information Technology Based Higher Education and Training (ITHET)*, pp. 1–4. <https://doi.org/10.1109/ITHET.2012.6246057>
- Al-Aulamie, A. (2013) *Enhanced technology acceptance model to explain and predict learners' behavioural intentions in learning management systems*, PhD thesis, Luton: University of Bedfordshire.
- Al-Debei, M. M. (2014) 'The Quality and Acceptance of Websites: an Empirical Investigation in the Context of Higher Education', *International Journal of Business Information Systems*, Vol. 15, No. 2, pp. 170–188. <https://doi.org/10.1504/IJBIS.2014.059252>
- Al-Fraihat, D., Joy, M., Masa'deh, R. and Sinclair, J. (2020) 'Evaluating E-learning Systems Success: An Empirical Study', *Computers in Human Behavior*, Vol. 102, pp. 67–86. <https://doi.org/10.1016/j.chb.2019.08.004>
- Al-Gahtani, S. S. (2016) 'Empirical Investigation of E-learning Acceptance and Assimilation: a Structural Equation Model', *Applied Computing and Informatics*, Vol. 12, No. 1, pp. 27–50. <https://doi.org/10.1016/j.aci.2014.09.001>
- Al-Rahmi, W. M., Alzahrani, A. I., Yahaya, N., Alalwan, N. and Kamin, Y. B. (2020) 'Digital Communication: Information and Communication Technology (ICT) Usage for Education Sustainability', *Sustainability (Switzerland)*, Vol. 12, No. 12, p. 5052. <https://doi.org/10.3390/su12125052>
- Almaiah, M. A., Al-khasawneh, A. and Althunibat, A. (2020) 'Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic', *Education and Information Technologies*, Vol. 25, No. 6, pp. 5261–5280. <https://doi.org/10.1007/s10639-020-10219-y>
- Almaiah, M. A., Jalil, M. A. and Man, M. (2016) 'Extending the TAM to Examine the Effects of Quality Features on Mobile Learning Acceptance', *Journal of Computers in Education*, Vol. 3, No. 4, pp. 453–485. <https://doi.org/10.1007/s40692-016-0074-1>
- Alshammari, S. H. (2020) 'The Influence of Technical Support, Perceived Self-efficacy, and Instructional Design on Students' Use of Learning Management Systems', *Turkish Online Journal of Distance Education*, Vol. 21, No. 3, pp. 112–141. <https://doi.org/10.17718/TOJDE.762034>
- Arief, U.M. et al. (2021) 'Computer Self-Efficacy and Organizational Culture In Affecting Technology Acceptance Model', *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, Vol. 12, No. 5, pp. 484–489. <https://doi.org/10.17762/turcomat.v12i5.997>
- Barz, N., Benick, M., Dörrenbächer-Ulrich, L. and Perels, F. (2024) 'Students' acceptance of e-learning: extending the technology acceptance model with self-regulated learning and affinity for technology', *Discover Education*, Vol. 3, No. 114, pp. 1–20. <https://doi.org/10.1007/s44217-024-00195-7>
- Benson, O. O. et al. (2020) 'Students' Perception of Teachers' Pedagogical Skills and its Influence on their Attitude towards Science, Implication for Science, Technology and Engineering Careers', *International Journal of Mechanical and Production Engineering Research and Development*, Vol. 10, No. 3, pp. 14701–14714. <https://doi.org/10.24247/ijmperdjun20201397>

- Cahyanti, A. D., Hermanto, F. Y., Rosy, B., Atmojo, P. J., Utami, A. T. and Maulida, M. S. (2025) 'Technology acceptance model on pedagogical skills: The moderating role of Android game addiction', *Journal of Theoretical and Applied Information Technology*, Vol. 103, No. 12, pp. 5203–5212.
- Ching-Ter, C., Hajiyev, J. and Su, C.R. (2017) 'Examining the Students' Behavioral Intention to Use E-learning in Azerbaijan? The General Extended Technology Acceptance Model for E-learning Approach', *Computers and Education*, Vol. 111, pp. 128–143. <https://doi.org/10.1016/j.compedu.2017.04.010>
- Creswell, J. W. and Creswell, J. D. (2018) *Research design: Qualitative, quantitative, and mixed methods approaches*, 5th edition, Los Angeles: SAGE Publications.
- Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly*, Vol. 13, No. 3, pp. 319–340. <https://doi.org/10.2307/249008>
- Davis, F. D. (1993) 'User Acceptance of Information Technology: System Characteristics, User Perceptions and Behavioral Impacts', *International Journal of Man-Machine Studies*, Vol. 38, No. 3, pp. 475–487. <https://doi.org/10.1006/imms.1993.1022>
- Downey, J. P. and Kher, H. V. (2015) 'A Longitudinal Examination of the Effects of Computer Self-efficacy Growth on Performance during Technology Training', *Journal of Information Technology Education: Research*, Vol. 14, pp. 91–111. <https://doi.org/10.28945/2114>
- Espinosa, E. O. C., Ruiz, J. A. C. and Mercado, M. T. C. (2021) 'Computational Self-Efficacy and Information Skills in Postgraduate Students in Administration at the Times of Covid-19', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 14, No. 2, pp. 118–129. <https://doi.org/10.7160/eriesj.2021.140205>
- Fornell, C. and Larcker, D. F. (1981) 'Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics', *Journal of Marketing Research*, Vol. 18, No. 3, pp. 382–388. <https://doi.org/10.2307/3150980>
- Goodhue, D. L. and Thompson, R. L. (1995) 'Task-Technology Fit and Individual Performance', *MIS Quarterly*, Vol. 19, No. 2, pp. 213–236. <https://doi.org/10.2307/249689>
- Gurban, M. A. and Almogren, A. S. (2022) 'Students' Actual Use of E-Learning in Higher Education During the COVID-19 Pandemic', *Sage Open*, Vol. 12, No. 2, pp. 1–16. <https://doi.org/10.1177/21582440221091250>
- Hai, L., Sang, G., Wang, H., Li, W. and Bao, X. (2022) 'An Empirical Investigation of University Students' Behavioural Intention to Adopt Online Learning: Evidence from China', *Behavioral Sciences*, Vol. 12, No. 10, p. 403. <https://doi.org/10.3390/bs12100403>
- Hair, J. and Alamer, A. (2022) 'Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example', *Research Methods in Applied Linguistics*, Vol. 1, No. 3, Article 100027. <https://doi.org/10.1016/j.rmal.2022.100027>
- Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M. and Sarstedt, M. (2017) *A primer on partial least squares structural equation modeling (PLS-SEM)*, 2nd edition, Los Angeles: SAGE Publications.
- Hair, J. F., Ringle, C. M. and Sarstedt, M. (2011) 'PLS-SEM: Indeed a silver bullet', *Journal of Marketing Theory and Practice*, Vol. 19, No. 2, pp. 139–152. <https://doi.org/10.2753/MTP1069-6679190202>
- Halomoan, H., Hakiki, M., Putra, B. A. W., Hamid, M. A., Utami, R., Saputro, I. N., Azizah, W. A., Sabir, A., Hidayah, Y. and Yassin, A. (2026) 'Deep Learning Methods Towards a Pedagogical Framework and Implementation Strategy: A Study of Information Technology Education Curriculum Development in Indonesia', *Journal of Teaching and Learning*, Vol. 20, No. 1, pp. 185–205. <https://doi.org/10.22329/jtl.v20i1.9970>
- Hamari, J., Koivisto, J. and Sarsa, H. (2014) 'Does gamification work? A literature review of empirical studies on gamification', *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS 2014)*, Waikoloa, HI, pp. 3025–3034. <https://doi.org/10.1109/HICSS.2014.377>
- Hanif, A., Jamal, F. Q. and Imran, M. (2018) 'Extending the Technology Acceptance Model for Use of e-Learning Systems by Digital Learners', *IEEE Access*, Vol. 6, pp. 73395–73404. <https://doi.org/10.1109/ACCESS.2018.2881384>
- Henseler, J., Ringle, C. M. and Sarstedt, M. (2015) 'A new criterion for assessing discriminant validity in variance-based structural equation modeling', *Journal of the Academy of Marketing Science*, Vol. 43, No. 1, pp. 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hermanto, F. Y., Sutirman, Sholikah, M. and Ranu, M. E. (2021) 'The effectiveness of distance practice learning for facing COVID-19 pandemic in Indonesia', *Journal of Theoretical and Applied Information Technology*, Vol. 99, No. 12, pp. 2925–2936.
- Hung, S. Y., Tsai, J. C. A. and Chou, S. T. (2016) 'Decomposing Perceived Playfulness: a Contextual Examination of Two Social Networking Sites', *Information and Management*, Vol. 53, No. 6, pp. 698–716. <https://doi.org/10.1016/j.im.2016.02.005>
- Hwang, H., Cho, G. and Choo, H. (2024) 'GSCA Pro—Free Stand-Alone Software for Structural Equation Modeling', *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 31, No. 4, pp. 696–711. <https://doi.org/10.1080/10705511.2023.2272294>
- Hwang, H., Ho, M.-H.R. and Lee, J. (2010) 'Generalized Structured Component Analysis with Latent Interactions', *Psychometrika*, Vol. 75, No. 2, pp. 228–242. <https://doi.org/10.1007/s11336-010-9157-5>
- Hwang, H. and Takane, Y. (2004) 'Generalized Structured Component Analysis', *Psychometrika*, Vol. 69, No. 1, pp. 81–99. <https://doi.org/10.1007/BF02295841>
- Im, T. (2021) 'Online and blended learning in vocational training institutions in South Korea', *Knowledge Management & E-Learning*, Vol. 13, No. 2, pp. 194–208. <https://doi.org/10.34105/j.kmel.2021.13.011>
- Jiang, M. Y., Jong, M. S., Lau, W. W., Meng, Y., Chai, C. and Chen, M. (2021) 'Validating the General Extended Technology Acceptance Model for E-Learning: Evidence From an Online English as a Foreign Language Course Amid COVID-19', *Frontiers in Psychology*, Vol. 12, p. 671615. <https://doi.org/10.3389/fpsyg.2021.671615>
- Castiblanco Jimenez, I. A., Cepeda García, L. C., Violante, M. G., Marcolin, F. and Vezzetti, E. (2021) 'Commonly Used External TAM Variables in E-learning, Agriculture and Virtual Reality Applications', *Future Internet*, Vol. 13, No. 7, pp. 1–21. <https://doi.org/10.3390/fi13010007>
- Kazu, İ. Y. and Yalçın, C. K. (2022) 'The Relationship Between Secondary School Teachers and Students' Readiness of Using Flipped Classroom', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 15, No. 1, pp. 1–9. <https://doi.org/10.7160/eriesj.2022.150101>

- Khan, S. A., Zainuddin, M., Mahi, M. and Arif, I. (2021) 'Behavioral intention to use online learning during COVID-19: An analysis of the technology acceptance model', in *1 International Conference on Innovative Methods of Teaching and Technological Advancements in Higher Education (IMTTA 2020): Book of Abstracts*, Tbilisi: Publishing House Universal, pp. 39–41. Available at: <https://eu.edu.ge/sites/default/files/conference/Book%20of%20abstracts%20IMTTA%202020.pdf>
- Khoirunisa, K. R., Warsono, S. and Mirza, BR, A. D. (2023) 'Expanding Technology Acceptance Model 3 Use Innovation Diffusion Theory on Accounting Learning During Pandemic: Insight from Indonesia', *The Indonesian Journal of Accounting Research*, Vol. 26, No. 01, pp. 25–48. <https://doi.org/10.33312/ijar.656>
- Larionova, V., Brown, K., Bystrova, T. and Sinitsyn, E. (2018) 'Russian Perspectives of Online Learning Technologies in Higher Education: An Empirical Study of a MOOC', *Research in Comparative and International Education*, Vol. 13, No. 1, pp. 70–91. <https://doi.org/10.1177/1745499918763420>
- Lee, B. C., Yoon, J. O. and Lee, I. (2009) 'Learners' Acceptance of E-learning in South Korea: Theories and Results', *Computers and Education*, Vol. 53, No. 4, pp. 1320–1329. <https://doi.org/10.1016/j.compedu.2009.06.014>
- Llamas, M.C.A.R., Vilela-Malabanan, C.M. and Dinawanao, D.D. (2025) 'Needs and Challenges of Academics and Students in Using a Learning Management System: a User Experience Approach', *Journal of Education and Learning*, Vol. 19, No. 3, pp. 1662–1669. <https://doi.org/10.11591/edulearn.v19i3.22350>
- Mohammadi, H. (2015) 'Computers in Human Behavior Investigating users' perspectives on e-learning: An integration of TAM and IS success model', *Computers and Education*, Vol. 45, pp. 359–374. <https://doi.org/10.1016/j.chb.2014.07.044>
- Muhaimin, M., Habibi, A., Mukminin, A., Pratama, R., Asrial, A. and Harja, H. (2019) 'Predicting Factors Affecting Intention to Use Web 2.0 in Learning: Evidence from Science Education', *Journal of Baltic Science Education*, Vol. 18, No. 4, pp. 595–606. <https://doi.org/10.33225/jbse/19.18.595>
- Ni, A. and Cheung, A. (2023) 'Understanding Secondary Students' Continuance Intention to adopt AI-powered Intelligent Tutoring System for English Learning', *Education and Information Technologies*, Vol. 28, No. 3, pp. 3191–3216. <https://doi.org/10.1007/s10639-022-11305-z>
- Octavia, A. C. B. and Nugraha, J. (2024) 'Influence of Enjoyment and Trust on the Use of Artificial Intelligence-based Voice Assistant in Vocational Students Using Technology Acceptance Model (TAM)', *Journal of Office Administration: Education and Practice*, Vol. 4, No. 1, pp. 10–23. <https://doi.org/10.26740/joaep.v4n1.p10-23>
- OECD (2021) *21st-century readers: Developing literacy skills in a digital world, PISA, Paris: OECD Publishing*. <https://doi.org/10.1787/a83d84cb-en>
- Or, C. (2024) 'Watch That Attitude! Examining the Role of Attitude in the Technology Acceptance Model through Meta-Analytic Structural Equation Modelling', *Journal of Technology in Education and Science (IJTES)*, Vol. 8, No. 4, pp. 558–582. <https://doi.org/10.46328/ijtes.575>
- Park, S. Y. (2009) 'An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use e-Learning', *Educational Technology and Society*, Vol. 12, No. 3, pp. 150–162.
- Park, S. Y., Nam, M. W. and Cha, S. B. (2012) 'University Students' Behavioral Intention to Use Mobile Learning: Evaluating the Technology Acceptance Model', *British Journal of Educational Technology*, Vol. 43, No. 4, pp. 592–605. <https://doi.org/10.1111/j.1467-8535.2011.01229.x>
- Revythi, A. and Tselios, N. (2019) 'Extension of Technology Acceptance Model by Using System Usability Scale to Assess Behavioral Intention to Use E-learning', *Education and Information Technologies*, Vol. 24, No. 4, pp. 2341–2355. <https://doi.org/10.1007/s10639-019-09869-4>
- Rogers, J.K.B., Mercado, T.C.R. and Decano, R.S. (2025) 'Moodle Interactions and Academic Performance: Educational Data Mining in a Philippine University', *Journal of Education and Learning*, Vol. 19, No. 1, pp. 542–550. <https://doi.org/10.11591/edulearn.v19i1.21549>
- Rohana, W.D., Nugraha, J. and Koblianska, I. (2025) 'Application of TAM to Explore Awareness and Enjoyment of the Use of Artificial Intelligence ChatGPT in Students', *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, Vol. 13, No. 2, pp. 498–512. <https://doi.org/10.26740/jpap.v13n2.p498-512>
- Rokhim, R., Mayasari, I., Wulandari, P. and Haryanto, H. C. (2024) 'Analysis of the Extrinsic and Intrinsic Aspects of the Technology Acceptance Model Associated with the Learning Management System during the COVID-19 Pandemic', *VINE Journal of Information and Knowledge Management Systems*, Vol. 54, No. 6, pp. 1289–1314. <https://doi.org/10.1108/VJKMS-04-2022-0113>
- Salloum, S. A., Qasim Mohammad Alhamad, A., Al-Emran, M., Abdel Monem, A. and Shaalan, K. (2019) 'Exploring Students' Acceptance of E-Learning Through the Development of a Comprehensive Technology Acceptance Model', *IEEE Access*, Vol. 7, pp. 128445–128462. <https://doi.org/10.1109/ACCESS.2019.2939467>
- Sánchez-Prieto, J. C., Hernández-García, Á., García-Peñalvo, F. J., Chaparro-Peláez, J. and Olmos-Migueláñez, S. (2019) 'Break the Walls! Second-Order Barriers and the Acceptance of mLearning by First-Year Pre-Service Teachers', *Computers in Human Behavior*, Vol. 95, pp. 158–167. <https://doi.org/10.1016/j.chb.2019.01.019>
- Snoussi, T. (2019) 'Learning Management System in Education: Opportunities and Challenges', *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, No. 12S, pp. 664–667. <https://doi.org/10.35940/ijtee.11161.10812s19>
- Solichin, A. and Wijaya, R. (2021) 'Acceptance of E-Learning System at Private University in Indonesia During the COVID-19 Pandemic: Students' Perspectives', *Jurnal Infotel*, Vol. 13, No. 3, pp. 104–113. <https://doi.org/10.20895/infotel.v13i3.662>
- Solimun, S. and Fernandes, A. A. R. (2017) 'Investigation the Mediating Variable: What Necessary? (Case Study in Management Research)', *International Journal of Law and Management*, Vol. 59, No. 6, pp. 1059–1067. <https://doi.org/10.1108/IJLMA-09-2016-0077>
- Söllner, M., Bitzer, P., Janson, A. and Leimeister, J. M. (2018) 'Process is king: Evaluating the performance of technology-mediated learning in vocational software training', *Journal of Information Technology*, Vol. 33, No. 3, pp. 233–253. <https://doi.org/10.1057/s41265-017-0046-6>
- Sorensen, K. N. (2019) Millennials' acceptance of voice activated shopping, Master's thesis, University of Nebraska–Lincoln. Available at: <https://digitalcommons.unl.edu/textilesdiss/13> [Accessed 16 September 2026].

- Surjono, H. D., Muhtadi, A. and Trilisiana, N. (2019) 'The effects of online activities on student learning outcomes in blended learning environment', *Proceedings of the 2019 3rd International Conference on Education and Multimedia Technology (ICEMT 2019)*, pp. 107–110. <https://doi.org/10.1145/3345120.3345167>
- Sutirman, Hermanto, F. Y., Adinata, R. E., Atmojo, P. J., Yuliansah and Suyanto, U. Y. (2026) 'Technology Acceptance and Self-regulated Learning among Vocational Students: The Mediating Role of Interest in Learning', *Cogent Education*, Vol. 13, No. 1, pp. 1–20. <https://doi.org/10.1080/2331186X.2026.2687173>
- Teo, T., Sang, G., Mei, B. and Hoi, C. K. W. (2018) 'Investigating Pre-service Teachers' Acceptance of Web 2.0 Technologies in Their Future Teaching: a Chinese Perspective', *Interactive Learning Environments*, Vol. 27, No. 4, pp. 530–546. <https://doi.org/10.1080/10494820.2018.1489290>
- Unlu, H. and Suel, E. (2014) 'Computer Self-Efficacy of Prospective Physical Education Teachers', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 7, No. 2, pp. 33–36. <https://doi.org/10.7160/eriesj.2014.070202>
- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003) 'User Acceptance of Information Technology: Toward a Unified View', *MIS Quarterly*, Vol. 27, No. 3, pp. 425–478. <https://doi.org/10.2307/30036540>
- Venkatesh, V. and Davis, F.D. (2000) 'A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies', *Management Science*, Vol. 46, No. 2, pp. 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Wang, Y. and Lin, K. (2021) 'Understanding Continuance Usage of Mobile Learning Applications: The Moderating Role of Habit', *Frontiers in Psychology*, Vol. 12, p. 736051. <https://doi.org/10.3389/fpsyg.2021.736051>

FROM KNOWLEDGE TO INSTRUCTION: EXAMINING THE ROLES OF USEFULNESS AND TEACHING SELF-EFFICACY ON POSTGRADUATE STUDENTS' WILLINGNESS TO TEACH ETHNOMATHEMATICS

Ebenezer Kwesi Lotey✉
Ebenezer Bonyah

¹Department of Mathematics Education, University of Skills Training and Entrepreneurial Development, Kumasi-Ghana, Ghana

✉ ebenezerlotey92@gmail.com

ABSTRACT

Teacher education programmes invest in ethnomathematics knowledge, but its value depends on whether it is translated into instructional practice. This study examined whether ethnomathematics usefulness (EU) and teaching self-efficacy (ETSE) mediated relationships of ethnomathematics content knowledge (ECK) and pedagogical knowledge (EPK) with willingness to teach ethnomathematics (WTE) among Ghanaian postgraduate students. Guided by a pragmatic paradigm, data were collected from 159 participants through an online questionnaire and telephone interviews with eight purposively selected participants. Quantitative data were analysed using covariance-based structural equation modelling in AMOS 31 and qualitative data through reflexive thematic analysis. ECK and EPK were positively associated with EU and ETSE. WTE was positively predicted by EPK, EU, and ETSE, while ECK showed a small negative direct association but a positive total effect. EU and ETSE showed significant indirect effects in both the ECK–WTE and EPK–WTE relationships. Interviewees expressed strong willingness to use cultural artefacts in teaching while identifying curricular and assessment constraints. The findings suggest that ethnomathematics knowledge is more likely to translate into teaching willingness when accompanied by perceived usefulness and confidence, while institutional conditions may influence the feasibility of culturally inclusive mathematics teaching.

KEYWORDS

Efficiency in education, ethnomathematics, Ghanaian postgraduate education, pedagogical content knowledge, teaching self-efficacy, willingness to teach

HOW TO CITE

Lotey E. K., Bonyah E. (2026) 'From Knowledge to Instruction: Examining the Roles of Usefulness and Teaching Self-Efficacy on Postgraduate Students' Willingness to Teach Ethnomathematics', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 254–268. <http://dx.doi.org/10.7160/eriesj.2026.190306>

Article history

Received

August 4, 2025

Received in revised form

July 30, 2026

Accepted

September 7, 2026

Available on-line

September 30, 2026

Highlights

- *Ethnomathematics content and pedagogical knowledge are positively associated with usefulness and self-efficacy beliefs.*
- *Perceived usefulness and self-efficacy positively predict willingness to teach ethnomathematics.*
- *Both beliefs show significant indirect effects linking knowledge and willingness.*
- *Effective preparation should build perceived value and confidence, not knowledge alone.*

INTRODUCTION

Globally, mathematics educators are expected not only to help students master abstract concepts but also to show how mathematics emerges from and functions within the social and cultural environments in which they live. This expectation has stimulated growing interest in culturally responsive approaches

to mathematics teaching that recognise learners' cultural identities, experiences, and indigenous knowledge as valuable resources for learning. Rather than presenting mathematics as a culturally neutral body of universal truths, researchers such as D'Ambrosio (1985), Rosa and Orey (2016), Owusu-Darko et al. (2023), and Kyeremeh et al. (2025) advocate

instructional practices that demonstrate how mathematical ideas are embedded in everyday activities, historical traditions, and community practices. Such approaches are intended to make mathematics more meaningful, promote equitable participation, and strengthen learners' appreciation of the discipline across diverse educational contexts. Among them, ethnomathematics—a field pioneered by D'Ambrosio (1985) that examines relationships between mathematical thinking and cultural practices—has received considerable research attention.

Ethnomathematics challenges the long-standing assumption that mathematics is exclusively a product of Western intellectual traditions by recognising that societies develop mathematical ideas in response to problems arising from their environments, occupations, technologies, and social organisation (Babbitt et al., 2015; Rosa et al., 2016; D'Ambrosio and Rosa, 2017; Owusu-Darko et al., 2023). Activities such as counting, architecture, textile design, navigation, trading, craftsmanship, agriculture, and artistic expression can embody sophisticated mathematical reasoning that has evolved over generations (Gerdes, 2005; Kyeremeh et al., 2023; Payadnya et al., 2024). Ethnomathematics therefore broadens conceptions of legitimate mathematical knowledge and encourages a more inclusive perspective on mathematics education.

The educational value of ethnomathematics extends beyond recognising cultural diversity. Research suggests that connecting mathematics instruction to learners' cultural experiences can support conceptual understanding, classroom participation, motivation, and positive attitudes towards mathematics (Babbitt et al., 2015; Sunzuma et al., 2021; Khalil, 2023). Students may engage more meaningfully with mathematical ideas when they encounter familiar cultural contexts in instruction, allowing them to connect informal knowledge acquired in their communities with formal mathematical concepts. Ethnomathematics has therefore increasingly been advocated as a component of culturally responsive and socially just mathematics education, particularly in contexts characterised by rich indigenous knowledge systems (Palhares, 2012; Rosa and Orey, 2016; D'Ambrosio and Knijnik, 2020; Wulandari et al., 2024).

Despite these recognised benefits, successful integration of ethnomathematics into formal education depends substantially on teachers' preparation. Unlike conventional mathematics instruction, ethnomathematics requires educators to identify mathematical ideas embedded in local cultural practices, evaluate their mathematical authenticity, and design learning experiences that connect indigenous knowledge meaningfully with formal school mathematics. Teachers therefore need sufficient disciplinary knowledge to recognise authentic mathematical ideas in cultural practices and the pedagogical expertise to transform those ideas into effective learning experiences. Without adequate preparation, attempts to integrate ethnomathematics may remain superficial cultural illustrations rather than mathematically meaningful learning opportunities.

The importance of teacher preparation is underscored by evidence that many educators lack confidence in implementing culturally responsive instruction. Mania and Alam (2021), for example, reported that more than 60% of mathematics teachers in their

Indonesian study perceived their training in ethnomathematics approaches as insufficient. Similarly, Kyeremeh et al. (2025) documented inadequate preparation as a barrier to integrating cultural knowledge into mathematics instruction in Ghana. These findings highlight the need to understand factors that may enable or constrain educators' willingness to adopt ethnomathematics in professional practice.

Shulman's (1986) theory of teacher knowledge provides a useful framework for explaining this requirement. According to the theory, effective teaching depends upon multiple forms of professional knowledge, particularly content knowledge and pedagogical knowledge. Content knowledge refers to teachers' understanding of the concepts, principles, and structures within a discipline, whereas pedagogical knowledge concerns the ability to organise instruction, select appropriate teaching strategies, and facilitate meaningful learning. Within the context of ethnomathematics, content knowledge enables educators to recognise mathematical ideas embedded in cultural practices, while pedagogical knowledge equips them to transform these ideas into effective classroom learning experiences. Both content and pedagogical knowledge in ethnomathematics provide the professional foundation necessary for implementing culturally responsive mathematics instruction.

However, possessing professional knowledge does not necessarily guarantee instructional implementation. Educators may understand an educational innovation yet remain reluctant to adopt it because they question either its educational value or their own capability to implement it successfully (Davis, 2003; Ponticell, 2003; Buabeng-Andoh, 2012). Understanding instructional intentions therefore also requires attention to psychological factors that may shape how professional knowledge is translated into practice. One such factor is ethnomathematics usefulness (EU), defined here as educators' beliefs about the extent to which integrating ethnomathematics can improve teaching, enhance students' mathematical understanding, strengthen cultural identity, and support meaningful learning. Educational innovations perceived as beneficial are generally more likely to be accepted and implemented than those considered ineffective or irrelevant (Papagiannis et al., 1982; Serdyukov, 2017). Educators who perceive ethnomathematics as educationally valuable may therefore express stronger willingness to incorporate culturally grounded mathematical activities into instruction. Moreover, those with stronger content and pedagogical knowledge may be better positioned to recognise these potential benefits because they understand both the mathematical richness of cultural practices and the pedagogical opportunities they provide.

In addition to perceiving ethnomathematics as useful, educators must believe that they are capable of implementing it effectively. Bandura's (1977) Social Cognitive Theory identifies self-efficacy as individuals' beliefs in their capability to organise and execute the actions required to achieve desired outcomes. In educational contexts, teacher self-efficacy has been associated with instructional innovation, persistence, instructional quality, and classroom effectiveness (Kwarteng et al., 2025; Lotey et al., 2026a). Educators with stronger self-efficacy are generally more willing to experiment with unfamiliar approaches, persist through instructional difficulties,

and sustain innovative practices. Ethnomathematics presents distinctive demands: educators must identify authentic mathematical ideas in cultural contexts, adapt curriculum materials, facilitate culturally sensitive discussions, and connect indigenous knowledge with formal mathematical reasoning (Mania and Alam, 2021; Kyeremeh et al., 2025). Ethnomathematics teaching self-efficacy (ETSE) may therefore represent an important psychological mechanism linking professional knowledge with willingness to teach ethnomathematics. Educators with stronger ECK and EPK may also report greater confidence in implementing culturally responsive mathematics instruction.

Postgraduate programmes devote limited contact hours, specialist expertise, and public resources to developing ethnomathematics knowledge. The instructional value of this investment, however, ultimately depends on whether participants are willing and able to use that knowledge in practice. Identifying the beliefs associated with the translation of professional knowledge into instructional willingness can therefore inform how teacher-preparation resources are prioritised. A mathematics education that systematically excludes the counting systems, measurement practices, textile geometries, and craft techniques of learners' communities may also fail to represent the cultural knowledge learners bring to school (Rosa et al., 2016; Sari and Yüce, 2020). Teachers occupy an important position in determining whether such knowledge is recognised in instruction. Examining knowledge, beliefs, and willingness together therefore speaks to both the effective use of teacher-preparation resources and the responsible representation of learners' cultural knowledge within school mathematics.

Although previous research has documented the educational benefits of ethnomathematics and examined teachers' knowledge, attitudes, or classroom practices separately (Babbitt et al., 2015; Ergene et al., 2020; Sunzuma and Maharaj, 2021; Khalil, 2023), limited empirical evidence has examined how ethnomathematics content knowledge and pedagogical knowledge are related to willingness to teach through EU and ETSE within a single model. Existing studies have largely focused on classroom interventions or student outcomes, leaving the psychological processes associated with educators' willingness to adopt ethnomathematics less well understood. Addressing this gap, the present study examines relationships among ECK, EPK, EU, ETSE, and WTE among Ghanaian postgraduate students in mathematics education. By integrating Shulman's Theory of Teacher Knowledge with Bandura's (1977) Social Cognitive Theory, the study seeks to clarify educators' readiness to implement ethnomathematics and to inform teacher education, curriculum development, and culturally responsive mathematics education in Ghana and comparable contexts.

THEORETICAL FRAMEWORK

The successful implementation of ethnomathematics in formal mathematics education requires considerably more than the inclusion of cultural examples in classroom instruction. Although ethnomathematics has been recognised as a culturally responsive approach capable of connecting school mathematics

with learners' lived experiences, many educators continue to encounter challenges in its implementation. Wulandari et al. (2024) observed that inadequate preparation among teachers often results in suboptimal ethnomathematics instruction, thereby limiting students' opportunities to experience culturally meaningful mathematics learning. Similarly, Sari and Yüce (2020) argued that curriculum constraints, limited instructional resources, and insufficient teacher preparation remain significant barriers to the effective integration of ethnomathematics into mathematics classrooms. While these institutional factors undoubtedly influence implementation, teachers' professional knowledge, confidence, and perceptions regarding ethnomathematics may be equally important determinants of whether they are willing to adopt culturally responsive instructional practices. Therefore, understanding the relationship among these factors may provide a stronger theoretical basis for explaining educators' willingness to teach ethnomathematics.

The present study is primarily anchored in Shulman's (1986) Theory of Teacher Knowledge, which proposes that effective teaching depends upon multiple categories of professional knowledge, particularly content knowledge and pedagogical knowledge. According to Shulman (1986), content knowledge refers to teachers' comprehensive understanding of the concepts, structures, and principles within a discipline, whereas pedagogical knowledge concerns the ability to organise, represent, and communicate subject matter in ways that facilitate students' understanding. These complementary knowledge domains enable teachers not only to master disciplinary content but also to transform that knowledge into meaningful learning experiences. Within mathematics education, Shulman's framework has become one of the most influential theoretical perspectives for explaining differences in instructional quality and teachers' classroom effectiveness. Applied to ethnomathematics, content knowledge refers to educators' understanding of mathematical concepts embedded within cultural practices, indigenous technologies, artistic expressions, architecture, traditional games, measurement systems, and other community-based activities. Teachers possessing strong ethnomathematics content knowledge are able to recognise that mathematical reasoning exists within diverse cultural contexts and appreciate how indigenous knowledge contributes to mathematical thinking. Ethnomathematics pedagogical knowledge, on the other hand, concerns teachers' ability to design culturally responsive lessons, select appropriate instructional strategies, develop authentic assessment activities, and facilitate classroom discussions that meaningfully connect cultural practices with formal mathematical concepts. These competencies enable educators to move beyond merely presenting cultural artefacts and instead engage learners in constructing mathematical understanding through culturally familiar experiences.

The importance of these professional knowledge domains is particularly evident within African educational contexts, where numerous cultural practices provide authentic opportunities for mathematics instruction. Rosa et al. (2016) demonstrated that African cultures contain abundant mathematical ideas embedded in textile weaving, traditional symbols, sculpture,

architecture, engineering, symmetry, patterns, and spatial reasoning. Such cultural resources provide meaningful contexts through which abstract mathematical concepts can be taught while simultaneously promoting learners' appreciation of their cultural heritage. However, the educational potential of these resources depends largely on teachers' ability to recognise their mathematical significance and translate them into classroom instruction. In this way, educators with stronger ethnomathematics content knowledge and pedagogical knowledge are expected to demonstrate greater readiness to implement ethnomathematics within formal education.

Nevertheless, knowledge alone may not be sufficient to explain teachers' instructional decisions. Many educators possess adequate professional knowledge but remain reluctant to implement innovative instructional approaches because they lack confidence in their own teaching capabilities (Ponticell, 2003; Ertmer and Ottenbreit-Leftwich, 2010; Buabeng-Andoh, 2012). To explain this psychological dimension, the present study also draws upon Bandura's (1977, 1997) Social Cognitive Theory, particularly the construct of self-efficacy. Bandura (1977) defined self-efficacy as individuals' beliefs about their capability to organise and execute the actions necessary to accomplish specific tasks successfully. Within educational settings, teacher self-efficacy influences instructional decision-making, persistence when encountering classroom challenges, openness to innovation, and commitment to improving student learning (Kwarteng et al., 2025; Lotey et al., 2026a). Teachers who believe they possess the capability to implement new instructional approaches are generally more willing to adopt them than teachers who doubt their own instructional competence.

Ethnomathematics presents unique instructional demands that make teaching self-efficacy particularly important. Effective implementation requires educators to identify authentic cultural examples, establish meaningful links between indigenous knowledge and formal mathematical concepts, facilitate culturally sensitive classroom interactions, and respond appropriately to students' diverse cultural experiences. These responsibilities require confidence in one's instructional abilities in addition to professional knowledge. Accordingly, this study conceptualises Ethnomathematics Teaching Self-Efficacy as educators' confidence in their ability to effectively teach mathematics through culturally relevant contexts and practices. Teachers who possess stronger ethnomathematics content knowledge and pedagogical knowledge are more likely to develop this confidence because they are better prepared to address the instructional complexities associated with culturally responsive mathematics teaching. Hence, teaching self-efficacy is expected to function as an important motivational factor that may link professional knowledge with willingness to teach ethnomathematics.

Beyond confidence in instructional capability, teachers'

willingness to implement ethnomathematics may also depend upon their perceptions of its educational value. In the context of this study, ethnomathematics usefulness refers to teachers' beliefs that integrating cultural knowledge into mathematics instruction enhances students' conceptual understanding, motivation, engagement, cultural identity, and overall learning outcomes. Teachers who perceive ethnomathematics as educationally valuable are more likely to invest time and effort in incorporating culturally responsive instructional practices into their teaching. Empirical evidence supports the educational benefits associated with ethnomathematics. Babbitt et al. (2015), for example, reported that Ghanaian students who learned geometry through Adinkra symbols demonstrated significantly higher learning outcomes than those receiving conventional instruction. Similarly, Gerdes (2005) illustrated how indigenous construction techniques and artistic practices in Mozambique provide rich mathematical contexts for classroom instruction, while Palhares (2012) highlighted the mathematical significance of professional activities such as boat construction and craftsmanship.

More recently, Payadnya et al. (2024) found that teachers viewed ethnomathematics as an effective means of stimulating learners' curiosity, increasing classroom engagement, and fostering appreciation for cultural diversity within mathematics education. Eglash et al. (2011) developed culturally situated software based on African fractal patterns, allowing students to investigate recursive mathematical structures through traditional hairstyles and jewellery designs. Lachney (2017) further demonstrated how African hair-braiding techniques inspired the development of the Cornrow Curves software, illustrating the potential of culturally grounded computational tools to enhance mathematics and computer science education. These studies reinforce the argument that ethnomathematics is not merely a cultural enrichment strategy but an educational innovation capable of supporting interdisciplinary learning and technological development. Teachers who understand the educational benefits of ethnomathematics are likely to perceive it as a valuable instructional approach, thereby strengthening their willingness to implement it.

Hypotheses

Our hypothesised model is presented in Figure 1. We hypothesise that (H1) ECK positively predicts EU; (H2) ECK positively predicts WTE; (H3) ECK positively predicts ETSE; (H4) EPK positively predicts EU; (H5) EPK positively predicts WTE; (H6) EPK positively predicts ETSE; (H7) EU positively predicts WTE; (H8) ETSE positively predicts WTE; (H9) EU mediates the relationship between ECK and WTE; (H10) EU mediates the relationship between EPK and WTE; (H11) ETSE mediates the relationship between ECK and WTE; and (H12) ETSE mediates the relationship between EPK and WTE.

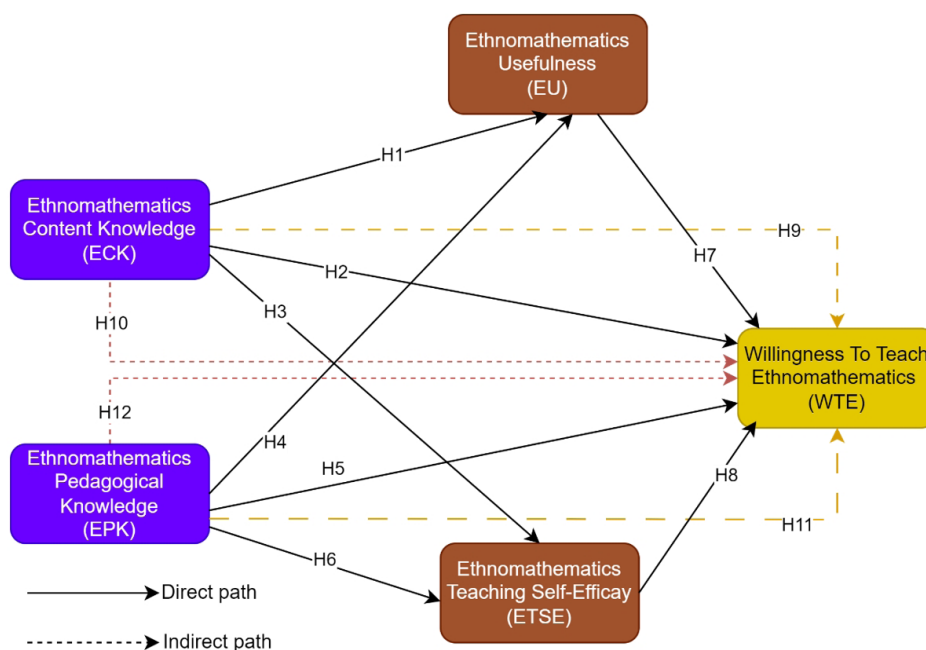


Figure 1: Conceptual model of the relationships among ethnomathematics content knowledge (ECK), ethnomathematics pedagogical knowledge (EPK), ethnomathematics usefulness (EU), ethnomathematics teaching self-efficacy (ETSE), and willingness to teach ethnomathematics (WTE) (Source: Authors' conceptualisation).

MATERIALS AND METHODS

Participants

The study involved 159 postgraduate students from mathematics education programmes at four public universities in Ghana: 141 males (88.7%) and 18 females (11.3%). Of these participants, 57 (35.8%) were enrolled in master's programmes and 102 (64.2%) in doctoral programmes. Eleven participants (6.9%) were younger than 35 years, 139 (87.4%) were aged 35–44 years, and 9 (5.7%) were older than 44 years. The study used quantitative and qualitative methods to obtain a more comprehensive understanding of factors associated with postgraduate students' willingness to teach ethnomathematics. The quantitative phase used a structured questionnaire, whereas the qualitative phase used semi-structured interviews to explore participants' experiences and provide context for the quantitative findings.

Quantitative data were collected through Google Forms using closed-ended items. Ethical review of the study was formally waived by the Institutional Ethical Review Committee of the University of Skills Training and Entrepreneurial Development (USTED), as the study constituted a PhD programme requirement by first author in a coursework titled, 'Socio-Cultural and Ethnological Foundations in Mathematics (MTE 921)' in the form of a term paper. The waiver took effect on 9 June 2025. The term paper forms part of the programme's scholarly requirements, including publication of the resulting research. Informed consent was obtained from all participants before data collection, and the data reported in this article are presented in aggregate form to protect participants' confidentiality. Participants' demographic characteristics are presented in Table 1.

Demographics		Frequency	Percentage (%)
Gender	Male	141	88.7
	Female	18	11.3
Educational level	Master's degree	57	35.8
	Doctor's degree	102	64.2
Age group (years)	< 35 years	11	6.9
	35 – 44 years	139	87.4
	> 44 years	9	5.7

Table 1: Respondents' demographic characteristics

Instruments

Items measuring Ethnomathematics Content Knowledge (ECK) and Ethnomathematics Pedagogical Knowledge (EPK) were adapted from Schmidt et al. (2009) and contextualised for ethnomathematics. ECK was operationalised as participants' self-perceived knowledge and understanding of ethnomathematical concepts, practices, and approaches

to understanding ethnomathematics. Thus, ECK concerns perceived knowledge of ethnomathematical content. Four items assessed ECK, including 'I have sufficient knowledge of ethnomathematics' and 'I know different approaches for developing my understanding of ethnomathematics.' EPK was operationalised as participants' self-perceived knowledge of approaches for integrating ethnomathematics into mathematics

teaching. In other words, EPK concerns perceived knowledge of how ethnomathematics can be taught. Four items assessed EPK, including ‘I can adapt my teaching strategies to incorporate culturally relevant mathematical practices’ and ‘I can employ a variety of instructional approaches to effectively teach ethnomathematics.’ Thus, ECK and EPK represent perceived knowledge rather than objectively demonstrated knowledge or mastery. The Ethnomathematics Usefulness (EU) scale was adapted from Khalil (2023) and comprised four items. Sample items included ‘Ethnomathematics enhances students’ understanding of the relevance of mathematics in everyday life’ and ‘Integrating ethnomathematics helps learners appreciate and value their cultural heritage.’ Ethnomathematics Teaching Self-Efficacy (ETSE) was measured using four items adapted from the teacher self-efficacy instruments developed by Tschannen-Moran and Woolfolk Hoy (2001) and contextualised for ethnomathematics instruction. ETSE was conceptualised as participants’ confidence in their capability to successfully perform specific ethnomathematics teaching tasks. In simple terms, ETSE concerns confidence in one’s capability to perform teaching tasks in ethnomathematics. Examples included ‘I am confident in my ability to teach mathematics using culturally relevant examples,’ ‘I can successfully design lessons that integrate ethnomathematics,’ and ‘I can effectively facilitate classroom discussions that connect cultural practices with mathematical concepts.’ Willingness to Teach Ethnomathematics (WTE) was measured using four items adapted from Davis (1989) and contextualised to the teaching of ethnomathematics. The adaptation involved replacing the original technology-use target behaviour with the integration of ethnomathematics into mathematics teaching, while retaining the underlying behavioural-intention content of the original items. Thus, the items were reworded to assess respondents’ willingness, intention, and readiness to incorporate ethnomathematical knowledge and practices into their mathematics teaching. The adapted items were designed to preserve the conceptual meaning of behavioural intention while aligning the measure with the specific instructional context of ethnomathematics. Sample items included ‘I intend to integrate ethnomathematics into my future mathematics teaching,’ ‘I am willing to regularly incorporate culturally relevant mathematical activities into my lessons,’ and ‘I will encourage students to explore mathematical ideas embedded in their cultural practices.’ the items were therefore intended to capture participants’ willingness or intention to incorporate ethnomathematics into their future teaching rather than their demonstrated classroom behaviour. All questionnaire items were rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

To establish content validity, the questionnaire was reviewed by a panel of ethnomathematics specialists. Their feedback addressed item clarity, relevance, appropriateness, and representativeness, and the instrument was revised accordingly before administration. The qualitative phase used a semi-structured interview guide developed from the quantitative findings and the study’s conceptual framework. Questions explored participants’ experiences of learning ethnomathematics, perceptions of its usefulness, confidence in

teaching it, factors influencing their willingness to implement it, and suggestions for improving ethnomathematics instruction in teacher-education programmes. The interviews were intended to provide contextual explanations for relationships identified in the structural equation model and to support integration of the quantitative and qualitative findings.

Data Collection Procedure

Data collection followed an explanatory sequential mixed-methods design. In the first phase, the quantitative questionnaire was administered electronically through Google Forms. With the assistance of postgraduate class representatives, the survey link was distributed through WhatsApp and Telegram to students in mathematics education programmes at the selected Ghanaian universities. Snowball sampling was used, with initial participants invited to share the survey link with other eligible postgraduate students within their institutions. Participation was voluntary, and electronic informed consent was obtained before questionnaire completion. Quantitative data collection lasted two weeks. After the quantitative analysis, eight participants were purposively selected for the qualitative phase to provide variation in reported levels of knowledge, usefulness, self-efficacy, and willingness to teach ethnomathematics. Semi-structured interviews were conducted by telephone and, with participants’ permission, audio-recorded and transcribed verbatim. Participants were informed of the study purpose, the voluntary nature of participation, their right to withdraw without penalty, and the confidentiality of the information provided. Personal identifiers were removed during analysis, and data were stored securely. The quantitative dataset contained no missing values because responses to all Google Forms items were required.

Data Analysis

Completed questionnaires were downloaded from Google Forms into Microsoft Excel for preliminary screening and data cleaning and were subsequently imported into IBM SPSS Statistics version 32 for descriptive analysis. The dataset was examined for outliers, normality, and multicollinearity before the structural analysis. Covariance-based structural equation modelling (CB-SEM) with latent variables was then conducted using IBM SPSS AMOS version 31 with maximum likelihood estimation. The measurement model was first evaluated using confirmatory factor analysis (CFA). Internal consistency reliability was assessed using Cronbach’s alpha, McDonald’s omega, and composite reliability (CR). Convergent validity was evaluated using average variance extracted (AVE) and standardised factor loadings, while discriminant validity was assessed using the Fornell–Larcker criterion. Overall model fit was evaluated using the chi-square statistic, chi-square/degrees of freedom ratio (χ^2/df), comparative fit index (CFI), Tucker–Lewis index (TLI), and root mean square error of approximation (RMSEA). Because the data exhibited evidence of multivariate non-normality (see Table 5), bootstrap procedures were employed to strengthen statistical inference. A Bollen–Stine bootstrap procedure based on 5,000 bootstrap samples was used specifically to assess the overall model fit under non-normality. The Bollen–Stine procedure was distinguished

from the bootstrap procedure used for parameter inference. Bias-corrected (BC) 95% bootstrap confidence intervals and corresponding bootstrap *p*-values were used to evaluate the statistical significance of the direct and indirect structural effects. An effect was considered statistically significant when its BC 95% confidence interval did not include zero. Thus, although conventional maximum-likelihood estimates were obtained, bootstrap confidence intervals and bootstrap-based *p*-values were used as the primary basis for testing the hypothesised structural relationships.

Qualitative data were analysed after the quantitative phase to explain and elaborate on the statistical findings in accordance with the sequential explanatory mixed-methods design. Audio-recorded interviews were transcribed verbatim and analysed using reflexive thematic analysis following the six phases proposed by Braun and Clarke (2006): familiarisation with the data, generation of initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. The analysis was primarily inductive, with the study questions and quantitative findings informing the subsequent integration and interpretation of the qualitative evidence rather than serving as predetermined themes. Coding was iterative and involved repeated engagement with the transcripts, initial codes, candidate themes, and the broader dataset. Themes were developed from recurring patterns of meaning across participants' accounts rather than being constructed to correspond directly to individual SEM paths or hypotheses. Reflexivity was maintained throughout the analytic process by critically examining the researcher's assumptions, theoretical commitments, and expectations arising from the quantitative findings. The researchers documented analytic decisions and reflected on how prior expectations might influence the interpretation of participants' accounts. Integration occurred at the interpretation stage

by comparing the qualitative themes with the quantitative findings. The qualitative evidence was used to explain, elaborate on, and contextualise statistically significant or theoretically important relationships identified through SEM. Accordingly, the qualitative themes were not intended to reproduce the quantitative model but to provide contextual insight into the processes and conditions that participants associated with the development of ethnomathematics knowledge, perceived usefulness, teaching self-efficacy, and willingness to teach ethnomathematics. Table 8 therefore presents qualitative themes and explanatory categories alongside the quantitative findings to which they are relevant. The quotations are illustrative excerpts that demonstrate participants' perspectives and are not treated as evidence of statistical effects. Evidence for the statistical relationships and mediation effects is derived from the SEM analysis, whereas the interview findings provide contextual and explanatory insight into how participants understood and experienced these relationships.

RESULTS

Validity and Reliability of the Measurement Model

The measurement model was evaluated using standardised factor loadings, Cronbach's alpha, McDonald's omega, composite reliability (CR), and average variance extracted (AVE), as reported in Tables 2 and 3. EU4 was removed because its factor loading was below the threshold of 0.500. The remaining standardised loadings ranged from 0.600 to 0.900. Reliability coefficients for all constructs exceeded 0.700, supporting internal consistency, and AVE values exceeded 0.500, supporting convergent validity (Hair et al., 2017; Lotey et al., 2026b).

<i>N</i>	Variable	Cronbach's Alpha	McDonald's Omega	Composite Reliability
4	ECK	0.904	0.905	0.905
4	EPK	0.820	0.819	0.819
4	ETSE	0.895	0.895	0.902
3	EU	0.818	0.819	0.817
4	WTE	0.864	0.865	0.866

Table 2: Internal consistency reliability statistics

ECK had a Cronbach's alpha of 0.904, McDonald's omega of 0.905, CR of 0.905, and AVE of 0.704. Corresponding values were 0.820, 0.819, 0.819, and 0.536 for EPK; 0.895, 0.895, 0.902, and 0.698 for ETSE; 0.818, 0.819, 0.817, and 0.600 for EU; and 0.864, 0.865, 0.866, and 0.619 for WTE. Overall, these results indicate acceptable internal consistency and convergent validity for the retained measurement model.

Discriminant Validity and Inter-Construct Correlations

Discriminant validity was assessed using the Fornell–Larcker criterion, as recommended by Fornell and Larcker (1981) and Ab Hamid et al. (2017). Under this criterion, the square root of

the AVE for each construct should exceed its correlations with the other constructs. As presented in Table 3, the square roots of the AVEs (bold diagonal values) ranged from 0.732 to 0.839 and exceeded the corresponding inter-construct correlations, supporting discriminant validity.

All study constructs were positively correlated ($p < 0.05$). ECK correlated positively with EPK ($r = 0.297$), ETSE ($r = 0.518$), EU ($r = 0.582$), and WTE ($r = 0.392$). ETSE was also positively correlated with EU ($r = 0.646$) and WTE ($r = 0.640$), while EU correlated positively with WTE ($r = 0.641$). These bivariate associations provide preliminary support for the proposed relationships while remaining below the corresponding Fornell–Larcker thresholds.

Variable	ECK	EPK	ETSE	EU	WTE
ECK	0.839				
EPK	0.297*	0.732			
ETSE	0.518*	0.321*	0.835		
EU	0.582*	0.323*	0.646*	0.775	
WTE	0.392*	0.355*	0.640*	0.641*	0.787

*Correlation is significant at 0.05; boldface= square root of AVE

Table 3: Correlations and Discriminant Validity

Descriptive Statistics

Table 4 presents descriptive statistics for the five study constructs. Mean scores ranged from 3.58 to 4.42, indicating

generally favourable self-reported perceptions, while standard deviations ranged from 0.784 to 0.883, indicating moderate variability in responses.

N	Variable	Mean	Standard Deviation	Tolerance	VIF
4	ECK	4.086	0.784	0.687	1.456
4	EPK	4.420	0.883	0.903	1.107
4	ETSE	4.059	0.829	0.632	1.583
3	EU	3.774	0.785	0.619	1.616
4	WTE	3.576	0.846	–	–

Table 4: Mean, Standard Deviation, and Collinearity Statistics

Potential multicollinearity among the predictor variables was assessed using tolerance and variance inflation factor (VIF) statistics. Tolerance values ranged from 0.619 to 0.903 and VIF values from 1.107 to 1.616, well within the recommended thresholds of tolerance > 0.20 and VIF < 5.00 (Kwarteng et al., 2025). These results indicate that problematic multicollinearity was not evident among the predictors.

Structural Model Results

The measurement model showed acceptable overall fit to the data. The chi-square-to-degrees-of-freedom ratio was 1.799 ($\chi^2 = 251.856$, $df = 140$). The incremental fit indices were strong, with $CFI = 0.985$ and $TLI = 0.982$, while $GFI = 0.963$ indicated satisfactory absolute fit. The residual indices were also low ($RMR = 0.026$; $SRMR = 0.028$). Based on $\chi^2 = 251.856$, $df = 140$, and $N = 159$, $RMSEA$ was 0.071 (90% CI [0.057, 0.085]), with $PCLOSE = 0.009$. Thus, the $RMSEA$ indicated moderate rather than close approximate fit, whereas the remaining indices supported an acceptable measurement model. The standardised factor loadings ranged from 0.610 to 0.893, exceeding the commonly recommended minimum of 0.50. The AVE values were 0.704 for ECK, 0.536 for EPK, 0.698 for ETSE, 0.600 for EU, and 0.619 for WTE, demonstrating adequate convergent validity across all constructs. Univariate normality was examined using skewness and kurtosis. Following Kline (2023), the observed skewness values (–1.553 to –0.413) and kurtosis values (–0.280 to 2.943) were within commonly accepted limits. However, the multivariate kurtosis critical ratio ($CR = 85.379$) indicated substantial departure from multivariate normality. The analysis therefore used 5,000 bootstrap resamples for more robust inference, including bias-corrected confidence intervals for indirect effects; the Bollen–Stine procedure was also used in AMOS in relation to overall model fit. The confirmatory factor analysis is presented in Figure 2.

The final structural model showed mixed fit to the data: $\chi^2 = 397.654$, $df = 143$, $p < 0.001$, $CMIN/DF = 2.781$. $GFI = 0.944$, $AGFI = 0.926$, $CFI = 0.966$, $TLI = 0.960$, and $SRMR = 0.060$ suggested acceptable fit on several indices. However, based on $\chi^2 = 397.654$, $df = 143$, and $N = 159$, $RMSEA$ was 0.106 (90% CI [0.094, 0.119]), with $PCLOSE < 0.001$; the Bollen–Stine bootstrap p-value was also 0.001. These latter indices indicate notable approximate model misfit, and the structural coefficients should therefore be interpreted with appropriate caution. The exogenous latent variables ECK and EPK were allowed to covary freely in the structural model to avoid specification error. The standardized covariance was 0.29 ($p < .001$), consistent with their bivariate correlation ($r = 0.297$). Additionally, correlated residuals between EPK3–EPK4 and ETSE1–ETSE4 were retained in the structural model based on modification indices and overlapping item wording. These specifications are visually represented in Figure 3.

The coefficients of determination (R^2) indicated that ECK and EPK jointly explained 36.0% of the variance in EU ($R^2 = 0.360$) and 30.0% of the variance in ETSE ($R^2 = 0.300$). ECK, EPK, EU, and ETSE collectively explained 46.0% of the variance in WTE ($R^2 = 0.460$), indicating moderate explanatory power for the proposed structural model.

As shown in Table 6, ECK was positively associated with EU ($\beta = 0.551$, BC 95% CI [0.477, 0.619], $p < 0.001$) and ETSE ($\beta = 0.487$, BC 95% CI [0.415, 0.554], $p < 0.001$). ECK also showed a statistically significant negative direct effect on WTE ($\beta = -0.108$, BC 95% CI [–0.207, –0.013], $p = 0.025$). Also, EPK was positively associated with EU ($\beta = 0.168$, BC 95% CI [0.081, 0.254], $p < 0.001$) and ETSE ($\beta = 0.188$, BC 95% CI [0.109, 0.268], $p < 0.001$). EPK also had a statistically significant positive direct effect on WTE ($\beta = 0.119$, BC 95% CI [0.037, 0.193], $p = 0.005$).

MFI: <i>CMIN</i> = 251.856; <i>DF</i> = 140; <i>CMIN/DF</i> = 1.799; <i>TLI</i> = 0.982; <i>CFI</i> = 0.985; <i>GFI</i> = 0.963; <i>RMR</i> = 0.026; <i>RMSEA</i> = 0.071; <i>SRMR</i> = .028; <i>PClose</i> = 0.009	Std. Estimate	Univariate Normality	
		Skewness	Kurtosis
ECK: <i>AVE</i> = 0.704			
ECK1	0.774	-1.409	2.943
ECK2	0.893	-1.300	2.182
ECK3	0.815	-1.211	1.931
ECK4	0.870	-1.443	2.901
EPK: <i>AVE</i> = 0.536			
EPK1	0.856	-0.432	0.774
EPK2	0.754	-0.607	0.692
EPK3	0.684	-0.413	0.344
EPK4	0.610	-0.772	0.399
ETSE: <i>AVE</i> = 0.698			
ETSE1	0.833	-1.295	1.730
ETSE2	0.852	-1.329	2.173
ETSE3	0.840	-1.390	2.448
ETSE4	0.815	-1.553	2.902
EU: <i>AVE</i> = 0.600			
EU1	0.840	-0.700	0.544
EU2	0.770	-0.840	0.653
EU3	0.708	-1.123	1.738
WTE: <i>AVE</i> = 0.619			
WTE1	0.762	-0.585	-0.280
WTE2	0.848	-0.883	0.684
WTE3	0.813	-0.981	0.988
WTE4	0.717	-0.719	0.350
		Multivariate kurtosis	Critical ratio
Multivariate normality		182.321	85.379

Table 5: Confirmatory Factor Analysis, Model fit indices, and Normality Test

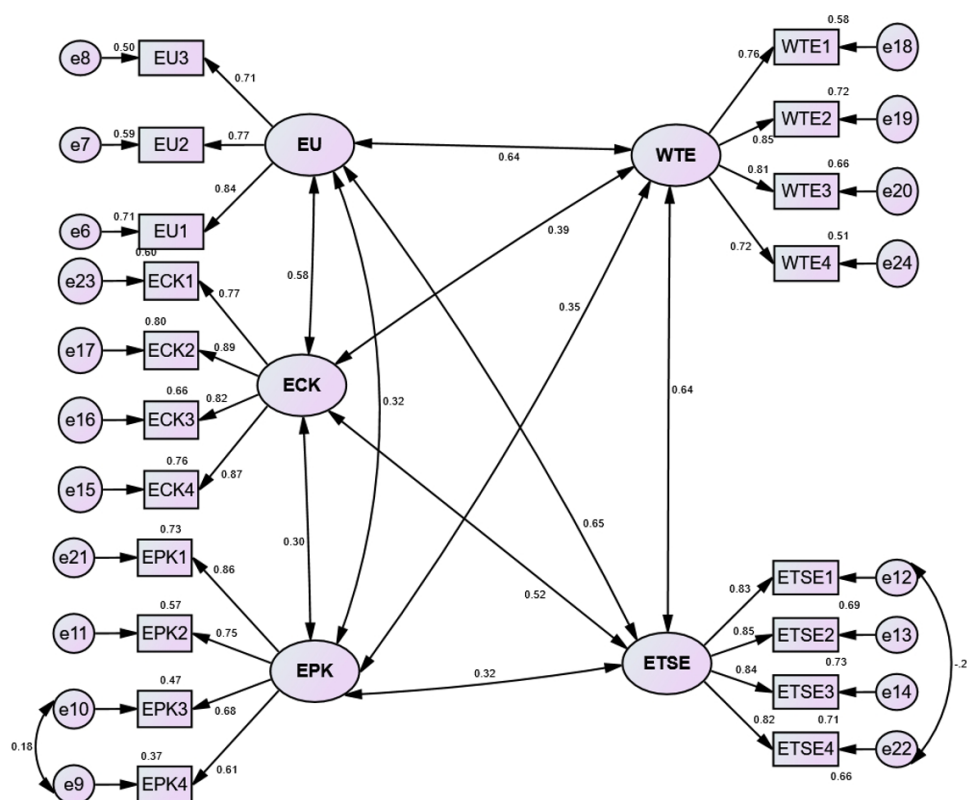


Figure 2: Confirmatory factor analysis

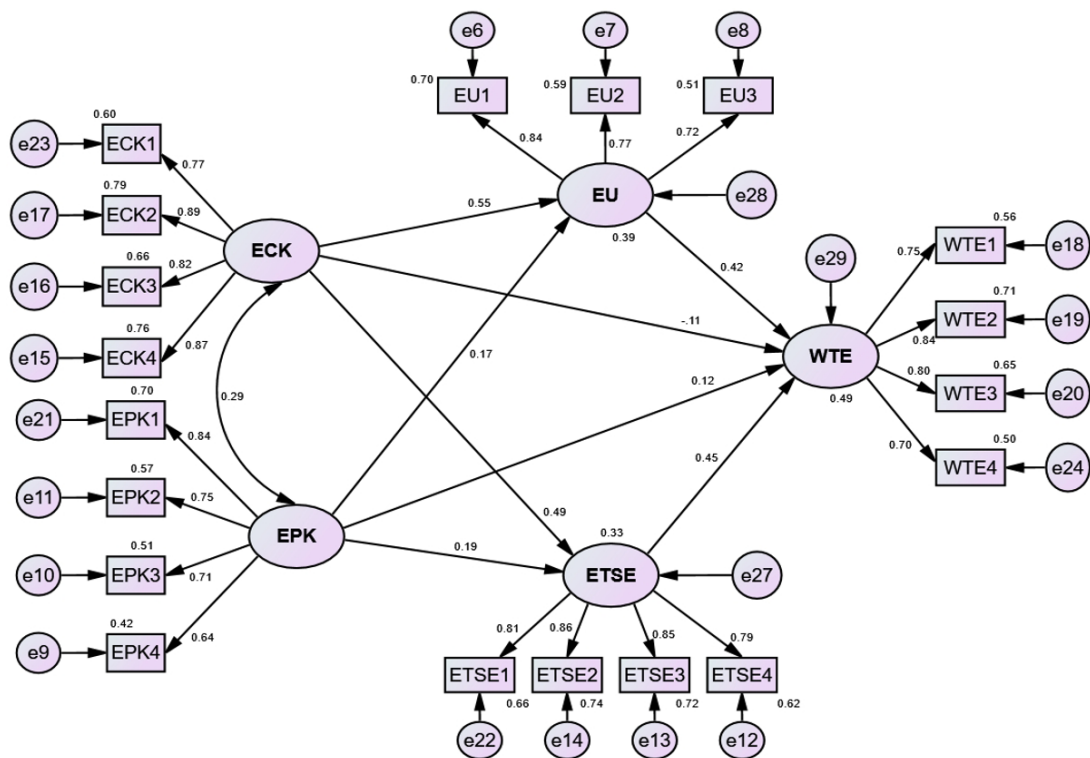


Figure 3: Structural Equation Modeling with latent variables

Test	Hypotheses	Direct effect	Total effect	BC 95% CI	Bootstrap p-value	Decision
H1	ECK→EU	0.551	0.569	[0.477, 0.619]	< 0.001	Supported
H2	ECK→WTE	-0.108	0.340	[-0.207, -0.013]	0.025	Supported
H3	ECK→ETSE	0.487	0.487	[0.415, 0.554]	< 0.001	Supported
H4	EPK→EU	0.168	0.168	[0.081, 0.254]	< 0.001	Supported
H5	EPK→WTE	0.119	0.273	[0.037, 0.193]	0.005	Supported
H6	EPK→ETSE	0.188	0.188	[0.109, 0.268]	< 0.001	Supported
H7	EU→WTE	0.419	0.419	[0.313, 0.519]	< 0.001	Supported
H8	ETSE→WTE	0.446	0.446	[0.353, 0.534]	< 0.001	Supported

Table 6: Hypothesis Testing of Direct and Total Effects ($p < 0.05$)

EU was positively associated with WTE ($\beta = 0.419$, BC 95% CI [0.313, 0.519], $p < 0.001$), indicating that participants who perceived greater educational value in ethnomathematics also reported greater willingness to use it in future teaching. ETSE was likewise positively associated with WTE ($\beta = 0.446$, BC 95% CI [0.353, 0.534], $p < 0.001$), indicating that stronger confidence in teaching ethnomathematics was associated with stronger instructional intentions.

Indirect effects were examined using 5,000 bias-corrected bootstrap resamples. As presented in Table 7, all four specific indirect effects were statistically significant. The indirect association between ECK and WTE through EU was $\beta = 0.244$ (95% BC CI [0.174, 0.326], $p = 0.001$), and the indirect association through ETSE was $\beta = 0.229$ (95% BC CI [0.173, 0.298], $p = 0.001$). For EPK, the indirect association with WTE through EU was $\beta = 0.070$ (95% BC CI [0.033, 0.118], $p = 0.001$), and the indirect association through ETSE was $\beta = 0.083$ (95% BC CI [0.046, 0.129], $p = 0.001$). The total

indirect effect of ECK on WTE was $\beta = 0.448$ (95% BC CI [0.367, 0.535], $p < .001$), and the total indirect effect of EPK on WTE was $\beta = 0.154$ (95% BC CI [0.094, 0.221], $p < .001$). This result shows that while the specific indirect effects of EPK through both EU and ETSE were statistically significant, their magnitudes ($\beta = 0.070$ and $\beta = 0.083$, respectively) were notably smaller than those observed for ECK ($\beta = 0.244$ and $\beta = 0.229$). This disparity is statistically explained by the weaker direct associations of EPK with EU ($\beta = 0.168$) and ETSE ($\beta = 0.188$) compared to the strong paths from ECK to these mediators. This suggests that ethnomathematics content knowledge may serve as the foundational cognitive basis for perceiving the approach as useful and developing confidence to teach it. While pedagogical knowledge is essential for translating that content into classroom practice, as evidenced by its direct association with WTE ($\beta = 0.119$), it may not independently drive beliefs about usefulness or self-efficacy as strongly as content knowledge does.

	Hypotheses	Indirect Effect	Lower BC	Upper BC	P-value
H9	ECK→EU→WTE	0.244	0.174	0.326	0.001
H10	EPK→EU→WTE	0.070	0.033	0.118	0.001
H11	ECK→ETSE→WTE	0.229	0.173	0.298	0.001
H12	EPK→ETSE→WTE	0.083	0.046	0.129	0.001

BC = bias-corrected confidence interval.

Table 7: Indirect Effect

DISCUSSION

ECK was positively associated with EU, suggesting that participants who reported stronger knowledge of ethnomathematical concepts were also more likely to recognise the educational relevance of connecting mathematics with culture. This finding is consistent with D'Ambrosio's (1985) argument that multiple cultural perspectives can broaden educators' conceptions of mathematics and with Rosa et al.'s (2016) emphasis on mathematical ideas embedded in communities. Interview accounts complemented this pattern: participants described greater exposure to ethnomathematics as helping them appreciate its potential contribution to conceptual understanding and cultural identity. The findings therefore suggest that teacher preparation should make the educational relevance of ethnomathematical knowledge explicit rather than treating ethnomathematics as content to be learned in isolation. Authentic local examples, such as Adinkra symmetry or Akan counting and measurement practices, may help participants connect theoretical knowledge with potential instructional use (Babbitt et al., 2015; Owusu-Darko et al., 2023).

ECK was also positively associated with ETSE, indicating that participants with greater self-reported ethnomathematics knowledge tended to report greater confidence in teaching culturally responsive mathematics. This pattern is consistent with Bandura's (1977) account of self-efficacy and with Shulman's (1986) emphasis on professional knowledge in teaching. Interview participants similarly described ethnomathematics learning as reducing uncertainty and increasing confidence in developing culturally relevant lessons. These accounts support the inclusion of applied components such as lesson planning, classroom simulation, and microteaching alongside conceptual coverage in teacher preparation (Bandura, 1997; Ertmer and Ottenbreit-Leftwich, 2010).

ECK showed a small but statistically significant negative direct effect on WTE after EU and ETSE were included in the model ($\beta = -0.108$, BC 95% CI $[-0.207, -0.013]$, $p = 0.025$), despite having a positive indirect effect ($\beta = 0.448$, BC 95% CI $[0.367, 0.535]$, $p < 0.001$) and a positive total effect ($\beta = 0.340$). This pattern indicates inconsistent mediation, whereby the direct and indirect effects operate in opposite directions. The finding should not be interpreted as evidence that greater ethnomathematics content knowledge reduces willingness to teach ethnomathematics but rather, the positive total effect indicates that the overall association between ECK and WTE remains favourable, with the positive indirect pathways through perceived usefulness and teaching self-efficacy outweighing the negative direct pathway. One possible explanation, suggested by participants' interview accounts,

is that greater knowledge may also heighten awareness of contextual conditions that could constrain the implementation of ethnomathematics, including curriculum demands, assessment expectations, limited resources, and institutional support. This interpretation is tentative, however, because the cross-sectional design does not permit conclusions about the temporal or causal mechanism underlying the negative direct effect. The finding nevertheless suggests that ethnomathematics content knowledge alone may not be sufficient to translate into willingness to teach when perceived usefulness and teaching self-efficacy are not simultaneously strengthened. The qualitative findings provide contextual insight into this possibility, with participants describing curricular and assessment conditions as potential constraints on the enactment of culturally inclusive mathematics teaching (Papagiannis et al., 1982; Sarı and Yüce, 2020).

EPK was positively associated with EU, ETSE, and WTE. Participants who reported stronger knowledge of strategies for integrating culture into mathematics lessons were more likely to value ethnomathematics, feel confident implementing it, and express stronger intentions to use it. This pattern is consistent with Wulandari et al. (2024), who emphasised the importance of both cultural knowledge and appropriate pedagogy for ethnomathematics implementation. Interview participants also described demonstrations, lesson planning, and discussion of culturally responsive strategies as increasing their appreciation of ethnomathematics and their readiness to use it. These findings suggest that pedagogical preparation is an important component of ethnomathematics-related teacher education, although the present study does not directly compare the efficiency of alternative programme designs (Shulman, 1986; Schmidt et al., 2009).

Both EU and ETSE were positively associated with WTE. Participants who perceived ethnomathematics as beneficial for conceptual understanding, learner engagement, and cultural appreciation expressed stronger intentions to integrate it into future teaching. Similarly, greater confidence in teaching ethnomathematics was associated with greater willingness to implement culturally responsive instruction. These findings are consistent with research indicating that educators' beliefs about innovations can be important alongside professional knowledge (Wallace and Priestley, 2011; Buabeng-Andoh, 2012; Serdyukov, 2017; Boudouaia et al., 2024). Guided practice, mentoring, and reflection may therefore complement existing professional knowledge by strengthening the beliefs and confidence associated with instructional intention (Davis, 2003; Ponticell, 2003).

EU and ETSE each showed significant indirect effects in the relationships of ECK and EPK with WTE. Statistically, these findings indicate that part of the association between

professional knowledge and willingness operates through participants' perceived value of ethnomathematics and their confidence in teaching it. Because the data are cross-sectional, these indirect effects should be interpreted as theoretically specified associations rather than evidence of a confirmed causal sequence. Nevertheless, the pattern highlights perceived usefulness and teaching self-efficacy as potentially important targets within teacher preparation. Content and pedagogical activities may have greater instructional relevance when they are accompanied by opportunities to recognise the value of ethnomathematics and to develop confidence in using it (Ertmer and Ottenbreit-Leftwich, 2010; Boudouaia et al., 2024). Broader curricular and assessment conditions may also determine whether such willingness can subsequently be enacted in practice.

Limitations of the Study

Several limitations qualify the interpretation of these findings. First, participants were postgraduate students in mathematics education at four public universities and were recruited through snowball sampling. The sample therefore does not represent the wider population of mathematics teachers and was disproportionately male (88.7%) and concentrated in the 35–44-year age group (87.4%). The observed levels of knowledge, usefulness, self-efficacy, and willingness should consequently not be interpreted as national estimates. The restricted composition of the sample may also have reduced variability in the constructs and affected the magnitude of observed associations. Replication with a more diverse, probability-based sample of pre-service and in-service teachers would strengthen generalisability.

Second, the cross-sectional design precludes causal conclusions. Although the structural model specifies theoretically informed directional relationships among the constructs, the data were collected at a single time point and therefore cannot establish temporal precedence or rule out reciprocal relationships among ethnomathematics knowledge, beliefs, and willingness to teach. Accordingly, the reported indirect effects should be interpreted as statistical mediation within the hypothesised SEM rather than as evidence of temporally ordered or causal mediation mechanisms. The study also cannot determine whether participants' reported willingness to teach ethnomathematics would be sustained under the curricular, assessment, resource, and institutional conditions of actual classroom practice. Longitudinal or experimental research is therefore needed to examine temporal ordering and causal mechanisms more directly (Kline, 2023).

Third, all constructs were measured by self-report at a single point in time, creating potential social-desirability and common-method bias (Podsakoff et al., 2003). Anonymity assurances may have reduced some of these risks but cannot eliminate them. Restricted sample variability could attenuate associations, whereas common-method bias could inflate them; the estimated coefficients should therefore be interpreted cautiously. More importantly, WTE represents an intention rather than observed teaching behaviour. The findings cannot establish whether participants subsequently enact culturally responsive

mathematics teaching. Classroom observations and lesson-plan analyses would provide stronger evidence of enactment (Wallace and Priestley, 2011; Buabeng-Andoh, 2012).

Fourth, the qualitative component involved only eight participants and was designed to contextualise the quantitative findings rather than generate theory. Consequently, the interpretation of the negative direct ECK–WTE association as reflecting heightened awareness of implementation constraints remains tentative. Alternative explanations, including statistical suppression, cannot be excluded. Further qualitative work focused specifically on participants with high knowledge but low willingness would provide a stronger test of this interpretation.

Finally, the model did not include contextual factors such as school leadership support, availability of culturally grounded teaching materials, or access to community knowledge holders (Mania and Alam, 2021; Wulandari et al., 2024). Although the model accounted for 46.0% of the variance in WTE, some associations attributed to individual beliefs may also reflect unmeasured institutional conditions. Future research incorporating both individual- and institution-level predictors could provide a more complete account of willingness to enact culturally responsive mathematics teaching.

CONCLUSION

This study examined relationships among ethnomathematics content knowledge (ECK), pedagogical knowledge (EPK), perceived usefulness (EU), teaching self-efficacy (ETSE), and willingness to teach ethnomathematics (WTE) among Ghanaian postgraduate students. ECK and EPK were positively associated with EU and ETSE. WTE was positively predicted by EPK, EU, and ETSE, while ECK showed a small negative direct association with WTE but a positive total effect. EU and ETSE also showed significant indirect effects linking both knowledge constructs with WTE. The qualitative findings complemented these statistical patterns by indicating that exposure to ethnomathematics and pedagogical preparation may strengthen participants' appreciation of, and confidence in, culturally responsive instruction, while curricular and assessment constraints may limit implementation. Overall, the findings suggest that knowledge acquisition should be accompanied by opportunities to develop perceived value and teaching confidence.

In practical terms, postgraduate programmes may benefit from integrating ethnomathematics content with lesson design, guided practice, and microteaching to support both competence and confidence. However, this study did not directly measure programme efficiency, contact-hour productivity, or cost-effectiveness. Therefore, claims regarding the comparative efficiency of specific preparation components are framed strictly as implications and hypotheses for future research. Curriculum developers and policymakers should consider contextual barriers, including assessment expectations and access to appropriate instructional resources. Sustainable integration of culturally responsive mathematics education is likely to require teacher preparation that addresses knowledge, beliefs, and instructional confidence alongside supportive curricular and institutional conditions.

Quantitative Finding	Illustrative Contextual Insight	Illustrative Participant Quotation	Integrated Interpretation
ECK→EU ($\beta = 0.551$)	Recognising the cultural and educational value of ethnomathematics.	<i>'The more I studied ethnomathematics, the more I realised that mathematics is deeply rooted in our culture and should be part of classroom teaching.'</i> (Participant 7)	Participants' accounts suggest that deeper knowledge of ethnomathematics can broaden teachers' recognition of its cultural and educational value, helping to explain the positive association between ECK and EU.
ECK→ETSE ($\beta = 0.487$)	Subject knowledge as a source of teaching confidence.	<i>'Understanding the mathematical concepts behind our cultural practices gave me the confidence that I could successfully teach them to students.'</i> (Participant 2)	Participants associated understanding the mathematical basis of cultural practices with greater confidence in their capacity to teach ethnomathematics, providing an experiential explanation for the positive ECK→ETSE relationship.
ECK→WTE ($\beta = -0.108$)	Knowledge-practice gap under contextual constraints.	<i>'Although I now know about ethnomathematics, I may still hesitate to teach it because I am unsure whether the curriculum and assessment system will support it.'</i> (Participant 4)	The qualitative evidence helps contextualise the small negative direct effect by suggesting that knowledge alone may not translate into willingness when teachers perceive curriculum, assessment, or institutional constraints.
EPK→EU ($\beta = 0.168$) and EPK→WTE ($\beta = 0.119$)	Pedagogical preparation and perceived classroom value.	<i>'When we learned practical ways of integrating culture into mathematics lessons, I began to see how useful ethnomathematics could be for improving students' understanding.'</i> (Participant 5)	Participants described practical pedagogical preparation as helping them recognise the instructional value of ethnomathematics and its feasibility for classroom use, providing a common explanation for the positive EPK→EU and EPK→WTE relationships.
EPK→ETSE ($\beta = 0.188$)	Instructional preparation as confidence building.	<i>'Preparing lesson plans based on cultural examples made me feel capable of teaching ethnomathematics in my future classroom.'</i> (Participant 3)	Participants' accounts indicate that opportunities to practise instructional strategies and lesson planning can strengthen perceived capability to implement ethnomathematics, supporting the positive EPK→ETSE relationship.
EU→WTE ($\beta = 0.419$)	Perceived educational relevance as a motivation for adoption.	<i>'Students learn better when mathematics reflects their own cultural experiences, so I definitely intend to use ethnomathematics whenever possible.'</i> (Participant 2)	Participants' recognition of the relevance and educational benefits of ethnomathematics was associated with stronger intentions to incorporate it into teaching, helping to explain the positive EU→WTE relationship.
ETSE→WTE ($\beta = 0.446$)	Teaching confidence and readiness for implementation.	<i>'Because I feel confident designing ethnomathematics lessons, I have no hesitation about teaching them.'</i> (Participant 6)	Participants linked confidence in lesson design and delivery with readiness to implement ethnomathematics, providing qualitative context for the positive ETSE→WTE relationship.
ECK→WTE indirect effect ($\beta = 0.448$)	Knowledge translated into willingness through perceived value and confidence.	<i>'Understanding ethnomathematics helped me appreciate its value, and that is what motivated me to teach it in the future.'</i> (Participant 8)	The quantitative analysis established a significant indirect effect of ECK on WTE. Participants' accounts suggest that developing content knowledge may increase willingness by enhancing perceived value and teaching confidence.
EPK→WTE indirect effect ($\beta = 0.154$)	Pedagogical preparation translated into willingness through value and confidence.	<i>'Practising how to teach ethnomathematics gave me the confidence to implement it when I become a lecturer.'</i> (Participant 2)	The quantitative analysis established a significant indirect effect of EPK on WTE. Participants' accounts suggest that pedagogical preparation may foster willingness by increasing perceived usefulness and confidence in implementation.

Table 8: Integration of Quantitative and Qualitative Findings

REFERENCES

- Ab Hamid, M. R., Sami, W. and Mohmad Sidek, M.H. (2017) 'Discriminant validity assessment: Use of Fornell & Larcker criterion versus HTMT criterion', *Journal of Physics: Conference Series*, Vol. 890, No. 1, p. 012163. <https://doi.org/10.1088/1742-6596/890/1/012163>
- Babbitt, W., Lachney, M., Bulley, E. and Eglash, R. (2015) 'Adinkra mathematics: A study of ethnocomputing in Ghana', *Multidisciplinary Journal of Educational Research*, Vol. 5, No. 2, pp. 110–135. <https://doi.org/10.17583/remic.2015.1399>
- Bandura, A. (1977) 'Self-efficacy: Toward a unifying theory of behavioral change', *Psychological Review*, Vol. 84, No. 2, pp. 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1997) *Self-efficacy: The exercise of control*, New York: W. H. Freeman and Company.
- Boudouaia, A., AL-Qadri, A. H., Houichi, A. and Diafi, S. (2024) 'A study on the effect of school culture on teachers' commitment to curriculum implementation: The mediating role of self-efficacy and job satisfaction', *Heliyon*, Vol. 10, No. 8, p. e29183. <https://doi.org/10.1016/j.heliyon.2024.e29183>
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, Vol. 3, No. 2, pp. 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Buabeng-Andoh, C. (2012) 'Factors influencing teachers' adoption and integration of information and communication technology into teaching: A review of the literature', *International Journal of Education and Development using Information and Communication Technology*, Vol. 8, No. 1, pp. 136–155. Available at: <https://www.learntechlib.org/p/188018/>
- D'Ambrosio, U. (1985) 'Ethnomathematics and its place in the history and pedagogy of mathematics', *For the Learning of Mathematics*, Vol. 5, No. 1, pp. 44–48.
- D'Ambrosio, U. and Knijnik, G. (2020) 'Ethnomathematics', in: Lerman, S. (ed.), *Encyclopedia of mathematics education*, 2nd edition, Cham: Springer International Publishing, pp. 283–288. https://doi.org/10.1007/978-3-030-15789-0_60
- D'Ambrosio, U. and Rosa, M. (2017) 'Ethnomathematics and its pedagogical action in mathematics education', in: Rosa, M., Shirley, L., Gavarrete, M. E. and Alangui, W. V. (eds.), *Ethnomathematics and its diverse approaches for mathematics education*, Cham: Springer International Publishing, pp. 285–305. https://doi.org/10.1007/978-3-319-59220-6_12
- Davis, K.S. (2003) 'Change is hard: What science teachers are telling us about reform and teacher learning of innovative practices', *Science Education*, Vol. 87, No. 1, pp. 3–30. <https://doi.org/10.1002/sce.10037>
- Eglash, R., Krishnamoorthy, M., Sanchez, J. and Woodbridge, A. (2011) 'Fractal simulations of African design in pre-college computing education', *ACM Transactions on Computing Education*, Vol. 11, No. 3, pp. 1–14. <https://doi.org/10.1145/2037276.2037281>
- Ergene, Ö., Ergene, B.Ç. and Yazıcı, E.Z. (2020) 'Ethnomathematics activities: Reflections from the design and implementation process', *Turkish Journal of Computer and Mathematics Education*, Vol. 11, No. 2, pp. 402–437. <https://doi.org/10.16949/turkbilm2.688780>
- Ertmer, P. A. and Ottenbreit-Leftwich, A. T. (2010) 'Teacher technology change: How knowledge, confidence, beliefs, and culture intersect', *Journal of Research on Technology in Education*, Vol. 42, No. 3, pp. 255–284. <https://doi.org/10.1080/15391523.2010.10782551>
- Fornell, C. and Larcker, D. F. (1981) 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, Vol. 18, No. 1, pp. 39–50. <https://doi.org/10.2307/3151312>
- Gerdes, P. (2005) 'Ethnomathematics, geometry and educational experiences in Africa', *Africa Development*, Vol. 30, No. 3, pp. 315–336. <https://doi.org/10.4314/ad.v30i3.22229>
- Hair, J. F., Matthews, L. M., Matthews, R. L. and Sarstedt, M. (2017) 'PLS-SEM or CB-SEM: Updated guidelines on which method to use', *International Journal of Multivariate Data Analysis*, Vol. 1, No. 2, p. 107. <https://doi.org/10.1504/ijmda.2017.10008574>
- Hu, L. T. and Bentler, P. M. (1999) 'Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives', *Structural Equation Modeling*, Vol. 6, No. 1, pp. 1–55. <https://doi.org/10.1080/10705519909540118>
- Khalil, I. A. (2023) 'Mathematics teachers' attitudes towards integrating ethnic mathematics in teaching mathematical topics: A mixed methods study', *International Journal of Instruction*, Vol. 16, No. 4, pp. 1061–1080. <https://doi.org/10.29333/iji.2023.16458a>
- Kline, R. B. (2023) *Principles and practice of structural equation modeling*, 5th edition, New York: The Guilford Press.
- Kwarteng, S., Appiagyei, E., Larbi, E. and Lotey, E. K. (2025) 'Predicting mathematics facilitators' teaching efficacy through their engagement in professional learning communities: A mixed-method approach', *Journal of Mathematics, Science and Technology Education*, Vol. 2, No. 2, pp. 117–128. <https://doi.org/10.12973/jmste.2.2.117>
- Kyeremeh, P., Awuah, F. K. and Dorwu, E. (2023) 'Integration of ethnomathematics in teaching geometry: A systematic review and bibliometric report', *Journal of Urban Mathematics Education*, Vol. 16, No. 2, pp. 68–89. <https://doi.org/10.21423/jume-v16i2a519>
- Kyeremeh, P., Awuah, F. K. and Orey, D. C. (2025) 'Challenges regarding the integration of ethnomathematical perspectives into geometry teaching: The faculty reflection', *Asian Journal for Mathematics Education*, Vol. 4, No. 3, pp. 416–430. <https://doi.org/10.1177/27527263251326330>
- Lachney, M. (2017) 'Culturally responsive computing as brokerage: Toward asset building with education-based social movements', *Learning, Media and Technology*, Vol. 42, No. 4, pp. 420–439. <https://doi.org/10.1080/17439884.2016.1211679>
- Lotey, E. K., Ofosua, B., Bonyah, E. and Boateng, F. O. (2025) 'Exploring LaTeX adoption: Verifying university students' intentions from a professional development perspective', *Social Sciences and Humanities Open*, Vol. 12, p. 102257. <https://doi.org/10.1016/j.ssaho.2025.102257>
- Lotey, E. K., Bonyah, E., Ofosua, B., Prodromou, T. and Adeniji, S. M. (2026a) 'Examining the influence of professional learning community participation on mathematics facilitators' teaching efficacy: A structural equation modeling approach based on Hord's model', *African Journal of Research in Mathematics, Science and Technology Education*, pp. 1–15. <https://doi.org/10.1080/18117295.2026.2678651>
- Lotey, E. K., Boateng, F. O. and Ofosua, B. (2026b) 'AI self-efficacy, motivation, and engagement among STEM learners: Mediating and moderating pathways to perceived mathematics performance in Ghanaian higher education', *Human Behavior and Emerging Technologies*, Vol. 2026, No. 1, p. 8637455. <https://doi.org/10.1155/hbe2/8637455>

- Mania, S. and Alam, S. (2021) 'Teachers' perception toward the use of ethnomathematics approach in teaching math', *International Journal of Education in Mathematics, Science and Technology*, Vol. 9, No. 2, pp. 282–298. <https://doi.org/10.46328/ijemst.1551>
- Owusu-Darko, I., Sabtiwu, R., Doe, F., Owusu-Mintah, B. and Ofosu, E.K. (2023) 'Akan ethnomathematics: Demonstrating its pedagogical action on the teaching and learning of mensuration and geometry', *Journal of Mathematics and Science Teacher*, Vol. 3, No. 2, p. em042. <https://doi.org/10.29333/mathsciteacher/13281>
- Palhares, P. (2012) 'Mathematics education and ethnomathematics: A connection in need of reinforcement', *Journal of Research in Mathematics Education*, Vol. 1, No. 1, pp. 79–92. <https://doi.org/10.4471/redimat.2012.04>
- Papagiannis, G. J., Klees, S. J. and Bickel, R. N. (1982) 'Toward a political economy of educational innovation', *Review of Educational Research*, Vol. 52, No. 2, pp. 245–290. <https://doi.org/10.3102/00346543052002245>
- Payadnya, I. P.A. A., Wulandari, I. G. A. P. A., Puspawati, K. R. and Saelee, S. (2024) 'The significance of ethnomathematics learning: A cross-cultural perspective between Indonesian and Thailand educators', *Journal for Multicultural Education*, Vol. 18, No. 4, pp. 508–522. <https://doi.org/10.1108/JME-05-2024-0049>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. and Podsakoff, N. P. (2003) 'Common method biases in behavioral research: A critical review of the literature and recommended remedies', *Journal of Applied Psychology*, Vol. 88, No. 5, pp. 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Ponticell, J.A. (2003) 'Enhancers and inhibitors of teacher risk taking: A case study', *Peabody Journal of Education*, Vol. 78, No. 3, pp. 5–24. https://doi.org/10.1207/S15327930PJE7803_02
- Rosa, M., D'Ambrosio, U., Orey, D. C., Shirley, L., Alangu, W. V., Palhares, P. and Gavarrete, M. E. (2016) *Current and future perspectives of ethnomathematics as a program*, Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-30120-4>
- Rosa, M. and Orey, D. C. (2016) 'Humanizing mathematics through ethnomodelling', *Journal of Humanistic Mathematics*, Vol. 6, No. 2, pp. 3–22. <https://doi.org/10.5642/jhummath.201602.03>
- Sari, M. H. and Yüce, E. (2020) 'Problems experienced in classrooms with students from different cultures', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 13, No. 2, pp. 90–100. <https://doi.org/10.7160/eriesj.2020.130204>
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J. and Shin, T. S. (2009) 'Technological pedagogical content knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers', *Journal of Research on Technology in Education*, Vol. 42, No. 2, pp. 123–149. <https://doi.org/10.1080/15391523.2009.10782544>
- Serdyukov, P. (2017) 'Innovation in education: What works, what doesn't, and what to do about it?', *Journal of Research in Innovative Teaching and Learning*, Vol. 10, No. 1, pp. 4–33. <https://doi.org/10.1108/JRIT-10-2016-0007>
- Shulman, L. S. (1986) 'Those who understand: Knowledge growth in teaching', *Educational Researcher*, Vol. 15, No. 2, pp. 4–14. <https://doi.org/10.3102/0013189X015002004>
- Sunzuma, G. and Maharaj, A. (2021) 'In-service mathematics teachers' knowledge and awareness of ethnomathematics approaches', *International Journal of Mathematical Education in Science and Technology*, Vol. 52, No. 7, pp. 1063–1078. <https://doi.org/10.1080/0020739X.2020.1736351>
- Sunzuma, G., Zezekwa, N., Gwizangwe, I. and Zinyeka, G. (2021) 'A comparison of the effectiveness of ethnomathematics and traditional lecture approaches in teaching consumer arithmetic: Learners' achievement and teachers' views', *Pedagogical Research*, Vol. 6, No. 4, p. em0103. <https://doi.org/10.29333/pr/11215>
- Tschannen-Moran, M. and Woolfolk Hoy, A. (2001) 'Teacher efficacy: Capturing an elusive construct', *Teaching and Teacher Education*, Vol. 17, No. 7, pp. 783–805. [https://doi.org/10.1016/S0742-051X\(01\)00036-1](https://doi.org/10.1016/S0742-051X(01)00036-1)
- Wallace, C. S. and Priestley, M. (2011) 'Teacher beliefs and the mediation of curriculum innovation in Scotland: A socio-cultural perspective on professional development and change', *Journal of Curriculum Studies*, Vol. 43, No. 3, pp. 357–381. <https://doi.org/10.1080/00220272.2011.563447>
- Wulandari, I.G.A.P.A., Payadnya, I.P.A.A., Puspawati, K.R. and Saelee, S. (2024) 'The role of ethnomathematics in South-East Asian learning: A perspective of Indonesian and Thailand educators', *Mathematics Teaching-Research Journal*, Vol. 16, No. 3, pp. 101–119.

BEYOND A-LEVEL POINTS: EVALUATING THEIR RELEVANCE TO RESPONSIBLE AND EFFICIENT UNIVERSITY ADMISSIONS IN THE DIGITAL ERA

Desmond Mwembe✉
Brighton Mpofu
Nicholas J Sibanda
Aaron Nyoni

National University of Science and
Technology, Zimbabwe

✉ desmond.mwembe@nust.ac.zw

ABSTRACT

In Zimbabwe, admission to university programmes is largely based on the number of points obtained at Advanced Level (A-Level). Because applicants come from schools with unequal access to educational resources, heavy reliance on examination scores may raise concerns about equity. This study examined structural associations among A-Level points, gender, faculty, and final degree classification using records from 1,512 graduates of five Zimbabwean public universities. A log-linear model was used to analyse these associations. The collapsed model [FA, FD] provided an adequate fit, indicating associations between faculty and A-Level points and between faculty and degree class, but no evidence of a direct association between A-Level points and degree class after accounting for faculty. In the four-variable model, gender was associated with faculty but not with A-Level points or degree class. Overall, the findings question the use of A-Level points as a general indicator of subsequent degree classification and suggest that any discipline-specific use of entry scores in admissions should be supported by further evidence.

KEYWORDS

A-Level points, higher education, log-linear modelling, responsible admissions, university admissions, Zimbabwe

HOW TO CITE

Mwembe D., Mpofu B., Sibanda N. J., Nyoni A. (2026) 'Beyond A-Level Points: Evaluating their Relevance to Responsible and Efficient University Admissions in the Digital Era', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 269–280. <http://dx.doi.org/10.7160/eriesj.2026.190307>

Article history

Received

July 16, 2025

Received in revised form

June 1, 2026

Accepted

September 7, 2026

Available on-line

September 30, 2026

Highlights

- Faculty was associated with both A-Level points and degree class among graduates.
- Gender composition differed across faculties, with female graduates more strongly represented in the non-STEM faculties.
- In the final collapsed model, faculty was associated with A-Level points and degree classification, while A-Level points and degree class were conditionally independent given faculty.
- The results do not support using A-Level points as a general predictor of degree classification across faculties.

INTRODUCTION

Admission to many Zimbabwean universities and other tertiary institutions, including colleges and polytechnics, relies heavily on Advanced Level (A-Level) results. Although factors such as experience and maturity may occasionally be considered, examination performance remains a dominant admission criterion. The predictive validity of A-Level points for subsequent university success, however, remains debated. Highly competitive programmes, such as medicine, may require 14 to 15 A-Level points. Some studies have reported positive associations between stronger pre-university performance and university outcomes (Alfan and Othman, 2005; Klüfa, 2015), whereas others have highlighted additional influences,

including socio-economic background and discipline-specific demands (Tomaszewski et al., 2025; Wang et al., 2024; Betts et al., 2008). Tamara et al. (2016), for example, reported that university performance could not be explained solely by students' previous school performance.

Previous studies have found that A-Level performance can predict university success in the United Kingdom (Rowbottom, 2013; Natalia et al., 2024), although the strength of this relationship varies by discipline and may be particularly relevant in Science, Technology, Engineering and Mathematics (STEM) fields. Universities have a responsibility to ensure that selection decisions are fair, ethical, and based on an appropriate assessment of applicants' potential rather than on a single

indicator alone. Admission policies should also identify students who are likely to succeed while supporting retention and completion. Accordingly, this study examines whether A-Level points are sufficiently associated with later university outcomes to justify their prominent role in admission decisions across different academic fields.

A-Level points also have important limitations in an increasingly digitalised educational environment. Examination scores provide a relatively narrow indicator of academic attainment and may not capture broader competencies such as critical thinking, creativity, collaboration, and digital literacy, all of which are increasingly valued in higher education and employment.

Moreover, students from under-resourced schools may have less access to high-quality teaching, digital resources, and academic support, placing them at a potential disadvantage in high-stakes examinations. University admission processes should therefore consider whether reliance on A-Level results adequately reflects applicants' potential and differing educational circumstances. This may require recognising relevant forms of achievement beyond traditional examination performance.

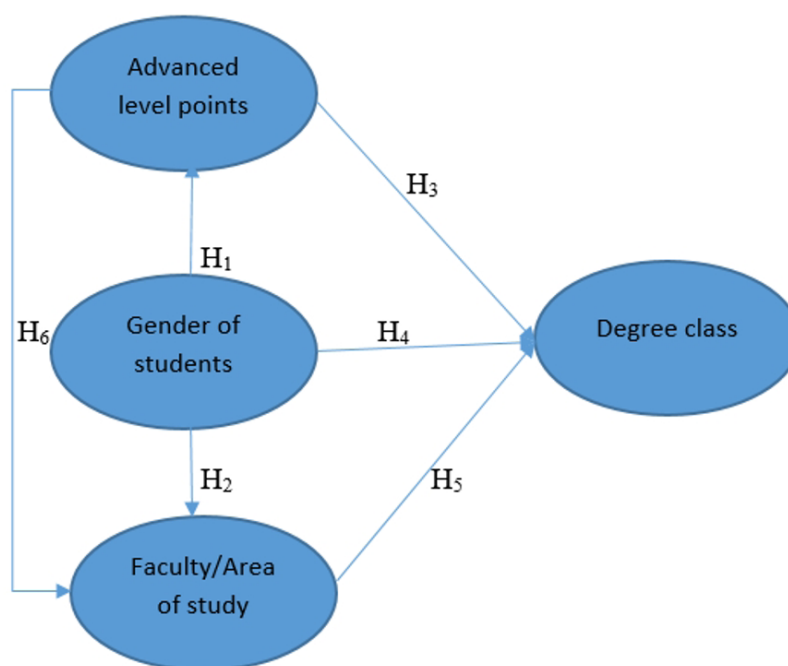
To address these concerns, universities may consider more holistic admission approaches that incorporate relevant contextual information, extracurricular involvement, and non-traditional learning experiences alongside academic results. Over-reliance on A-Level points may fail to capture the full range of applicants' abilities and circumstances, thereby raising concerns about fairness and inclusivity. A broader evidence base may help universities identify students who are

able to adapt and succeed in rapidly changing academic and professional environments.

This study contributes to this discussion by examining the Zimbabwean university context. Admission requirements in Zimbabwean universities are generally applied irrespective of applicants' geographical location or socio-economic circumstances. Such a practice may disadvantage students from rural or under-resourced schools if their A-Level results are systematically lower than those of students from better-resourced settings (Mwembe et al., 2022; Mwembe and Moyo, 2024).

The number of A-Level points is an important factor in determining eligibility for admission to particular programmes. Students with high A-Level scores are generally better positioned to compete for places, provided that their subject combinations meet programme requirements. Subject combinations also constrain the disciplines for which applicants are eligible, including science, arts, business, and engineering-related fields.

The present study therefore evaluates the relevance of A-Level points to later university outcomes, given the assumption that university achievement may be associated with prior academic attainment (Kučera et al., 2015). Specifically, it examines degree-class distributions among graduates classified as low-, medium-, and high-point entrants. Previous research suggests that gender (Ram et al., 2025; Eshetu et al., 2024), area of study (Blatt et al., 2020), and prior academic performance (Kipuru et al., 2024; Merchant et al., 2025) may be associated with tertiary-level outcomes. The conceptual framework motivating the study is presented in Figure 1.



H₀: There is no association between the variables,

H_i: There is an association between the variables, $i=1-6$.

Figure 1: Conceptual framework

The conceptual framework assumes potential associations among gender, A-Level points, choice of study area, and final degree classification. The study therefore assesses the structural relationships among A-Level points, gender, faculty, and degree class among university graduates. The findings are intended to inform discussion of whether admission policies that rely heavily on A-Level points are sufficiently evidence-based, equitable, and appropriate across fields of study.

The relationship between A-Level points and university outcomes, particularly final degree classification, remains uncertain. Although higher A-Level points are often assumed to be associated with stronger performance in STEM fields, this relationship may vary across disciplines. This creates a need to examine whether graduates entering university with low, medium, or high A-Level scores show different degree-class distributions and whether any association remains after accounting for faculty. Such evidence can contribute to the development of more effective and equitable admission strategies and to the identification of areas in which students may require additional academic support.

Literature Review of Empirical Research

Previous research indicates that higher education institutions have increasingly emphasised practical competencies alongside theoretical knowledge. Mahmud and Wong (2022), for example, highlighted the growing importance of twenty-first-century skills among university students. Similarly, Zeng and Shen (2026) proposed a Comprehensive Applicant Profile Score (CAPS) that extends beyond prior academic performance to incorporate extracurricular activities and other dimensions of applicants' profiles.

Ali et al. (2024), Kloos et al. (2025), and Lee (2025) likewise emphasise multidimensional approaches to evaluating applicants in a changing higher education environment. Reliance on A-Level scores alone may not fully reflect a student's potential to succeed at university. Alternative or complementary admission mechanisms may therefore provide a broader assessment of applicant readiness. Klüfa (2015), for instance, compared three approaches to admitting students to a Faculty of Informatics and Statistics, illustrating that admission decisions can be structured in different ways rather than relying on a single academic indicator.

Iannelli et al. (2025), Hinojosa et al. (2025), and Fox et al. (2025) also discuss alternative or combined approaches to university admission. These approaches may include prior academic performance, standardised assessments, contextual information, or other selection criteria. Such evidence supports examining whether reliance on A-Level performance alone is sufficient for responsible and effective admission decisions.

The present study focuses on the role of A-Level points in selection into university programmes. Heavy reliance on examination points may have equity implications because students from under-resourced schools may obtain lower scores despite having the potential to succeed in higher education. Consequently, admission systems that use A-Level points as a primary screening mechanism should be evaluated in relation to both their predictive relevance and their potential distributional consequences.

Universities have a responsibility to serve students from both well-resourced and under-resourced communities. From an efficiency perspective, admission and educational processes should help identify and develop students who can successfully acquire and apply relevant knowledge and skills. Nevertheless, inequalities in access to university programmes between students from differently resourced educational settings remain a concern (Koo et al., 2025).

MATERIALS AND METHODS

A retrospective cross-sectional design was used to analyse administrative records from students who completed undergraduate degrees at five randomly selected public universities in Zimbabwe. The sampling frame comprised of 18, 503 eligible students from which a sample of 1,512 graduates was selected, representing a sampling fraction of 8.2%. Stratified sampling ensured adequate representation across faculties and gender. In each stratum, graduates were assigned unique numbers and sampling was done using computer-generated randomisation. This afforded each graduate an equal chance of being included in the study which subsequently minimises selection bias. The sample size was considered adequate for robust statistical analysis of the relationship between advanced level points and degree classification attained by graduates.

The study included students who successfully graduated during 2021/2022 and 2022/2023 academic years, while students who did not graduate during these academic years were excluded. Graduates were from different universities and different academic programmes. Variables analysed were gender, A-Level points, faculty of enrolment and degree classification. There were no missing data and the study was restricted to graduates who successfully completed their studies. Non-completers, withdrawals, discontinued students and applicants who were not admitted were not represented in the dataset. The findings, therefore, reflect associations among graduates and cannot directly assess admission fairness or the predictive validity of A-Level points for university success in the broader admitted student population. Preliminary analyses were conducted in SPSS version 23. Descriptive statistics were calculated, followed by Pearson chi-square tests to assess bivariate associations among categorical variables. All inferential tests used a significance level of 0.05. Of the 1,512 graduates, 764 (50.5%) were male and 748 (49.5%) were female. Four faculty categories were represented: Applied Science (HAS), Built Environment (HAQS), Business and Economic Sciences (HCOM), and Communication and Information Sciences (HCIS).

Applied Science had the largest number of graduates ($n = 479$, 31.7%), including 225 women and 254 men. Business and Economic Sciences included 455 graduates (30.1%), of whom 236 were women and 219 were men. Built Environment included 299 graduates (19.8%), of whom 123 were women and 176 were men. Communication and Information Sciences had the smallest group ($n = 279$, 18.5%), including 164 women and 115 men.

Built Environment and Applied Science included more male than female graduates, whereas Business and Economic

Sciences and Communication and Information Sciences included more female than male graduates.

A total of 234 graduates (15.5%) received a distinction/first-class degree (above 75%; coded 1), including 110 women and 124 men. A further 701 graduates (46.4%) received an upper-second-class degree (65%–74%; coded 2.1), including 350 women and 351 men; 370

(24.5%) received a lower-second-class degree (60%–64%; coded 2.2), including 186 women and 184 men; and 207 (13.7%) received a pass classification (50%–59%; coded 3), including 102 women and 105 men. Degree-class distributions did not differ significantly by gender ($p = 0.868$) but did differ significantly by faculty ($p = 0.002$). These distributions are summarised in Table 1.

			Degree Class				Total	p-value
			Distinction	2.1	2.2	Pass		
Advanced level Points	Low	Count	70	227	132	76	505	0.202
		% of Total	4.6%	15.0%	8.7%	5.0%	33.4%	
	Medium	Count	84	280	143	69	576	
		% of Total	5.6%	18.5%	9.5%	4.6%	38.1%	
	High	Count	80	194	95	62	431	
		% of Total	5.3%	12.8%	6.3%	4.1%	28.5%	
	Total	Count	234	701	370	207	1512	
		% of Total	15.5%	46.4%	24.5%	13.7%	100.0%	
Faculty	HAS	Count	66	226	119	68	479	0.002
		% of Total	4.4%	14.9%	7.9%	4.5%	31.7%	
	HAQS	Count	44	122	84	49	299	
		% of Total	2.9%	8.1%	5.6%	3.1%	19.8%	
	HCOM	Count	78	241	93	43	455	
		% of Total	5.2%	15.9%	6.2%	2.8%	30.1%	
	HCIS	Count	46	112	74	47	279	
		% of Total	3.0%	7.4%	4.9%	3.1%	18.5%	
	Total	Count	234	701	370	207	1512	
		% of Total	15.5%	46.4%	24.5%	13.7%	100.0%	
Gender	Male	Count	124	351	184	105	764	0.868
		% of Total	8.2%	23.2%	12.2%	6.9%	50.5%	
	Female	Count	110	350	186	102	748	
		% of Total	7.3%	23.1%	12.3%	6.7%	49.5%	
Total		Count	234	701	370	207	1512	
		% of Total	15.5%	46.4%	24.5%	13.7%	100.0%	

Table 1: Distribution of graduates by A-Level points, faculty, gender, and degree class

Log-Linear Modelling

Four categorical variables were available for each graduate: gender, A-Level points, faculty, and degree class (Wrigley-Asante et al., 2023). Because the analysis concerned associations within a multiway contingency table, log-linear modelling was used to describe the structural relationships among these variables (Aravkin et al., 2024). Gender was categorised as male or female. A-Level points were grouped as low (2–7 points), medium (8–11 points), or high (12 points or above). These categories were study-defined analytical groups rather than as official policy classifications.

The possible score ranged from 0 to 15 points from three subjects. However, students obtaining 0 or 1 point would not

normally meet the minimum requirement for admission into university studies. Under the Zimbabwe Schools Examination Council (ZIMSEC) grading system, an A grade contributes 5 points, B contributes 4, C contributes 3, D contributes 2, E contributes 1, and F contributes 0 points (Kanyongo, 2005). It is acknowledged that categorising advanced level points leads to loss of information but categorisation provided a practical way of examining the association between degree class and number of advanced level points graduates had.

The Model

The general form of the log-linear model used to assess associations among the study variables is given in Equation 1:

$$\log \mu_{ijkl} = \tau + \tau_i^A + \tau_j^D + \tau_k^F + \tau_l^G + \tau_{ki}^{FA} + \tau_{kl}^{FG} + \tau_{kj}^{FD} + \tau_{il}^{AG} + \tau_{ij}^{AD} + \tau_{jl}^{GD} + \tau_{kil}^{FAG} + \tau_{kij}^{FAD} + \tau_{klj}^{FGD} + \tau_{ijl}^{AGD} + \tau_{ijkl}^{FAGD} \quad (1)$$

For $i = 1, 2, 3$; $j = 1, 2, 3, 4$; $k = 1, 2, 3, 4$; and $l = 1, 2$, where i denotes A-Level point category (low, medium, high), j denotes

degree-class category (1, 2.1, 2.2, 3), k denotes faculty (HAS, HAQS, HCOM, HCIS), and l denotes gender (male, female).

where

$\log \mu_{ijkl}$ = Log odds

τ = Overall effect,

τ_i^A = Main effect of advanced level points (at low, mid and high levels),

τ_j^D = Main effect of Degree class (that is 1, 2.1, 2.2 and 3),

τ_k^F = Main effect of Faculty (HAS, HAQS, HCOM and HCIS),

τ_l^G = Main effect of Gender (male and female),

τ_{ki}^{FA} = Second order interaction effect of faculty and advanced level points,

τ_{kl}^{FG} = Second order interaction effect of faculty and gender,

τ_{kj}^{FD} = Second order interaction effect of faculty and degree class,

τ_{il}^{AG} = Second order interaction effect of advanced level points and gender,

τ_{ij}^{AD} = Second order interaction effect of advanced level points and degree class,

τ_{lj}^{GD} = Second order interaction effect of gender and degree class,

τ_{kil}^{FAG} = Third order interaction effect of faculty, advanced level points and gender,

τ_{kij}^{FAD} = Third order interaction effect of faculty, advanced level points and degree class,

τ_{klj}^{FGD} = Third order interaction effect of faculty, gender and degree class,

τ_{ilj}^{AGD} = Third order interaction effect of advanced level points, gender and degree class,

τ_{ijkl}^{FAGD} = Fourth order interaction effect of faculty, advanced level points, gender and degree class.

Subsequent analyses also examined reported odds for selected degree-class contrasts by A-Level point category.

RESULTS

Associations between variables

Pearson chi-square tests identified statistically significant associations between faculty and A-Level points (chi-square = 14.069, $p = 0.029$), faculty and gender (chi-square = 20.224, $p < 0.001$), and faculty and degree class (chi-square = 25.491, $p = 0.002$). Thus, the distribution of A-Level point categories differed across faculties, gender composition differed across faculties, and degree-class distributions also differed across faculties. Table 2 summarises the bivariate associations.

			Faculty	Advanced level Points	Gender	Degree Class
Pearson Chi-Square	Faculty	Pearson Chi-square	1.000			
		Sig. (2-tailed)	.			
	Advanced level Points	Pearson Chi-square	14.069*	1.000		
		Sig. (2-tailed)	.029	.		
	Gender	Pearson Chi-square	20.224**	1.009	1.000	
		Sig. (2-tailed)	.000	.604	.	
	Degree Class	Pearson Chi-square	25.491**	8.522	0.734	1.000
		Sig. (2-tailed)	.002	.202	.868	.

Table 2: Chi-square tests for the study variables

Pearson's chi-square test revealed a statistically significant association between faculty and degree classification, $\chi^2(9) = 25.49$, $p = 0.002$. Examination of adjusted standardised residuals indicated that HCOM had significantly more students obtaining upper second-class degree (2.1) than expected and significantly fewer students obtaining a pass as expected. Applying the Holm correction for multiple comparisons for upper-second class and pass categories, the deviations remained statistically significant with $p = 0.012$ and $p = 0.025$, respectively. The remaining faculty-degree-class combinations did not show statistically significant deviations from the expected frequencies after adjustment. The results are presented in Table 3.

Log-Linear Modelling

Structural relationships among the variables were analysed in SPSS version 23.0. A K-way analysis was first used to determine the highest-order effects required in the model. The results indicated that the third-order interactions (F^*A^*G , F^*A^*D , F^*G^*D , and A^*G^*D) and the four-way interaction ($F^*A^*G^*D$) were not statistically significant, whereas lower-order effects required further evaluation. Table 4 summarises the K-way and higher-order tests.

The K-way and higher-order tests indicated that including three-way and four-way interactions did not materially improve model fit ($p = 0.880$ for the three-way effects and $p = 0.976$ for the four-way effect). A more parsimonious model containing

Faculty	Degree class	Adjusted residual	Holm-adjusted <i>p</i> -value	Conclusion
HCOM	Upper-second class (2.1)	3.36	0.012	Above expectation
HCOM	Pass (3)	-3.15	0.025	Below expectation

Table 3: Comparisons of performance among faculties

	<i>K</i>	<i>df</i>	Likelihood Ratio		Pearson	
			Chi-Square	Sig.	Chi-Square	Sig.
K-way and Higher Order Effects ^a	1	95	565.783	.000	680.228	.000
	2	86	112.646	.029	110.234	.040
	3	57	44.875	.878	44.769	.880
	4	18	8.179	.976	8.160	.976
K-way Effects ^b	1	9	453.137	.000	569.994	.000
	2	29	67.771	.000	65.465	.000
	3	39	36.696	.575	36.609	.579
	4	18	8.179	.976	8.160	.976

a. Tests whether k-way and higher-order effects are zero.

b. Tests whether k-way effects are zero.

Table 4: K-way and higher-order effects

lower-order effects was therefore considered appropriate. One-way and two-way effects were statistically significant overall ($p < 0.05$).

Partial association tests were then examined to determine which specific lower-order terms were required. Faculty ($p < 0.001$), A-Level points ($p < 0.001$), and degree class ($p < 0.001$) showed significant main effects, whereas the gender main effect was not significant ($p = 0.823$). Because hierarchical log-linear models retain lower-order terms that are components of

significant higher-order interactions, however, interpretation of individual main-effect tests must be made in conjunction with the interaction structure.

Among the two-way interactions, F*A ($p = 0.023$), F*G ($p = 0.001$), and F*D ($p = 0.002$) were statistically significant, whereas A*G ($p = 0.575$), A*D ($p = 0.284$), and G*D ($p = 0.795$) were not. Gender therefore remained in the hierarchical model because the F*G interaction was significant. Table 5 presents the partial association tests.

Effect	Abbreviations	<i>df</i>	Partial Chi-Square	Sig.
F*A*G	FAG	6	4.649	.590
F*A*D	FAD	18	15.312	.640
F*G*D	FGD	9	13.455	.143
A*G*D	AGD	6	3.025	.806
F*A	FA	6	14.684	.023
F*G	FG	3	17.701	.001
A*G	AG	2	1.105	.575
F*D	FD	9	26.222	.002
A*D	AD	6	7.417	.284
G*D	GD	3	1.028	.795
F	F	3	84.927	.000
A	A	2	16.715	.000
G	G	1	.050	.823
D	D	3	351.444	.000

Table 5: Partial associations among the variables

The resulting hierarchical log-linear model is given in Equation 2:

$$\log \mu_{ijkl} = \tau + \tau_i^A + \tau_j^D + \tau_k^F + \tau_l^G + \tau_{ki}^{FA} + \tau_{kl}^{FG} + \tau_{kj}^{FD} \quad (2)$$

For $i = 1, 2, 3$; $j = 1, 2, 3, 4$; $k = 1, 2, 3, 4$; and $l = 1, 2$.

The model containing the F*A, F*G, and F*D interactions provided an adequate fit. The likelihood-ratio chi-square was 54.201 with 68 degrees of freedom ($p = 0.888$), and the Pearson chi-square was 54.573 ($p = 0.881$). The non-significant

goodness-of-fit tests indicate no evidence of substantial lack of fit for this model.

Results from the Log-Linear Model

The fitted model [FG, FA, FD], summarised in Table 6, indicated that faculty was associated with gender, A-Level points, and degree class. Gender composition therefore differed across faculties, as did the distribution of A-Level point categories and degree classifications. These associations describe the observed distribution of graduates and should not be interpreted as causal effects of faculty.

A-Level points were not associated with gender, and gender was not associated with degree class after accounting for the model structure. Likewise, the A*D interaction was not statistically significant, indicating no

evidence of an association between A-Level point category and degree class in the fitted four-variable model. Thus, the statistically supported two-way associations were F*G, F*A, and F*D.

Steps in Modelling ^a		Deleted Effects	Abbreviations	Chi-Square ^c	df	p-value
0	Generating Class ^b	F*A*G*D	FAGD	.000	0	.000
	Deleted Effect	1 F*A*G*D	FAGD	8.179	18	.976
1	Generating Class ^b	F*A*G, F*A*D, F*G*D, A*G*D	FAG, FAD, FGD, AGD	8.179	18	.976
	Deleted Effect	1 F*A*G	FAG	4.649	6	.590
		2 F*A*D	FAD	15.312	18	.640
		3 F*G*D	FGD	13.455	9	.143
		4 A*G*D	AGD	3.025	6	.806
2	Generating Class ^b	F*A*G, F*A*D, F*G*D	FAG, FAD, FGD	11.203	24	.987
	Deleted Effect	1 F*A*G	FAG	4.049	6	.670
		2 F*A*D	FAD	15.207	18	.648
		3 F*G*D	FGD	13.546	9	.139
3	Generating Class ^b	F*A*D, F*G*D, A*G	FAD, FGD, AG	15.253	30	.988
	Deleted Effect	1 F*A*D	FAD	15.497	18	.628
		2 F*G*D	FGD	14.055	9	.120
		3 A*G	AG	1.045	2	.593
4	Generating Class ^b	F*G*D, A*G, F*A, A*D	FGD, AG, FA, AD	30.750	48	.975
	Deleted Effect	1 F*G*D	FGD	14.125	9	.118
		2 A*G	AG	1.114	2	.573
		3 F*A	FA	14.692	6	.023
		4 A*D	AD	7.425	6	.283
5	Generating Class ^b	F*G*D, F*A, A*D	FGD, FA, AD	31.864	50	.979
	Deleted Effect	1 F*G*D	FGD	14.116	9	.118
		2 F*A	FA	14.787	6	.022
		3 A*D	AD	7.305	6	.294
6	Generating Class ^b	F*G*D, F*A	FGD, FA	39.169	56	.957
	Deleted Effect	1 F*G*D	FGD	14.116	9	.118
		2 F*A	FA	14.809	6	.022
7	Generating Class ^b	F*A, F*G, F*D, G*D	FA, FG, FD, GD	53.285	65	.850
	Deleted Effect	1 F*A	FA	14.809	6	.022
		2 F*G	FG	17.803	3	.000
		3 F*D	FD	26.242	9	.002
		4 G*D	GD	.916	3	.822
8	Generating Class ^b	F*A, F*G, F*D	FA, FG, FD	54.201	68	.888
	Deleted Effect	1 F*A	FA	14.809	6	.022
		2 F*G	FG	17.599	3	.001
		3 F*D	FD	26.037	9	.002
9	Generating Class ^b	F*A, F*G, F*D	FA, FG, FD	54.201	68	.888

Table 6: Initial log-linear model estimates

The Final Model

The four-variable hierarchical log-linear model revealed a statistically significant faculty and gender [FG] interaction which shows that the distribution of male and female students differed significantly across faculties. However, the gender and advanced level points [GA] and gender and degree classification [GD] associations were statistically not significant. One of the researchers' motivation was to examine the relationship between faculty and degree

classification. A marginal analysis collapsing over gender was therefore conducted to assess faculty and degree classification [FD] association across male and female graduates. This approach was considered appropriate due to the fact that there was no evidence of a significant direct gender and degree classification association. It should be noted that collapsing over gender does not mean that gender is globally insignificant. Summary of the marginal model is shown in Figure 2 below.

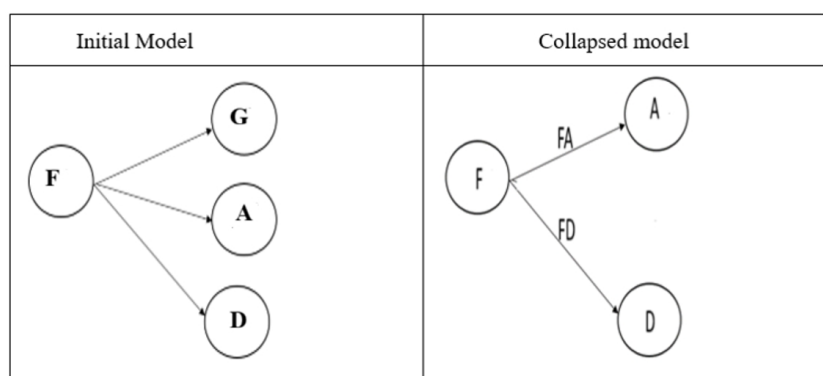


Figure 2: Diagrammatic representation of the log-linear models

The final model after collapsing over gender is given in Equation 3:

$$\log \mu_{ijk} = \tau + \tau_i^A + \tau_j^D + \tau_k^F + \tau_{ki}^{FA} + \tau_{kj}^{FD} . \quad (3)$$

The collapsed model indicated an association between A-Level points and faculty and an association between degree class

and faculty. The absence of an [AD] term implies conditional independence between A-Level points and degree class given faculty. The goodness-of-fit test for the [FA, FD] model was non-significant ($p = 0.527$), indicating that the model provided an adequate representation of the three-way table. Table 7 summarises the model-selection results.

Step ^a		Effects	Abbreviations	Chi-Square ^c	df	Sig.
0	Generating Class ^b	Faculty*A' Points*Degree Class	FAD	.000	0	.000
	Deleted Effect	1 Faculty*A' Points*Degree Class	FAD	15.566	18	.623
1	Generating Class ^b	Faculty*A' Points, Faculty*Degree Class, A' Points*Degree Class	FA FD AD	15.566	18	.623
	Deleted Effect	1 Faculty*A' Points	FA	14.787	6	.022
	Deleted Effect	2 Faculty*Degree Class	FD	26.014	9	.002
	Deleted Effect	3 A' Points*Degree Class	AD	7.305	6	.294
	Generating Class ^b	Faculty*A' Points, Faculty*Degree Class	FA,FD	22.871	24	.527
2	Deleted Effect	1 Faculty*A' Points	FA	14.809	6	.022
	Deleted Effect	2 Faculty*Degree Class	FD	26.037	9	.002
3	Generating Class ^b	Faculty*A' Points, Faculty*Degree Class	FA,FD	22.871	24	.527

a. At each step, the effect with the largest significance level for the likelihood-ratio change is deleted, provided that the significance level exceeds 0.050.

b. Statistics are displayed for the best model at each step after Step 0.

c. For a deleted effect, this value is the change in chi-square after the effect is removed from the model.

Table 7: Results of the final log-linear model

Descriptive Probabilities and Odds of Distinction and Pass by Faculty

The study also descriptively examined selected outcomes for graduates in the low- and high-point categories within each faculty. The medium-point category was not included in these contrasts. These comparisons were based on the observed counts and are presented to illustrate the distributions in the data; they are not formal tests of an A-Level-points-by-degree-class association within faculties. The primary inferential model remained the parsimonious [FA][FD] model, which implies conditional independence between A-Level points and degree class given faculty. The three-way F*A*D interaction was also not statistically significant. $\Delta G^2 = 10.593$. Alternative hierarchical models were examined as a sensitivity

check and are summarised in Table 8. Although models containing an A-Level-points-by-degree-class [AD] term can also show acceptable absolute fit, the [AD] term was not statistically significant and its inclusion was not required to obtain an adequate model. The [FA][FD] model was therefore retained for inferential interpretation on grounds of parsimony. The descriptive faculty-specific probabilities reported below should consequently not be interpreted as evidence that A-Level points predict degree classification within faculty or that faculty modifies such a relationship.

Accordingly, Table 8 is presented as a comparison of alternative model specifications rather than as a basis for replacing the final [FA][FD] model. The inferential conclusions of the study are based on the final parsimonious model reported in Table 7.

Model	G ² (likelihood ratio χ^2)	df	p-value	AIC	Decision
[F][A][D]	72.98	39	< 0.001	337.019	Reject
[FA][FD]	31.45	24	0.141	325.50	Acceptable
[FA][AD]	46.86	27	0.010	334.90	Reject
[FD][AD]	36.19	24	0.053	330.23	Marginal fit
[FA][FD][AD]	20.86	18	0.286	326.91	Acceptable
[FAD]	0	0	-	312.04	Acceptable

Table 8: Comparison of possible models

Faculty of Applied Sciences

For descriptive purposes, attention was given to two particular outcomes:

1. the odds of obtaining a distinction among graduates with low A-Level points, and
2. the odds of obtaining a pass-class degree among graduates with high A-Level points.

As an illustration, among the 170 HAS graduates with low points, 16 attained distinctions. This means:

$$P(\text{Distinction} | \text{Low, HAS}) = \frac{16}{170} = 0.0941.$$

This corresponds to an observed probability of 9.41% and odds of 0.104.

The observed probability of graduating with a pass was also examined. Of the 112 HAS graduates with high A-Level points, 18 obtained a pass-class degree, corresponding to a probability of 16.07% and odds of 0.191.

Odds within the Faculty of the Built Environment

For the Built Environment, the observed odds of obtaining a distinction among graduates with low A-Level points were 0.167, corresponding to the probability shown below:

$$P(\text{Distinction} | \text{Low, HAQS}) = \frac{15}{105} = 0.1429.$$

This corresponds to 14.29%. In the Built Environment faculty, 16 of the 93 graduates with high A-Level points obtained a pass-class degree, corresponding to a probability of 17.20% and odds of 0.208. The same descriptive calculations were applied to the other faculties.

Odds within the Faculty of Business and Economic Sciences

For Business and Economic Sciences, the probability of attaining a distinction given that the graduate had low points was 15.44% and the odds were 0.183. The probability of graduating with a pass-degree classification while with high advanced level points was found to be 9.66% with corresponding odds of 0.107.

Odds within the Faculty of Communication and Information Sciences

For Communication and Information Sciences, the estimated odds were 0.250 and the probability of getting a distinction given that a graduate had low points was 20%. The probability of graduating with a pass-degree classification given that the graduate had high advanced level points was found to be 14.89% with corresponding odds of 0.175. HCIS recorded the highest observed probability of distinction among graduates with low A-Level points (20.00%), followed by HCOM (15.44%), HAQS (14.29%), and HAS (9.41%). HCOM had the lowest observed probability of a pass among graduates with high A-Level points (9.66%), whereas HAQS had the highest (17.20%). These descriptive differences illustrate variation in the observed faculty distributions but do not establish an A-Level-points-by-degree-class association within faculties.

Faculty Patterns

Descriptive probabilities and odds of distinction and pass among low- and high-point graduates were also compared across faculties. Among low-point graduates, the observed odds of obtaining a distinction were 0.104 in Applied Science, 0.167 in the Built Environment, 0.183 in Business and Economic Sciences, and 0.250 in Communication and Information Sciences. These comparisons are summarised in Table 9.

Faculty	P(Distinction Low A-Level points)	Odds of Distinction Low A-Level points	P(Pass High A-Level points)	Odds of Pass High A-Level points
HAS	9.41%	0.104	16.07%	0.191
HAQS	14.29%	0.167	17.20%	0.208
HCOM	15.44%	0.183	9.66%	0.107
HCIS	20.00%	0.250	14.89%	0.175

Table 9: Comparison of odds of distinction and pass among faculties

Taken descriptively, these comparisons show variation in observed degree-class outcomes across faculties. They do not demonstrate that faculty modifies the relationship between A-Level points and degree classification, because the F*A*D interaction was not statistically significant. The relatively high distinction rate among low-point graduates in HCIS and the relatively low pass rate among high-point graduates in HCOM therefore should be interpreted as descriptive patterns rather than evidence of faculty-specific A-Level effects.

DISCUSSION

The analysis found no evidence of associations between A-Level points and gender and gender and degree class in the fitted four-variable model. These findings differ from those of Wrigley-Asante et al. (2023) and Matteucci and Mignani (2021), who reported gender differences in academic performance, particularly in STEM-related contexts. Differences in study populations, outcomes, educational systems, and disciplinary composition may contribute to the contrasting results.

In the present sample, gender was not associated with either A-Level point category or final degree class. This finding is broadly consistent with studies reporting that gender alone does not consistently explain academic performance once other factors are considered. However, direct comparison with Suchithra et al. (2022) should be interpreted cautiously because that study examined prediction of school-level academic performance using machine-learning methods rather than university degree classification.

At the faculty level, faculty was associated with A-Level points, gender, and degree class. These findings indicate that the composition of graduates differed across faculties with respect to prior A-Level attainment, gender, and final degree classification. Because the study is observational and includes graduates only, these associations should not be interpreted as evidence that faculty itself caused the observed differences.

The association between gender and faculty is consistent with literature documenting gender differences in participation across STEM and non-STEM fields. Alam (2022) and Tandrayen-Ragoobur and Gokulsing (2022), for example, discuss lower female participation in some STEM pathways, while De las Cuevas et al. (2022) examined gender-related factors associated with higher-education choices. The present findings similarly show that gender composition differed across the four faculty categories.

Faculty was also associated with degree classification, indicating that degree-class distributions differed across faculties. Such differences may reflect multiple factors, including programme characteristics, assessment practices, class size, student composition, and available resources, none of which were directly measured in the present study. Consequently, the observed faculty differences should not be attributed to a single mechanism. Related research has likewise documented variation in student outcomes across disciplinary contexts (Costello et al., 2023; Hatfield et al., 2022).

After collapsing over gender, the final model retained associations between faculty and A-Level points and between faculty and degree class. Thus, A-Level point distributions differed across faculties, as did degree-class distributions. However, the final model did not retain an A-Level-points-by-degree-class interaction, indicating conditional independence between these variables given faculty. This result suggests that differences in prior A-Level attainment and degree classification are structured primarily by faculty in the fitted model rather than by a direct association between A-Level points and degree class.

The observed differences in degree-class distributions across faculties are consistent with previous evidence that academic outcomes vary across fields of study (Otavová and Sýkorová, 2016). Potential explanations include differences in class size, teaching and assessment practices, staffing, and resource availability (Kara et al., 2021; Kaspříková, 2012), but these factors were not measured

in the present study and therefore remain hypotheses for future research. Similarly, although gender differences in academic performance have been reported elsewhere (Mensah et al., 2024; Barra et al., 2024), gender was not associated with degree class in this dataset.

The observed gender composition suggests that efforts to support female participation in STEM-related programmes may remain relevant. Universities may therefore consider admission and outreach strategies that promote equitable access across fields while avoiding assumptions about academic potential based on gender. More generally, the findings do not support using A-Level points as a universal stand-alone indicator of degree classification. Admission decisions should therefore be based on evidence appropriate to the programme and should consider fairness, transparency, and the broader context in which prior attainment was achieved.

CONCLUSION

This study analysed records from 1,512 graduates to examine structural associations among gender, A-Level points, faculty, and final degree classification. The four-variable model retained the F*G, F*A, and F*D interactions. After collapsing over gender, the final three-variable model was [FA, FD], indicating associations between faculty and A-Level points and between faculty and degree class. The absence of an A-Level-points-by-degree-class [AD] interaction indicates conditional independence between A-Level point category and degree class given faculty. Accordingly, the results do not provide evidence that A-Level point category generally predicts final degree classification once faculty is taken into account.

Several limitations should be considered. First, the analysis included graduates only; it did not include applicants, rejected applicants, admitted students who withdrew, or students who had not yet graduated. The findings therefore describe associations among graduates and cannot, on their own, establish the predictive validity or fairness of admission decisions. Second, school resourcing, rural or urban background, socio-economic status, teaching methods, class size, and other potentially relevant contextual variables were not available, so their effects could not be tested directly. Third, the data covered only two academic years, limiting temporal generalisability and the ability to assess changes in admission policies or student cohorts over time. Future studies should include broader cohorts and contextual variables to evaluate whether admission criteria operate differently across educational and socio-economic settings.

ACKNOWLEDGEMENT

The authors acknowledge the support of university admissions staff and the permission granted by the NUST Institutional Review Board (IRB; NUST/IRB/2024/186) to access and analyse the data used in this study.

- Alam, A. (2022) 'Psychological, sociocultural, and biological elucidations for gender gap in STEM education: A call for translation of research into evidence-based interventions', in: Samanta, S. R. and Gochhayat, J. (eds.), *Proceedings of the 2nd International Conference on Sustainability and Equity (ICSE-2021)*, Atlantis Highlights in Social Sciences, Education and Humanities, Vol. 2, Dordrecht: Atlantis Press, pp. 95–107. <https://doi.org/10.2991/ahssch.k.220105.012>
- Ali, M., Raza, S. A., Puah, C. H. and Qazi, W. (2024) 'Micro-credential as a digital enabler for higher education ecosystems', *International Journal of Educational Management*, Vol. 38, No. 4, pp. 893–896. <https://doi.org/10.1108/IJEM-05-2024-655>
- Alfan, E. and Othman, N. (2005) 'Undergraduate students' performance: the case of University of Malaya', *Quality Assurance in Education*, Vol. 13, No. 4, pp. 329–343. <https://doi.org/10.1108/09684880510626593>
- Aravkin, A., Choromanska, A., Deng, L., Heigold, G., Jebara, T., Kanevsky, D. and Wright, S. J. (eds.) (2024) *Log-Linear Models, Extensions, and Applications*, Cambridge, MA: The MIT Press.
- Barra, C., Destefanis, S., Sena, V. and Zotti, R. (2024) 'Students' performance and faculty efficiency. Assessing the role of gender through a metafrontier', *Journal of Economic Studies*, Vol. 51, No. 9, pp. 323–337. <https://doi.org/10.1108/JES-01-2024-0037>
- Bautista, R., Morada, M. H. and Pineda, A. L. (2021) 'A Comparative Study of the Academic Performance of STEM and Non-STEM Graduates in College and Advanced Algebra', *Southeast Asian Journal of Science and Technology*, Vol. 6, No. 2 (SI), pp. 114–122. Available at: <https://sajournal.pti.edu.ph/index.php/sajst/article/view/257>
- Betts, L. R., Elder, T. J., Hartley, J. and Blurton, A. (2008) 'Predicting university performance in Psychology: The role of previous performance and discipline specific knowledge', *Educational Studies*, Vol. 34, No. 5, pp. 543–556. <https://dx.doi.org/10.1080/03055690802288528>
- Blatt, L., Schunn, C. D., Votruba-Drzal, E. and Rottman, B. M. (2020) 'Variation in which key motivational and academic resources relate to academic performance disparities across introductory college courses', *International Journal of STEM Education*, Vol. 7, No. 1, p. 58. <https://doi.org/10.1186/s40594-020-00253-0>
- Boháčková, I. and Brožová, I. (2012) 'Assessing Students' Level of Knowledge in Profile Courses of Chosen Economic Study Programmes at the FEM CULS', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 5, No. 3, pp. 148–156. <https://dx.doi.org/10.7160/eriesj.2012.050304>
- Costello, R. A., Salehi, S., Ballen, C. J. and Burkholder, E. (2023) 'Pathways of opportunity in STEM: comparative investigation of degree attainment across different demographic groups at a large research institution', *International Journal of STEM Education*, Vol. 10, No. 1, p. 46. <https://doi.org/10.1186/s40594-023-00436-5>
- De las Cuevas, P., García-Arenas, M. and Rico, N. (2022) 'Why not STEM? A study case on the influence of gender factors on students' higher education choice', *Mathematics*, Vol. 10, No. 2, p. 239. <https://doi.org/10.3390/math10020239>
- Eshetu, A.A., Alemu, S., Mulat, A., Abdu, A., Gobezie, B., Muluken, E., Anagaw, A., Mulugeta, E., Hailemariam, E. and Aragaw, S. (2024) 'Gender disparity in academic performance in higher education institutions: a case of Wollo University, Ethiopia', *Frontiers in Education*, Vol. 9, p. 1476112. <https://dx.doi.org/10.3389/educ.2024.1476112>
- Fox, J., Beattie, J., McGrail, J., Eley, D., Fuller, L., Hu, W., Keniry, C., Chennakesavan, S.K., Parker-Newlyn, L., Reid, K. and Walters, L. (2025) 'Medical school admission processes to target rural applicants: an international scoping review and mapping of Australian practices', *BMC Medical Education*, Vol. 25, No. 1, p. 659. <https://doi.org/10.1186/s12909-025-07234-3>
- Hatfield, N., Brown, N. and Topaz, C. M. (2022) 'Do introductory courses disproportionately drive minoritized students out of STEM pathways?', *PNAS nexus*, Vol. 1, No. 4, p. pgac167. <https://doi.org/10.1093/pnasnexus/pgac167>
- Hinojosa, M., Alfaro, M., Fuertes, G., Ternero, R., Santander, P. and Vargas, M. (2025) 'Optimizing University Admission Processes for Improved Educational Administration Through Feature Selection Algorithms: A Case Study in Engineering Education', *Education Sciences*, Vol. 15, No. 3, p. 326. <https://doi.org/10.3390/educsci15030326>
- Iannelli, C., McMullin, P. and Smyth, E. (2025) 'Higher education retention in Ireland and Scotland: The role of admissions policies', *Higher Education*, Vol. 89, No. 4, pp. 1021–1039. <https://doi.org/10.1007/s10734-024-01259-1>
- Ismail, N. A. and Othman, A. (2006) 'Comparing University Academic Performances of HSC Students at the Three Art-Based Faculties', *International Education Journal*, Vol. 7, No. 5, pp. 668–675. Available at: https://knova.um.edu.my/cgi/viewcontent.cgi?article=1196&context=research_publications_2006_2010
- Kanyongo, G. Y. (2005) 'Zimbabwe's Public Education System Reforms: Successes and Challenges', *International Education Journal*, Vol. 6, No. 1, pp. 65–74. Available at: <https://www.learntechlib.org/p/103477/>
- Kara, E., Tonin, M. and Vlassopoulos, M. (2021) 'Class size effects in higher education: Differences across STEM and non-STEM fields', *Economics of Education Review*, Vol. 82, p. 102104. <https://doi.org/10.1016/j.econedurev.2021.102104>
- Kaspříková, N. (2012) 'Data analysis of students' performance', in: Kvasnička, R. (ed.), *Proceedings of the 9th International Conference on Efficiency and Responsibility in Education (ERIE 2012)*, Prague: Czech University of Life Sciences Prague, pp. 213–218.
- Kipuru, M. M., Kibanja, G. M. and Matagi, L. (2024) 'From A-level success to college achievement: the mediating effect of academic engagement on cumulative grade point average (CGPA) and persistence among undergraduate students at Makerere University', *African Journal of Empirical Research*, Vol. 5, No. 4, pp. 1779–1797. <https://doi.org/10.51867/ajernet.5.4.149>
- Kloos, C. D., Alario-Hoyos, C., Kemcha, R., Moreno-Marcos, P. M., Estévez-Ayres, I., Callejo-Pinardo, P., Muñoz-Merino, P. J., Ibáñez, M.-B. and Muñoz-Organero, M. (2025) 'How challenges become opportunities: Micro-credentials and artificial intelligence', in: *Proceedings of the 2025 IEEE Global Engineering Education Conference (EDUCON)*, London, United Kingdom, 22–25 April 2025, pp. 1–10. <https://doi.org/10.1109/EDUCON62633.2025.11016509>
- Klůfa J. (2015) 'Comparison of The Ways of Acceptance Students at University', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 8, No. 3, pp. 72–76. <https://doi.org/10.7160/eriesj.2015.080304>
- Koo, A., Wong, Y.L. and Cheung, K. (2025) 'Access to top universities through alternative pathways: experiences of transfer students in Hong Kong', *Educational Review*, pp. 1–17, <https://doi.org/10.1080/00131911.2025.2551021>

- Kučera, P., Svatošová, L. and Pelikán, M. (2015) 'University study results as related to the admission exam results', in: Flégl, M. and Houška, M. (eds.), *Proceedings of the 12th International Conference on Efficiency and Responsibility in Education (ERIE 2015)*, Prague: Czech University of Life Sciences Prague, pp. 318–324.
- Lee, J. A. (2025) 'Adaptive merit framework: Merit-anchored fairness via SES correction', *arXiv*, arXiv:2512.13698 [Preprint]. <https://doi.org/10.2139/ssrn.5872762>
- Mahmud, M. M. and Wong, S. F. (2022) 'Digital age: The importance of 21st century skills among the undergraduates', *Frontiers in Education*, Vol. 7, p. 950553. <https://dx.doi.org/10.3389/feduc.2022.950553>
- Matete, R. E. and Kombe, G. G. (2025) 'Gender parity trends in STEM and non-STEM fields in Higher Education Institutions in Tanzania: A comparative analysis', *International Journal of Educational Development*, Vol. 114, p. 103233. <https://doi.org/10.1016/j.ijeducdev.2025.103233>
- Matteucci, M. and Mignani, S. (2021) 'Investigating gender differences in mathematics by performance levels in the Italian school system', *Studies in Educational Evaluation*, Vol. 70, p. 101022. <https://doi.org/10.1016/j.stueduc.2021.101022>
- Mensah, R. O., Amponsah, K. D., Adiza Babah, P. and Jibril, H. S. (2024) 'Factors affecting students' academic performance and teachers' efficiency in Ghana; a case study of Wa senior high school', *Cogent Arts & Humanities*, Vol. 11, No. 1. <https://doi.org/10.1080/23311983.2024.2412944>
- Merchant, A. A. H., Afzal, N., Rahim, K.A., Shah, S. A., Jamal, W. Z., Rahim, A., Bakhshi, S. K., Shaikh, N. Q., Noorali, A.A., Mahmood, S. B. Z. and Akbarali, L. (2025) 'Application to achievement: association between pre-admission factors, admission scores, and medical students' performance', *BMC Medical Education*, Vol. 25, No. 1, p. 223. <https://doi.org/10.1186/s12909-025-06800-z>
- Mwembe, D., Zulu, I. L. and Mdlongwa, P. (2022) 'Needs Assessment for Facilitators in Teaching of STEM Subjects in UnderResourced Rural Schools: Case of Lupane and Hwange Districts in Zimbabwe', *International Journal of Professional Development, Learners and Learning*, Vol. 4, No. 1, p. ep2202. <https://doi.org/10.30935/ijpdll/11666>
- Newsome, M. L., Pina, A. A., Mollazehi, M., Al-Ali, K. and Al-Shaboul, Y. (2022) 'The Effect of Gender and STEM/ Non-STEM Disciplines on Remote Learning: A National Study of Undergraduates in Qatar', *Electronic Journal of e-Learning*, Vol. 20, No. 4, pp. 360–373. Available at: <https://files.eric.ed.gov/fulltext/EJ1352417.pdf>
- Otavová M., and Sýkorová I. (2016) 'Differences in Results Obtained by Students of Different Faculties', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 9, No. 1, pp. 1–6. <https://doi.org/10.7160/eriesj.2016.090101>
- Petersen, N., Ruegg, R. and Hoang, H. (2024) 'Admissions pathways and international undergraduates' academic achievement', *Educational Research*, Vol. 66, No. 4, pp. 413–428. <https://doi.org/10.1080/00131881.2024.2409647>
- Owston, R., York, D. N., Malhotra, T. and Sitthiworachart, J. (2020) 'Blended learning in STEM and non-STEM courses: How do student performance and perceptions compare?', *Online Learning*, Vol. 24, No. 3, pp. 203–221. <https://doi.org/10.24059/olj.v24i3.2151>
- Rajendran, S., Chamundeswari, S. and Sinha, A. A. (2022) 'Predicting the academic performance of middle- and high-school students using machine learning algorithms', *Social Sciences & Humanities Open*, Vol. 6, No. 1, p. 100357. <https://doi.org/10.1016/j.ssaho.2022.100357>
- Ram, I., Shwartz, A. and Roll, I. (2025) 'Minding the gender gap in advanced STEM courses: effects of student preparedness and activity level', *Instructional Science*, Vol. 53, No. 6, pp.1457–1478. <https://doi.org/10.1007/s11251-025-09748-1>
- Rowbottom, N. (2013) 'A-level subject choice, systematic bias and university performance in the UK: The case of accounting', *Accounting Education*, Vol. 22, No. 3, pp. 248–267. <http://dx.doi.org/10.1080/09639284.2013.799840>
- Thiele, T., Singleton, A., Pope, D. and Stanistreet, D. (2014) 'Predicting students' academic performance based on school and sociodemographic characteristics', *Studies in Higher Education*, Vol. 41, No. 8, pp. 1424–1446. <https://dx.doi.org/10.1080/03075079.2014.974528>
- Tandrayen-Ragoobur, V. and Gokulsing, D. (2022) 'Gender gap in STEM education and career choices: what matters? ', *Journal of Applied Research in Higher Education*, Vol. 14, No. 3, pp. 1021–1040. <https://doi.org/10.1108/JARHE-09-2019-0235>
- Tomaszewski, W., Xiang, N. and Kubler, M. (2025) 'Socio-economic status, school performance, and university participation: evidence from linked administrative and survey data from Australia', *Higher education*, Vol. 89, No. 3, pp. 753–774. <https://doi.org/10.1007/s10734-024-01245-7>
- Wang, W., Han, L., Lu, Q., Lv, X., Liu, Y. and Wang, D. (2024) 'Research on the impact of the socio-educational environment on the academic performance of college students: the mediating role of study motivation', *Frontiers in Psychology*, Vol. 14, p. 1289064. <https://doi.org/10.3389/fpsyg.2023.1289064>
- Wrigley-Asante, C., Ackah, C.G. and Frimpong, L.K (2023) 'Gender differences in academic performance of students studying Science Technology Engineering and Mathematics (STEM) subjects at the University of Ghana', *SN Social Sciences*, Vol. 3, No. 12. <https://doi.org/10.1007/s43545-023-00608-8>
- Zeng, J. W. and Shen, J. (2026) 'Quantifying holistic review: A multimodal approach to college admissions prediction', in: Tiwari, S. and Shi, Z. (eds.), *International Conference on Image, Signal Processing, and Machine Learning (ISPML 2025)*, Proceedings of SPIE, Vol. 14058, Bellingham, WA: SPIE, pp. 99–113.

ASSESSMENT PRACTICES AND CHALLENGES OF SOCIAL STUDIES: TEACHERS DURING HEAT-INDEX CLASS SUSPENSIONS

Czar Justine D. Navalta

Nueva Vizcaya State University,
Philippines

✉ czadnavalta@nvsu.edu.ph

ABSTRACT

Heat-index-related class suspensions can disrupt instructional delivery and classroom assessment. Using an explanatory sequential mixed-methods design, this study examined assessment practices and operational challenges among 129 Junior High School Social Studies teachers in Nueva Vizcaya during Alternative Delivery Mode transitions. Quantitative data from an adapted 32-item survey were analysed using descriptive statistics, one-way ANOVA, Tukey post hoc comparisons, and repeated-measures ANOVA. Qualitative data from semi-structured questionnaires and follow-up interviews were analysed thematically. Written Works Assessment ($M = 3.40$, $SD = 0.41$) and Performance Tasks Assessment ($M = 3.49$, $SD = 0.47$) were practised to a great extent, while Reflective and Compassionate Assessment was practised to a very great extent ($M = 3.56$, $SD = 0.38$). Written Works scores differed significantly across age and years-in-service groups, while Reflective and Compassionate Assessment differed across years-in-service groups. Repeated-measures ANOVA indicated overall differences among the three assessment domains. Qualitative findings identified submission delays, hardcopy retrieval barriers, concerns about output authenticity, and logistical burdens related to remedial work. The findings support context-sensitive professional development and assessment-continuity planning for heat-related school disruptions.

KEYWORDS

Alternative delivery modes, classroom assessment, climate disruptions, educational vulnerability, social studies education

HOW TO CITE

Navalta C. J. D. (2026) 'Assessment Practices and Challenges of Social Studies: Teachers during Heat-index Class Suspensions', *Journal on Efficiency and Responsibility in Education and Science*, vol. 19, no. 3, pp. 281–290. <http://dx.doi.org/10.7160/eriesj.2026.190308>

Article history

Received

July 21, 2024

Received in revised form

June 8, 2025

Accepted

September 9, 2026

Available on-line

September 30, 2026

Highlights

- Heat-index-related class suspensions can affect teachers' assessment practices and the continuity of classroom assessment.
- Written Works practices differed by age and years in service, while Reflective and Compassionate Assessment also differed by years in service.
- Teachers reported the highest implementation for Reflective and Compassionate Assessment, followed by Performance Tasks and Written Works.

INTRODUCTION

Environmental disruptions can interrupt school operations and place substantial pressure on educational continuity. Elevated heat conditions are increasingly relevant to learning and school functioning, while emergency closures may require rapid transitions to Alternative Delivery Modes (ADMs) or other forms of home-based learning (Park et al., 2020; UNICEF-EAPRO, 2020; Cahapay, 2020b). Sustainable Development Goal 4 emphasises inclusive and equitable quality education and continuity of learning. In this context, heat-index-related class suspensions create practical challenges for maintaining curriculum-aligned assessment when face-to-face instruction is interrupted.

Assessment is a core component of curriculum implementation because it provides evidence about student learning, instructional decisions, and progress (Mussawy, 2009; Mikre, 2011; Walde, 2015; Rundle, 2016; Umugiraneza et al., 2017; Cagasan et al., 2020). When schools shift temporarily to non-face-to-face delivery, teachers must adapt assessment procedures to the constraints of remote and modular learning. Philippine continuity policies developed for emergency schooling provide relevant guidance on flexible assessment and alternative delivery arrangements (DepEd, 2020a; DepEd, 2020b; Sarmiento, 2020; Reyes et al., 2021).

In Social Studies (Araling Panlipunan), these disruptions create distinct pedagogical challenges. Unlike subjects that rely heavily

on procedural practice, Social Studies curricula emphasise civic competence, critical inquiry, historical contextualisation, and collaborative discourse (NCSS, 2015; Torney-Purta et al., 2015; Tantengco and Maramag, 2016; Reyes, 2018; Rogayan and Villanueva, 2019; Crumb, 2021; Nunvarova, 2023). Assessing higher-order cognitive and affective outcomes becomes more difficult when instruction shifts abruptly to self-learning modules (SLMs) or asynchronous tasks (Adam et al., 2018).

Theoretical Framework and Domain Assessment Model

This study draws on Bronfenbrenner’s ecological-systems perspective and on literature concerning educational vulnerability under environmental disruption. Heat-index-related class suspensions interrupt the school microsystem and increase reliance on home-based learning and home-school interaction. These changes can expose differences in access to resources, supervision, communication, and learning conditions (Park et al., 2020). To examine how educators navigate these disruptions, the study organises classroom assessment into three analytical domains. Written Works and Performance Tasks draw on assessment categories used in DepEd Order No. 31, s. 2020, while Reflective and Compassionate Assessment captures teacher adjustments intended to accommodate socio-emotional strain and environmental constraints during emergency transitions:

- 1. Written Works Assessment (Domain 1):** Evaluates cognitive recall, conceptual comprehension, and factual synthesis through structured instruments such as unit tests, quizzes, and short-answer activity sheets (Boonying & Chatruprachewin, 2017).
- 2. Performance Tasks Assessment (Domain 2):** Evaluates the practical application of knowledge, civic competence, critical thinking, and creative problem-solving through authentic products, portfolios, and contextualized projects (NCSS, 2015; Reyes, 2018).
- 3. Reflective and Compassionate Assessment (Domain 3):** Evaluates affective engagement, student self-reflection, and instructional adjustments made by educators to accommodate socio-emotional strain and environmental hardship during emergency transitions (Soliven, 2020; Tan-Choi et al., 2020).

Research Objectives

The primary aim of this study was to examine the implementation of these three assessment domains among Social Studies teachers during heat-index-related class suspensions, compare reported practices across demographic groups, and identify the operational challenges teachers encountered. Specifically, the study aimed to:

1. measure and compare reported implementation across the three assessment domains: Written Works, Performance Tasks, and Reflective and Compassionate Assessment.
2. determine whether teachers’ assessment-practice scores differ significantly across age and years-in-service groups.
3. evaluate whether reported implementation differs significantly across the three assessment domains within the same teachers.
4. identify qualitative operational challenges encountered during assessment under ADMs and describe strategies teachers used to maintain assessment continuity and integrity.

METHODOLOGY

Research Design

This study employed an explanatory sequential mixed-methods design. The quantitative phase measured assessment practices across demographic categories and assessment domains. The subsequent qualitative phase was used to contextualise and explain the quantitative patterns by examining operational challenges teachers encountered during heat-index-related transitions.

Participants and Sampling Technique

The target population comprised Junior High School (Grades 7–10) Social Studies teachers in the Division of Nueva Vizcaya. Stratified random sampling was used. The final quantitative sample consisted of $N = 129$ full-time public-school teachers who had implemented assessment adaptations during the Quarter III heat-index-related class suspensions. The demographic distribution is summarised in Table 1.

Age Group	F	%
24 years and below	10	7.8
25 – 30	32	24.8
31 – 35	18	14.0
36 – 40	18	14.0
41 – 45	13	10.1
46 – 50	7	5.4
51 – 55	22	17.1
56 and above	9	7.0
Years in the Service		
0 – 3	35	27.1
4 – 10	40	31.0
11 and above	54	41.9
Total: 129		

Table 1: Profile characteristics of the respondents

Research Instrumentation and Distribution

The quantitative instrument was adapted from the Classroom Assessment Practices Survey (Gonzales and Callueng, 2014) and aligned with DepEd Order No. 31, s. 2020 and the Southeast Asia Teachers Competency Framework (SEAMEO, 2018). The survey comprised 32 items rated on a four-point Likert-type scale (1 = Practiced Rarely, 2 = Practiced Sometimes, 3 = Practiced Very Often, 4 = Practiced At All Times).

The instrument comprised three assessment domains:

- **Domain A: Written Works Assessment** (12 items): Measures alignment with learning competencies, table of specifications (TOS) usage, item analysis, unit test formatting, and self-assessment integration. *Sample Item*: “I use table of specifications as I conduct long/unit tests for me to identify the least mastered competencies.”
- **Domain B: Performance Tasks Assessment** (11 items): Measures task contextualization, civic competence integration, rubric clarity, multi-step authentic performance design, and student portfolios. *Sample Item*: “I select performance tasks that hones learners’ civic competence and consciousness.”
- **Domain C: Reflective and Compassionate Assessment** (9 items): Measures accommodation of environmental constraints, flexible submission timelines, formative feedback loops, and alternative remedial tasks. *Sample Item*: “I set realistic expectations and professional judgment to find what is feasible performance tasks to accomplish remotely.”

Psychometric Properties

The overall survey demonstrated high internal consistency, with Cronbach’s alpha of $\alpha = .954$. Domain-specific internal consistency coefficients were also calculated:

- Domain A (Written Works): $\alpha = .882$
- Domain B (Performance Tasks): $\alpha = .914$
- Domain C (Reflective/Compassionate): $\alpha = .897$

Content relevance was reviewed by a panel of three assessment specialists and master teachers in Social Studies Education, with a reported Content Validity Index (CVI) of 0.92.

Data Collection and Distribution Protocol

Because data collection occurred while elevated heat indices were disrupting normal school operations, a two-phase distribution strategy was used to support response completion:

- **Phase 1 (Digital Deployment)**: During the suspension of face-to-face classes, the instrument was administered through Google Forms. The survey link was distributed through official DepEd workplace communication channels and verified email directories to teachers included in the stratified sample.
- **Phase 2 (On-Site Follow-Up)**: After local government units and school heads lifted the heat-index-related suspensions, targeted in-person follow-up visits were conducted to address non-response, verify submissions, clarify technical questions, and complete data collection for the final sample of 129 respondents.

For the qualitative strand, a semi-structured interview guide was developed to capture operational challenges teachers encountered when assessing critical thinking and civic competencies through ADMs.

Data Analysis

Quantitative Data Analysis

Data were processed using SPSS version 28. Statistical procedures were matched to the research objectives:

- **Descriptive statistics (mean and standard deviation)**: Summarised reported implementation within each domain. Means were interpreted as: 1.00–1.49 (Practiced to a Little Extent), 1.50–2.49 (Practiced to a Moderate Extent), 2.50–3.49 (Practiced to a Great Extent), and 3.50–4.00 (Practiced to a Very Great Extent).
- **One-way analysis of variance (ANOVA)**: Tested differences in assessment-practice scores across age and years-in-service groups.
- **Tukey’s HSD post hoc test**: Examined pairwise differences between demographic groups when the omnibus ANOVA F-test was statistically significant ($\alpha = .05$).
- **Repeated-measures ANOVA**: Tested within-subject differences across the three assessment-domain scores. Mauchly’s test of sphericity was evaluated; when sphericity was violated, the Greenhouse-Geisser correction was applied to the degrees of freedom.

Qualitative Methodological Framework

Qualitative data were collected through semi-structured open-ended questionnaires and follow-up interviews with a subset of respondents ($n = 20$) selected through maximum-variation sampling across years-in-service groups.

Qualitative data were analysed using a six-phase thematic-analysis procedure informed by Braun and Clarke (2022):

1. **Familiarisation**: Repeated reading of the narrative data.
2. **Initial code generation**: Line-by-line inductive coding in NVivo to identify operational issues (e.g., “delayed hardcopy delivery” and “parental intervention in output”).
3. **Searching for themes**: Grouping related codes into candidate themes concerning assessment barriers and adaptations.
4. **Reviewing themes**: Checking candidate themes against the full qualitative dataset for coherence and distinctiveness.
5. **Defining and naming themes**: Establishing explicit definitions for the final themes and subthemes.
6. **Producing the report**: Selecting illustrative verbatim excerpts and integrating qualitative patterns with the quantitative findings.

Methodological rigour and trustworthiness were supported through member-checking, investigator triangulation, thick description of the emergency ADM context, an audit trail of coding decisions, and reflexive memoing. Mixed-methods integration occurred during interpretation, when qualitative themes were used to contextualise the quantitative patterns.

RESULTS

Section 1. Level of assessment practices among Junior High School Social Studies teachers

Table 2 presents the mean ratings and standard deviations for the three assessment domains. Reflective and

Compassionate Assessment had the highest mean ($M = 3.56$, $SD = 0.38$, Very Great Extent), followed by Performance Tasks Assessment ($M = 3.49$, $SD = 0.48$, Great Extent). Written Works Assessment had the lowest mean ($M = 3.40$, $SD = 0.41$, Great Extent).

Practices on Assessment	Mean	SD	Qualitative Description
Written Works Assessment	3.40	.41	Practiced at a great extent
Performance Tasks Assessment	3.49	.47	Practiced to a great extent
Reflective and Compassionate Assessment	3.56	.38	Practiced to a very great extent

Legend: 1.00-1.49: Practiced at a little extent; 1.50-2.49: Practiced at a moderate extent; 2.50 – 3.49: Practiced at a great extent; 3.50-4.00: Practiced at a very great extent

Table 2: Level of assessment practices of social studies junior high school teachers

Section 2. Comparison of Junior High School Social Studies teachers' assessment practices by profile characteristics

Table 3 summarises one-way ANOVA comparisons of assessment-practice scores across age groups. A statistically significant difference was observed for Written Works Assessment ($F(7,121) = 2.662$, $p = .013$). No statistically significant differences were

observed for Performance Tasks ($F(7,121) = 1.601$, $p = .141$) or Reflective and Compassionate Assessment ($F(7,121) = 1.964$, $p = .065$).

Tukey's HSD post hoc analysis for Written Works Assessment indicated that teachers aged ≥ 56 years ($M = 3.74$, $SD = 0.25$) scored significantly higher than teachers aged ≤ 24 years ($M = 3.30$, $SD = 0.34$, $p = .021$) and those aged 31–35 years ($M = 3.11$, $SD = 0.38$, $p = .004$).

Age Group	N	Mean	SD	F	Sig.
A. Written works assessment					
24 years and below	10	3.30	.34	2.662	.013*
25 – 30	32	3.48	.38		
31 – 35	18	3.11	.38		
36 – 40	18	3.42	.37		
41 – 45	13	3.42	.39		
46 – 50	7	3.40	.53		
51 – 55	22	3.43	.45		
56 and above	9	3.74	.25		
Total	129	3.40	.41		
B. Performance tasks assessment					
24 years and below	10	3.44	.58	1.601	.141
25 – 30	32	3.53	.50		
31 – 35	18	3.31	.47		
36 – 40	18	3.61	.38		
41 – 45	13	3.62	.39		
46 – 50	7	3.37	.57		
51 – 55	22	3.35	.50		
56 and above	9	3.78	.16		
Total	129	3.49	.47		
C. Reflective and compassionate assessment					
24 years and below	10	3.51	.30	1.964	.065
25 – 30	32	3.55	.37		
31 – 35	18	3.33	.51		
36 – 40	18	3.68	.34		
41 – 45	13	3.45	.26		
46 – 50	7	3.62	.52		
51 – 55	22	3.71	.34		
56 and above	9	3.65	.17		
Total	129	3.56	.38		

Legend: 1.00-1.49: Practiced at a little extent (PLE); 1.50-2.49: Practiced at a moderate extent (PME); 2.50 – 3.49: Practiced at a great extent (PGE); 3.50-4.00: Practiced at a very great extent (PGVE)

*Significant at 0.05

Table 3: Comparison of the level of assessment when grouped according to age

Table 4 presents assessment-practice scores across years-in-service groups. Statistically significant differences were observed for Written Works Assessment ($F(2,126) = 5.413, p = .006$) and Reflective and Compassionate Assessment ($F(2,126) = 4.732, p = .010$). Performance Tasks Assessment showed no statistically significant difference across service groups ($F(2,126) = 0.837, p = .436$).

Tukey's HSD post hoc analysis indicated that teachers with ≥ 11 years of service had significantly higher Written Works scores ($M = 3.53, SD = 0.38$) than teachers with 4–10 years of service ($M = 3.27, SD = 0.42, p = .004$). Teachers with at least 11 years of service also had higher Reflective and Compassionate Assessment scores ($M = 3.68, SD = 0.32$) than teachers with 0–3 years of service ($M = 3.46, SD = 0.37, p = .012$).

Years in Service	N	Mean	SD	F	Sig
A. Written works assessment					
0–3	35	3.36	.39	5.413	.006*
4–10	40	3.27	.42		
11 and above	54	3.53	.38		
Total	129	3.40	.41		
B. Performance tasks assessment					
0–3	35	3.47	.50	.837	.436
4–10	40	3.42	.48		
11 and above	54	3.55	.45		
Total	129	3.49	.47		
C. Reflective and compassionate assessment					
0–3	35	3.46	.37	4.732	.010*
4–10	40	3.50	.42		
11 and above	54	3.68	.32		
Total	129	3.56	.38		

Legend: 1.00–1.49: Practiced at a little extent (PLE); 1.50–2.49: Practiced at a moderate extent (PME); 2.50–3.49: Practiced at a great extent (PGE); 3.50–4.00: Practiced at a very great extent (PGVE)

*Significant at 0.05

Table 4: Comparison of the level of assessment when grouped according to years in service

Section 3. Comparison of Junior High School Social Studies teachers' assessment practices across the three domains

A repeated-measures ANOVA was conducted to test whether teachers' mean assessment-practice scores differed across the three domains. Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2(2) = 13.725, p = .001$). Therefore,

degrees of freedom were corrected using the Greenhouse-Geisser estimate ($\epsilon = 0.907$).

As shown in Table 5, using Greenhouse-Geisser-corrected degrees of freedom (1.814, 232.19), the overall effect of assessment domain was statistically significant ($F() = 12.413, p < .001, \eta_p^2 = .088$). The corrected partial eta squared was $\eta_p^2 = .088$, indicating a moderate omnibus domain effect.

Type of Assessment			Mean	Std. Deviation
Written Works Assessment			3.40	.41
Performance Task Assessment			3.49	.47
Reflective and compassionate assessment			3.56	.38
Type	Df	F	Sig	Partial Eta Squared
Greenhouse-Geisser*	1.814	12.413	< .001	.088

* Mauchly's test indicated that the sphericity assumption was violated, $\chi^2(2) = 13.725, p = .001$; Greenhouse-Geisser $\epsilon = .907$.

Table 5: Descriptive statistics and test within subjects' effects for the types of assessment

Section 4. Challenges and responses related to assessment during heat-index-related class suspensions

Subtheme 1: Delays in Submission and Accessibility Barriers

The first subtheme concerns delayed submission associated with structural, environmental, and socio-economic barriers. Teachers reported that learners with stable internet access

and digital devices could submit more readily, whereas others experienced delays associated with weak connectivity, limited device access, household responsibilities, and lack of adult supervision.

Illustrative verbatim responses from participants include:

"Late submission of outputs and slow internet connection."

"Written works were submitted on time to others who have gadgets, but others were late because they have no cellphones for updates."

"Most students late in submitting their activity sheets because most of them are working and helping their parents in their fields."

"Poor signal due to weather conditions; unreachable because some students does not have internet connectivity."

"Late submission of outputs, difficulty of learners in answering the modules because no one is guiding them."

Subtheme 2: Logistical Impediments in Collecting Hardcopy Outputs

The second subtheme concerns administrative and logistical difficulties in retrieving printed self-learning modules (SLMs) and activity sheets. For learners in remote or rural areas, hardcopy collection created delays, and teachers reported receiving incomplete or unanswered modules, particularly for more demanding performance tasks.

Key narrative statements supporting this subtheme include:

"The difficulty in collecting the written work (hardcopies) especially in the barrios."

"Hard to collect output and hard to reach out students."

"After retrieving learner's unit test, I don't undergo for item analysis sometimes because of too much work."

"Some learners leave the performance task unanswered."

"Many learner were not able to perform the task given to them because of limited resources."

Subtheme 3: Compromised Authenticity and Output Quality

The third subtheme concerns the perceived authenticity, rigour, and quality of student work. Teachers reported instances of verbatim internet copying, superficial task completion, and work completed by parents or older siblings. They also described difficulties with performance tasks requiring materials, technology, or higher-order creative work. Participant perspectives illustrating these integrity and quality concerns include:

"Almost all students rely their answers from google."

"Some are just copying what's provided in the net without reading."

"I am challenged that their answers are not really theirs or is a product of copying."

"There is no authenticity of learning since some activity sheets were answered by parents or guardian."

"Some of their work do not meet standards, lack creativity, and originality. I find it difficult evaluating their outputs because a few meet the standards and even go beyond while many do not."

"Performance tasks that requires much creativity (e.g. videos, social media advocacy campaigns) were not accomplished by all students due maybe to parental guidance for most students, and lack of enough know how technology use."

Subtheme 4: Implementation of Remedial and Compassionate Adjustments

The fourth subtheme illustrates how teachers used reflective and compassionate adjustments to support learners facing incomplete submissions and environmental constraints. Reported strategies included extended deadlines, simplified

tasks, follow-up feedback, remediation, and home visits.

Excerpts reflecting this empathetic stance include:

"Always consider the needs of your learners and adjust to their capacities."

"Be human and compassionate in giving assessment."

"Students are lazy to answer or do performance task that we need to return back their papers and call them from time to time."

"I, as a teacher is responsible in giving feedbacks, follow-through, reinforcements, remediation, and even home visitations because these mechanisms are crucial in the teaching and learning process."

Subtheme 5: Pedagogical Adaptations Across Learner Profiles

The final subtheme examines how teachers differentiated assessment practices according to learners' independence and the availability of household support. Teachers described performance tasks as particularly difficult to supervise remotely because learners differed in home resources, guidance, and access to technology.

To address these variations, teachers tailored their strategies according to four distinct learner profiles:

1. Independent Learners with Supportive Adults:

Teachers provided standard modules, online resources, and extension tasks requiring deeper critical reflection. Communication was maintained asynchronously.

o *Jas (15 years in service): "Through written works and performance task that is suitable to their level and sometimes if the task in the module is not enough we give them additional activities. I also monitor them via call/messenger app."*

2. Dependent Learners with Supportive Adults:

Teachers established structured home-school agreements, providing direct instructional guides to parents to facilitate learning supervision.

o *Lorie (6 years in service): "With my supervision I let my student to do the task and I also explain the task to the parents so that they could teach their child if they have free time."*

3. Independent Learners with Challenged Adults:

Teachers streamlined module requirements and developed simplified, teacher-made Learning Activity Sheets (LAS) to enable autonomous study.

o *Cesar (4 years in service): "Streamlining the activities given and sometimes giving them teacher made LAS that is made easy for them to finish..."*

4. Dependent Learners with Challenged Adults:

Teachers prioritized maximum compassionate assessment, conducting targeted home visitations, simplifying core competencies, and offering extensive remedial options.

DISCUSSION

The quantitative and qualitative findings illustrate how Junior High School Social Studies teachers reported adapting assessment during heat-index-related class suspensions. Descriptively, Reflective and Compassionate Assessment had the highest

mean score, followed by Performance Tasks and Written Works. The repeated-measures analysis indicated an overall difference among domain scores, while the qualitative findings showed how teachers adjusted deadlines, feedback, task demands, and communication in response to practical constraints.

From an ecological-systems perspective, heat-index-related suspensions interrupt the school setting and increase dependence on home-based learning and home-school interaction. When instruction moves temporarily to remote or modular arrangements, differences in connectivity, available materials, household support, and teacher access can become more consequential. In Social Studies, where learning often involves discussion, critical inquiry, and civic reasoning (NCSS, 2015; Reyes, 2018; Rogayan and Villanueva, 2019), these conditions can complicate the assessment of higher-order outcomes.

The reported adaptations are broadly consistent with DepEd Order No. 31, s. 2020, which provides guidance for flexible assessment under continuity constraints. Teachers described extensions, remedial work, home visits, and communication through tools such as calls and messaging applications. These practices illustrate the role of home-school communication in maintaining assessment processes when face-to-face instruction is interrupted.

The demographic comparisons showed differences in selected assessment domains. Teachers with ≥ 11 years of service reported higher Written Works scores than teachers with 4–10 years of service and higher Reflective and Compassionate Assessment scores than teachers with 0–3 years of service. These cross-sectional group differences may reflect experience, role demands, professional learning, or other unmeasured factors; they do not establish that longer service causes greater assessment competence.

The demographic comparisons may also identify groups that could benefit from differentiated professional support. However, the cross-sectional design does not establish that teaching experience itself caused the observed differences, and the results should not be interpreted as evidence that early-career teachers are inherently less capable of adapting assessment practices.

The repeated-measures ANOVA indicated an overall difference among the three assessment-domain scores ($\eta_p^2 = .088$). Descriptively, Written Works had the lowest mean, followed by Performance Tasks and Reflective and Compassionate Assessment. Because pairwise within-subject domain contrasts were not reported, the omnibus result should not be interpreted as establishing a statistically significant difference between every pair of domains. The pattern nevertheless indicates that teachers did not report identical levels of implementation across all three forms of assessment.

The qualitative findings also identified several operational challenges accompanying these assessment adaptations:

- **Submission Bottlenecks and Socio-Economic Constraints:** Teachers reported that limited devices, weak internet connectivity, household responsibilities, and agricultural work contributed to delayed or incomplete submissions (Penamora, 2021; Pascual, 2014; Krejčová, Rymešová and Chýlová, 2023).

- **Logistical and Workload Strain:** Retrieving physical modules from remote barangays created administrative burdens and could delay instructional follow-up. Similar workload and retrieval difficulties have been reported in Philippine remote-learning contexts (Robosa et al., 2021).
- **Threats to Assessment Validity:** Teachers described internet copying and work completed by parents or other family members, raising concerns about whether submitted outputs accurately represented students' own learning (Agayon, Agayon and Pentang, 2022).
- **Variability in Home Learning Support:** Teachers reported adapting tasks and follow-up according to learners' available support and resources. Home-school collaboration can facilitate remote learning, but uneven access to guidance and materials may require additional scaffolding, simplified Learning Activity Sheets (LAS), contextualised rubrics, or home visits (Miguel and Abulon, 2016; Soliven, 2020; Tan-Choi et al., 2020).

Overall, reflective and compassionate adaptations may help sustain assessment participation during heat-related disruptions, but the present study did not measure failure rates, learning gains, or competency mastery. The findings therefore support conclusions about reported assessment practices and operational challenges rather than the effectiveness of those practices for student outcomes.

CONCLUSIONS AND RECOMMENDATIONS

During heat-index-related class suspensions, Social Studies teachers reported using written work, performance tasks, and reflective or compassionate assessment practices, with the highest mean score observed for Reflective and Compassionate Assessment. Assessment-practice scores differed across age and years-in-service groups in selected domains: Written Works differed by age and years in service, while Reflective and Compassionate Assessment differed by years in service. These cross-sectional differences should not be interpreted as evidence that age or experience causes greater adaptive capacity. Qualitative findings further identified challenges involving submission delays, hardcopy retrieval, output authenticity, and teacher workload.

To strengthen assessment continuity during future climate-related disruptions, policymakers and school administrators may consider the following targeted actions:

1. DepEd and school divisions could develop low-resource, climate-resilient repositories of Social Studies performance tasks and adaptable rubrics that can be activated during ADM transitions.
2. Professional-development programmes could provide targeted support in emergency assessment design, authentic assessment, rubric use, and asynchronous feedback, particularly where local needs assessments identify gaps.
3. School divisions could also consider mentoring arrangements that enable experienced and less experienced teachers to exchange practical strategies for balancing assessment requirements with flexible and compassionate support.

LIMITATIONS

This study has several methodological limitations that should be considered when interpreting the findings:

1. The sample was restricted to public Junior High School teachers within a single division (Nueva Vizcaya), and some demographic subgroups were small. Regional variation in infrastructure, climate exposure, and ADM resources limits generalisation beyond this context, and subgroup ANOVA results should be interpreted cautiously.
2. Quantitative domain scores relied on self-reported survey responses, which may introduce social desirability bias regarding reported levels of compassionate and standard-compliant assessment practices.

3. Data collection was conducted across a single academic quarter (Quarter III) during acute heat suspensions. Longitudinal studies tracking student learning retention across multiple climate-disrupted school years were beyond the scope of this project.
4. The study focused on teacher-reported assessment practices. Student achievement, failure rates, home temperature exposure, and household socio-economic status were not directly measured; consequently, the study cannot determine whether the reported adaptations improved learning outcomes or educational effectiveness.

ACKNOWLEDGMENT

The author gratefully acknowledges the teachers who participated in the study and contributed their time and experiences.

REFERENCES

- Adam, M., Bekoe, S.O. and Poatob, S. (2018) 'Formative Assessment in the Social Studies Classroom: How Senior High School (SHS) Teachers in Ghana Actualise It', *European Journal of Training and Development Studies*, Vol. 5, No. 1, pp. 14–23.
- Agayon, A. J. D., Agayon, A. K. R. and Pentang, J. T. (2022) 'Teachers in the New Normal: Challenges and Coping Mechanisms in Secondary Schools', *Journal of Humanities and Education Development*, Vol. 4, No. 1, pp. 67–75. <https://doi.org/10.22161/jhed.4.1.8>
- Boonying, V. and Chatrurachewin, C. (2017) 'Teacher Professional Learning Community (TPLCs) on Social Studies Toward 21st Century Learning: A Case Study of Primary Schools in Sukhothai Province, Thailand', *Képzés és Gyakorlat*, Vol. 15, No. 3, pp. 9–20. <https://doi.org/10.17165/tp.2017.3.1>
- Braun, V. and Clarke, V. (2022) 'Conceptual and Design Thinking for Thematic Analysis', *Qualitative Psychology*, Vol. 9, No. 1, pp. 3–26. <https://doi.org/10.1037/qup0000196>
- Cagasan, L., Care, E., Robertson, P. and Luo, R. (2020) 'Developing a Formative Assessment Protocol to Examine Formative Assessment Practices in the Philippines', *Educational Assessment*, Vol. 25, No. 4, pp. 259–275. <https://doi.org/10.1080/10627197.2020.1766960>
- Cahapay, M. B. (2020a) 'Reshaping Assessment Practices in a Philippine Teacher Education Institution during the Coronavirus Disease 2019 Crisis', *Pedagogical Research*, Vol. 5, No. 4, p. em0079. <https://doi.org/10.29333/pr/8535>
- Cahapay, M.B. (2020b) 'Rethinking Education in the New Normal Post-COVID-19 Era: A Curriculum Studies Perspective', *Aquademia*, Vol. 4, No. 2, p. ep20018. <https://doi.org/10.29333/aquademia/8315>
- Calveric, S.B. (2010) *Elementary Teachers' Assessment Beliefs and Practices*, PhD dissertation, Richmond, VA: Virginia Commonwealth University. <https://doi.org/10.25772/OXTP-8W56>
- Crumb, J. S. (2021) *How Does the Use of a Social Studies Video Game Affect Student Learning?*, doctoral dissertation, Atlanta, GA: Georgia State University. Available at: <https://hdl.handle.net/20.500.14694/10752>
- De Vera, J.L. (2020) 'Challenges and Teacher Resilience: The New Normal Classroom Instruction Using Social Media in Philippine Context', in Barman, P. (ed.) *Social Media: Leisure, Health and Education*, New Delhi: New Delhi Publishers, pp. 83–96.
- Department of Education (2020a) *DO 012, s. 2020: Adoption of the Basic Education Learning Continuity Plan for School Year 2020–2021 in the Light of the COVID-19 Public Health Emergency*, [Online]. Available at: <https://www.deped.gov.ph/2020/06/19/june-19-2020-do-012-2020-adoption-of-the-basic-education-learning-continuity-plan-for-school-year-2020-2021-in-the-light-of-the-covid-19-public-health-emergency/>
- Department of Education (2020b) *DO 031, s. 2020: Interim Guidelines for Assessment and Grading in Light of the Basic Education Learning Continuity Plan*, [Online]. Available at: <https://www.deped.gov.ph/2020/10/02/october-2-2020-do-031-s-2020-interim-guidelines-for-assessment-and-grading-in-light-of-the-basic-education-learning-continuity-plan/>
- Department of Education (2020c) *Memorandum OUCI-2020-307: Suggested Measures to Foster "Academic Ease" During the COVID-19 Pandemic*, [Online]. Available at: <https://www.deped.gov.ph/wp-content/uploads/2020/11/OUCI-2020-307-ACADEMIC-EASE-DURING-COVID-19-final-version-01Nov2020.pdf>
- Ferretti, F., Santi, G. R. P., Del Zozzo, A., Garzetti, M. and Bolondi, G. (2021) 'Assessment Practices and Beliefs: Teachers' Perspectives on Assessment during Long Distance Learning', *Education Sciences*, Vol. 11, No. 6, p. 264. <https://doi.org/10.3390/educsci11060264>
- Gonzales, R. D. L. C. and Callueng, C. M. (2014) 'Classroom Assessment Practices of Filipino Teachers: Measurement and Impact of Professional Development', in Magno, C. and Gaerlan, M.J. (eds.) *Essentials on Counseling and Education: A Festschrift of Prof. Rose Marie Salazar-Clemena*, Manila: Asian Psychological Services and Assessment (APSA), pp. 220–242. <https://doi.org/10.13140/2.1.1950.3685>
- Goodman, L. A., Wilson, J. M., Helms, J. E., Greenstein, N. and Medzhitova, J. (2018) 'Becoming an Advocate: Processes and Outcomes of a Relationship-Centered Advocacy Training Model', *The Counseling Psychologist*, Vol. 46, No. 2, pp. 122–153. <https://doi.org/10.1177/0011000018757168>
- Hopkins, M. T. (2013) 'Teachers' Practices and Beliefs Regarding Teaching Tuning in Elementary and Middle School Group String Classes', *Journal of Research in Music Education*, Vol. 61, No. 1, pp. 97–114. <https://doi.org/10.1177/0022429412473607>

- Keçe, M. (2014) 'Problems Related to the Teaching of Social Studies and Suggestions for Solution: Teachers' Opinions Based on a Qualitative Research', *Procedia – Social and Behavioral Sciences*, Vol. 122, pp. 388–392. <https://doi.org/10.1016/j.sbspro.2014.01.1359>
- Koloi-Keaikitse, S. (2012) *Classroom Assessment Practices: A Survey of Botswana Primary and Secondary School Teachers*, PhD dissertation, Muncie, IN: Ball State University. Available at: <https://eric.ed.gov/?id=ED543182>
- Krejčová, K., Ryměšová, P. and Chýlová, H. (2023) 'Self-compassion as a Newly Observed Dimension of the Student's Personality', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 16, No. 2, pp. 140–148. <https://doi.org/10.7160/eriesj.2023.160205>
- Meziobi, D. I., Oyeoku, E. K. and Ezegbe, B.N. (2012) 'The Challenges of Social Studies Educators for Effective Teacher Preparation and Implementation of the Universal Basic Education Programme in Nigerian Junior Secondary Schools', *US-China Education Review B*, No. 11, pp. 947–955.
- Miguel, F. F. and Abulon, E. L. R. (2016) 'Teacher-Parent Collaborative Efforts in Facilitating Students' Homework', *The Normal Lights*, Special Issue 2016, pp. 18–32. <https://doi.org/10.56278/tnl.v1i1.273>
- Mikre, F. (2010) 'Review Article: The Roles of Assessment in Curriculum Practice and Enhancement of Learning', *Ethiopian Journal of Education and Sciences*, Vol. 5, No. 2, pp. 101–114. <https://doi.org/10.4314/ejesc.v5i2.65376>
- Mussawy, S. A. J. (2009) *Assessment Practices: Students' and Teachers' Perceptions of Classroom Assessment*, Master's Capstone Project, Amherst, MA: University of Massachusetts Amherst. Available at: https://scholarworks.umass.edu/cie_capstones/9/
- Nasab, F. G. (2015) 'Alternative versus Traditional Assessment', *Journal of Applied Linguistics and Language Research*, Vol. 2, No. 6, pp. 165–178. Available at: <http://www.jallr.com/index.php/JALLR/article/view/136>
- National Council for the Social Studies (NCSS) (2015) *National Curriculum Standards for Social Studies: Executive Summary*. Available at: <https://www.socialstudies.org/standards/national-curriculum-standards-social-studies-executive-summary>
- Nunvarova, J., Poulouva, P. and Prazak, P. (2023) 'Digital Storytelling in Economics Subjects and its Effectiveness on Student Learning Outcomes by Gender and Different Economic Knowledge', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 16, No. 3, pp. 208–219. <https://doi.org/10.7160/eriesj.2023.160305>
- Oducado, R. M. F., Rabacal, J. S., Moralista, R. B. and Tamdang, K. A. (2021) 'Perceived Stress Due COVID-19 Pandemic Among Employed Professional Teachers', *IJERI: International Journal of Educational Research and Innovation*, No. 15, pp. 305–316. <https://doi.org/10.46661/ijeri.5284>
- Robosa, J., Paras, N. E., Perante, L., Alvez, T. and Tus, J. (2021) 'The Experiences and Challenges Faced of the Public School Teachers Amidst the COVID-19 Pandemic: A Phenomenological Study in the Philippines', *International Journal of Advance Research and Innovative Ideas in Education*, Vol. 7, No. 1, pp. 1342–1361.
- Park, R. J., Goodman, J., Hurwitz, M. and Smith, J. (2020) 'Heat and Learning', *American Economic Journal: Economic Policy*, Vol. 12, No. 2, pp. 306–339. <https://doi.org/10.1257/pol.20180612>
- Pascual, N. T. (2014) 'Factors Affecting High School Students' Career Preference: A Basis for Career Planning Program', *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, Vol. 16, No. 1, pp. 1–14.
- Penamora, A. M. (2021) 'ML O LM: Stories of Mobile Legend Gamers in Accomplishing Their Learning Modules Amidst COVID-19 Pandemic: A Case Study', *International Journal of Discoveries and Innovations in Applied Sciences*, Vol. 1, No. 7, pp. 17–35.
- Rahman, A. A., Muhamad Hanafi, N., Ibrahim Mukhtar, M. and Ahmad, J. (2014) 'Assessment Practices for Competency Based Education and Training in Vocational College, Malaysia', *Procedia – Social and Behavioral Sciences*, Vol. 112, pp. 1070–1076. <https://doi.org/10.1016/j.sbspro.2014.01.1271>
- Reyes, J. R. S., Grajo, J. D. L., Comia, L. N., Talento, M. S. D. P., Ebal, L.P. A. and Mendoza, J. J. O. (2021) 'Assessment of Filipino Higher Education Students' Readiness for e-Learning During a Pandemic: A Rasch Technique Application', *Philippine Journal of Science*, Vol. 150, No. 3, pp. 1007–1018. <https://doi.org/10.56899/150.03.34>
- Robosa, J., Paras, N. E., Perante, L., Alvez, T. and Tus, J. (2021) 'The Experiences and Challenges Faced of the Public School Teachers Amidst the COVID-19 Pandemic: A Phenomenological Study in the Philippines', *International Journal of Advance Research and Innovative Ideas in Education*, Vol. 7, No. 1, pp. 1342–1361.
- Rogayan, D. V., Jr. and Villanueva, E. E. N. (2019) 'Implementation Status of K12 Social Studies Program in Philippine Public Schools', *PEOPLE: International Journal of Social Sciences*, Vol. 5, No. 3, pp. 233–250. <https://doi.org/10.20319/pijss.2019.53.233250>
- Rundle, N. (2016) *Assessment. Teaching and Learning - University of Tasmania, Australia*. Available at: <https://www.teaching-learning.utas.edu.au/assessment>
- Sarmiento, C. P., Morales, M. P. E., Elipane, L. E. and Palomar, B. C. (2020) 'Assessment Practices in Philippine Higher STEAM Education', *Journal of University Teaching & Learning Practice*, Vol. 17, No. 5, p. 18. <https://doi.org/10.53761/1.17.5.17>
- Serafico-Reyes, N. M.A. (2018) 'Civic Competency in Junior High School – Mapping through the Content and Performance Standards of the Philippine Social Studies (Araling Panlipunan) Curriculum', in *Proceedings of the 2nd Asian Education Symposium (AES 2017)*, Vol. 1, pp. 467–472. SciTePress. <https://doi.org/10.5220/0007306304670472>
- Soliven, S. (2020) *Assessing Learning with Compassion While Using Modular and Online Distance Educational Learning [presentation]*.
- Stifel, S. W. F., Feinberg, D. K., Zhang, Y., Chan, M. K. and Wagle, R. (2020) 'Assessment During the COVID-19 Pandemic: Ethical, Legal, and Safety Considerations Moving Forward', *School Psychology Review*, Vol. 49, No. 4, pp. 438–452. <https://doi.org/10.1080/2372966X.2020.1844549>
- Suryawati, E. and Osman, K. (2018) 'Contextual Learning: Innovative Approach towards the Development of Students' Scientific Attitude and Natural Science Performance', *EURASIA Journal of Mathematics, Science and Technology Education*, Vol. 14, No. 1, pp. 61–76. <https://doi.org/10.12973/ejmste/79329>
- Tan-Choi, A., Tinio, V.L., Castillo-Canales, D., Lim, C.P., Modesto, J.G. and Pouzevara, S.R. (2020) *Teacher's Guide for Remote Learning During School Closures and Beyond*, Quezon City: Foundation for Information Technology Education and Development. Available at: <https://tpdatascalecoalition.org/publication/teachers-guide-to-remote-learning/>

- Tantengco, N.S. and Maramag, R.L. (2016) 'Examining Gender Responsiveness of the Philippine Basic Education Reform: An Analysis of the K-12 Araling Panlipunan or Social Studies Curriculum', *MIMBAR PENDIDIKAN*, Vol. 1, No. 1, pp. 37–54. <https://doi.org/10.17509/mimbardik.v1i1.1752>
- Teachers' Council of Thailand (2018) *Southeast Asia Teachers Competency Framework (SEA-TCF)*, Bangkok: Teachers' Council of Thailand.
- Torney-Purta, J., Cabrera, J. C., Roohr, K. C., Liu, O. L. and Rios, J. A. (2015) *Assessing Civic Competency and Engagement in Higher Education: Research Background, Frameworks, and Directions for Next-Generation Assessment*, ETS Research Report No. RR-15-34, Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/ets2.12081>
- UNICEF Regional Office for South Asia (2020) *Guidance on Distance Learning Modalities*, Kathmandu: UNICEF Regional Office for South Asia. Available at: <https://www.unicef.org/rosa/reports/guidance-distance-learning-modalities-reach-all-children-and-youth-during-school-closures>
- Umugiraneza, O., Bansilal, S. and North, D. (2017) 'Exploring Teachers' Practices in Teaching Mathematics and Statistics in KwaZulu-Natal Schools', *South African Journal of Education*, Vol. 37, No. 2, pp. 1–13. <https://doi.org/10.15700/saje.v37n2a1306>
- United Nations Department of Economic and Social Affairs (n.d.) *Goal 4: Quality Education*, [Online]. Available at: <https://sdgs.un.org/goals/goal4>
- Walde, G.S. (2016) 'Assessment of the Implementation of Continuous Assessment: The Case of METTU University', *European Journal of Science and Mathematics Education*, Vol. 4, No. 4, pp. 534–544.
- World Health Organization (2020) *WHO Coronavirus (COVID-19) Dashboard*. Available at: <https://covid19.who.int/>
- Zapletal, F., Hudec, M., Švaňa, M., Chytilová, L., Hlaváček, K., Lokaj, A., Urbanek, A., Glova, J., Samartinho, J. P., Rodriguez, C. M. C. and Guðnason, S. (2023) 'The Complex Evaluation of the Impact of COVID-19 Pandemic at Universities: A Soft Computing Approach', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 16, No. 3, pp. 231–244. <https://doi.org/10.7160/eriesj.2023.160307>
- Zhang, C., Yan, X. and Wang, J. (2021) 'EFL Teachers' Online Assessment Practices During the COVID-19 Pandemic: Changes and Mediating Factors', *The Asia-Pacific Education Researcher*, Vol. 30, No. 6, pp. 499–507. <https://doi.org/10.1007/s40299-021-00589-3>