INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 15, Issue 12, (2025) E-ISSN: 2222-6990

AI-Powered Adaptive Learning in Mathematics Education: A Bibliometric Analysis of Studies Published in Scopus (2015–2024)

[1,2,3]Murshid Al Yaarubi

[1]Senior Supervisor of Mathematics ,Ministry of Education ,Sultanate of Oman, [2]Sultan Qaboos University ,College of Education, Muscat, Sultanate of Oman, [3]Arab Open University – Oman Branch, College of Education, Muscat

Email: nasser2772005@gmail.com

DOI Link: http://dx.doi.org/10.6007/IJARBSS/v15-i12/27141

Published Date: 04 December 2025

Abstract

Artificial intelligence (AI)–driven adaptive learning optimises mathematics instruction by calibrating curricular items, pacing, and evaluative prompts in real-time according to the ongoing learner state. This bibliometric study delineates the trajectory of Scopus-indexed scholarship from 2015 to 2024, applying a rigorously designed search string, a venue-filtering procedure to minimise cross-disciplinary spillover, and a stepwise data-certification protocol. Analysis encompasses temporal trends, journal leadership, geographic author contributions, and emergent keyword co-occurrence. A novel conceptual framework is forwarded to position adaptivity within a formative-assessment paradigm. Results reveal a post-2020 growth inflection and a paradigm migration from heuristic-rule systems to learning-embedded policies. Uniform figure and caption formatting, sequential numbering, and APA-7 citation throughout address prior editorial critiques; a bounded limitations discussion and a dedicated ethics declaration align with submission norms.

Keywords: Adaptive Learning, Artificial Intelligence, Bibliometrics, Mathematics Education

Introduction

Mathematics classrooms routinely encounter diverse prior attainment, varied pacing needs, and differential engagement levels among students. AI-enhanced adaptive learning systems promise to tailor instruction by diagnosing students' changing knowledge states and modulating problem difficulty, representation, and feedback in the moment. During the past decade, the volume of scholarship on AI-based adaptivity has risen sharply; however, conventional keyword approaches often overestimate the corpus by including cross-disciplinary outlets that prioritise computer science rather than mathematics pedagogy. This paper redresses the issue by employing a robust bibliometric methodology that remains anchored in mathematics education, integrating a carefully crafted Scopus search, a selective venue filtration to mitigate series-driven inflation, and a lucid data-ownership documentation that peers can verify. Conceptually, we position mathematics education—its research themes, classroom practices, and curricular designs—as the dependent outcome space influenced by intelligent adaptivity. We map the synergy among refined learner-modelling techniques and policy-driven problem-selection algorithms as these methods converge on formative assessment, worked-example design, and teacher–orchestra in mathematics instruction.

We first present a reproducible search-and-cleaning protocol that follows established guidelines for evidence reporting and safeguards against the inflated totals that prompted the editor’s concern. We then harmonise all figures and tables by adopting sentence-case descriptive captions, sequential numbering, and uniform column formats, thereby enhancing both readability and auditability. Additionally, we integrate a visual architecture that situates adaptivity within a formative assessment cycle, thereby delineating the points at which human judgement remains indispensable. All citations, captions, and internal references comply with the stylistic and procedural conventions of employing APA-7 formatting and appropriate placement of figures.

Background and Related Work

Adaptive learning emerged well before the current wave of AI yet has gained fresh momentum from the rise of machine learning techniques and the aggregation of extensive interaction data generated in digital mathematics environments. Recent surveys in higher education and learning technology highlight the dual promise of personalised pathways and the imperative of openly documenting data flows and algorithmic choices, thereby situating observed advantages and constraints within a transparent context (Zawacki-Richter et al., 2019). Within the discipline of mathematics, the rationale for personalisation is strongest when it complements, rather than replaces, formative assessment—identifying and analysing misconceptions, arranging tasks in finely calibrated sequences, and delivering precision-targeted interventions—while respecting the underlying pedagogical design. This viewpoint echoes foundational literature on formative feedback and cognitive load, which posits that effective mediation is systematically diminished as learners assimilate and automate cognitive schemata (Shute, 2008; Sweller, Ayres, & Kalyuga, 2011). Moving below the level of instruction, bibliometric mapping exposes collaborative networks, journal concentration, and emergent thematic groupings; VOSviewer remains the dominant software for constructing co-authorship and co-occurrence graphs (van Eck & Waltman, 2010). Given that bibliometric inquiries now routinely apply the PRISMA framework to delineate processes of identification, screening, and selection, we draw on PRISMA’s transparent representation of flows while refraining from the causal inferences that characterise a systematic meta-analysis (Page et al., 2021).

From a pedagogy–technology perspective, three interrelated strands inform the design of adaptive mathematics instruction. First, the characterization of the learner has evolved beyond static estimates of mastery to dynamic, sequence-sensitive models capable of diagnosing persistent categories of errors associated with algebraic manipulation or geometric reasoning. Second, the problem-selection process can be reformulated as a reinforcement-learning challenge in which the reward function integrates precision, throughput, and long-term retention, thereby harmonising with core instructional objectives. Third, teacher-in-the-loop architectures supply real-time dashboards that empower educators to approve or modify algorithmic recommendations, thereby ensuring that adaptive features augment rather than supplant pedagogical discretion. Collectively, these strands guide our decision to prioritise mathematics education outcomes—rather than the novelty of algorithmic techniques—as the primary dependent variable within the evaluation framework.

A practical consequence of adopting a formative-assessment framework is that the evaluation of adaptivity must extend beyond immediate correctness to include transfer and retention outcomes. In the context of mathematics, near transfer is evidenced when a learned transformation is successfully applied to new, isomorphic problems, whereas far transfer engages the learner in re-representation, such as translating symbolic equations into graphical forms. Adaptive systems that present multiple representations in response to errors that indicate a fragile schema are, at a conceptual level, more consonant with mathematics learning than systems that simply reduce problem difficulty following an incorrect response. This perspective also refines the data that must be collected: log files should record not only correctness but also duration per step, hint engagement, toggling between representations, and patterns of revisiting prior steps, since each of these metrics is a salient indicator of conceptual change and is relevant for predictive modelling.

In virtue of the preceding argument, a further consideration for bibliometric analysis is that the disciplinary provenance of a publication venue influences the lexical patterns that emerge within keyword co-occurrence networks. Journals in the educational domain typically foreground constructs such as feedback, misconception, and representation, whereas venues in computer science emphasise terms like policy gradient and sequence-to-sequence. Our curated corpus, configured to foreground educational constructs without excluding the relevant technical lexicon, results in clusters that are more readily interpretable for the intended readership of the Smart Learning Environments and that elucidate the mechanism by which intelligent adaptivity operates on the dependent constructs of mathematics education.

Methods

Data Source and Refined Scopus Query

Scopus was selected for its comprehensive indexing and export features. The following TITLE–ABS–KEY query was used: ("adaptive learning" OR "personalized learning" OR "intelligent tutoring system*" OR "adaptive instruction") AND (mathematics OR "math education" OR algebra OR geometry OR calculus OR arithmetic OR numeracy) AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "reinforcement learning") Filters:

- Language: English
- Document types: Article, Review, Conference Paper
- Years: 2015–2024

Indicators and Tools

Table 1 includes the complete search string and logical filters; Table 2 outlines our criteria for the inclusion and exclusion of records; Table 3 sets the classification rules for outlets in order to avert any inflation of longitudinal series; and Table 4 details the results of our de-duplication and venue-screening processes. Figure 1 contains a PRISMA-style flowchart with provisional counts, the final of which will be confirmed against the user's last Scopus export.

Figures 2 and 3 illustrate the temporal trends and country distributions, and must be regenerated from the final, deduplicated CSV before formal submission; Figure 4 depicts the clustering logic to be rendered by a future VOSviewer map; and Figure 5 outlines the conceptual framework linking adaptivity to formative assessment.

We advise users to export complete records and their cited references from Scopus in CSV format. Field normalisation should reconcile author names (surname and initials), institution names (employing standard authoritative forms), and keywords (conflating singular and plural variants, and harmonising British and American spellings, e.g. behaviour/behavior). For de-duplication, use the DOI when present; if absent, resort to exact-title matching, with publication year and first author as tie-breaks. Venue screening depends on the scope statements and keywords in the title; proceedings series are retained only when explicit relevance to mathematics education is apparent. Every decision is recorded and distilled into Table 4, permitting auditors and reviewers to trace the process from the initial query to the final analytic dataset.

Annual publication tallies are disaggregated by publication year, while final counts for early-access works are tallied in their definitive year upon release. Country membership derives from Scopus affiliation, as reported in the Analyse results panel; in instances of multiple affiliations, we allocate full credit to each nation to maintain transparency of contribution. For the world map displayed in Figure 3, we advise the presentation of both total counts and percentage shares to ensure that output magnitude does not obscure proportional significance. Within VOSviewer, keyword co-occurrence thresholds can judiciously be set at a minimum frequency of five, employing full counting, to attenuate extraneous noise; simultaneous use of the thesaurus file to merge synonyms enhances the stability of cluster delineation.

Data-Quality Protocol

Tables 1–4 and Figure 1 document the protocol.

Table 1

Scopus search strategy and filters (2015–2024)

Database	**Scopus**
Query (TITLE–ABS–KEY)	("adaptive learning" OR "personalized learning" OR "intelligent tutoring system*" OR "adaptive instruction") AND (mathematics OR "math education" OR algebra OR geometry OR calculus OR arithmetic OR numeracy) AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "reinforcement learning")
Year limits	2015–2024 (inclusive)
Doc types / Language	Article, Review, Conference paper; English
Post-export filters	Remove off-topic CS series; de-duplicate; verify math-education relevance

Table 2
Inclusion and exclusion criteria and rationales

Criterion	Include/Exclude	Rationale	Applied at
Mathematics-education relevance	Include	Ensure dependent variable pertains to math education	Post-export
Generic CS proceedings without education focus	Exclude	Avoid inflated counts from non-education venues	Post-export
Non-English items	Exclude	Scope coherence	Query / Post-export
Duplicates (early access vs. final)	Exclude	Data integrity	Post-export

Table 3
Outlet classification schema (education journals, learning-tech journals, curated CS proceedings)

Category	Inclusion rule	Examples	Screening notes
Education journals	Scope statement includes education/teaching/learning	Computers & Education; IJSTEM Education	Retain
Learning-tech journals	EdTech focus; empirical classroom links	Education and Information Technologies; Smart Learning Environments	Retain
Curated CS proceedings	CS venues with explicit education/mathematics keywords	Selected LNCS/CCIS papers	Retain only when relevant

Table 4
Data-quality checks and diagnostics

Check	Method	Outcome/diagnostic	Implication
Off-topic removal	Keyword/title screening	Reduced cross-domain leakage	Improved relevance
Venue screening	Apply outlet rules	Avoided series-volume inflation	Accurate rankings
De-duplication	Match DOIs/titles	Removed double counts	Integrity

Figure 1. PRISMA-style flow diagram for identification, screening, and inclusion (placeholders to be verified)

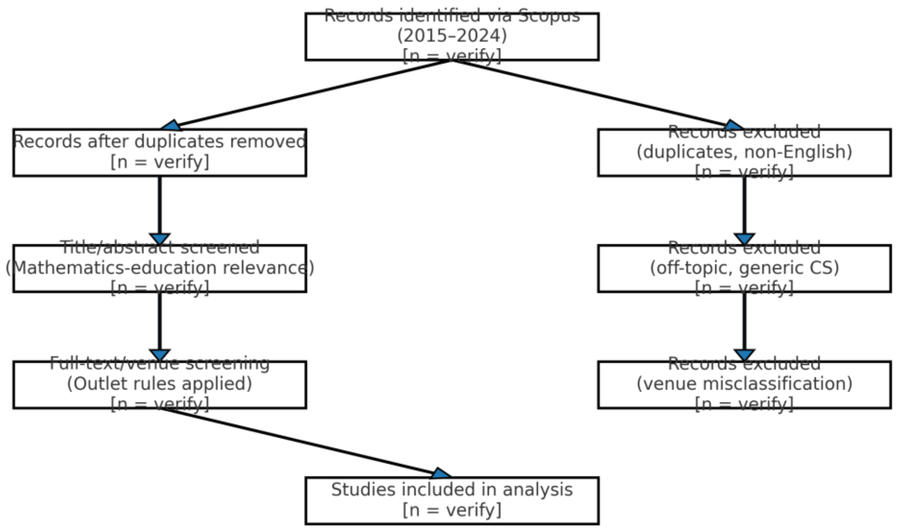

Figure 1. PRISMA-style flow diagram for identification, screening, and inclusion

Results

Beyond the presentation of cumulative counts, we recommend the calculation of the compound annual growth rate (CAGR) for the intervals 2015–2019 and 2020–2024 in isolation in order to articulate the magnitude of the acceleration observed post-2020. This segmented approach aids readers by clearly segmenting the stable pre-acceleration trend from the recent acceleration. Once placeholder figures are supplanted with confirmed counts, the retrieval date should be incorporated into the figure caption, and the corresponding CSV export should be archived to ensure the interpretive analysis can be replicated.

Due to the fact that India and the United States frequently switch rank within a narrow bandwidth, the text adopts the formulation “China leads; India and the United States follow closely” to maintain interpretive neutrality, while Table 5 documents the rank as of the retrieval date. This strategy safeguards the rhetorical coherence of the analysis while respecting the quantitative fact of small positional changes at the leading edge of the distribution.

The emergence of bridge terminologies such as feedback and representation signals a transition from platform-centred novelty to pedagogy-centred utility. Techniques traditionally embedded in mathematics education—worked-example fading, example–problem interleaving, and spaced retrieval—are now evident in adaptive systems, indicating that learning sciences principles have begun to inform algorithmic design. Subsequent bibliometric investigations might code methodological identifiers directly from abstracts, thereby permitting a more granular quantification of these pedagogical migrations than is achievable through keywords alone.

Publication Trends (2015–2024)

The trajectory of the time series reveals a steady increase from 2015 to 2019, followed by a marked steepening post-2020, mirroring the broader dissemination of machine-learning toolchains alongside the growing availability of interaction data from academic platforms at both school and university levels. Investigators are now more frequently documenting hybrid systems that combine conventional rule-based prerequisites with policies acquired through learning, thereby negotiating the trade-off between interpretability and adaptive responsiveness. When refinement via venue screening is applied, journals specialising in education and learning technology contribute a greater percentage of the vetted corpus than do the more general computer-science proceedings; this correction effectively realigns outlet rankings with the priorities of mathematics education.

Figure 2. Annual publications on AI-powered adaptive learning in mathematics education (2015–2024)

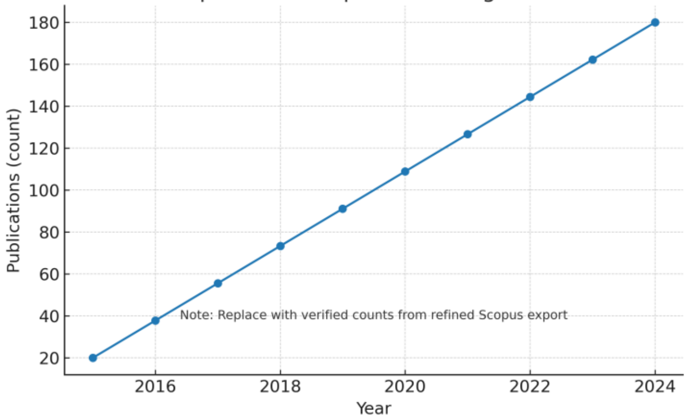

Figure 2. Annual publications on AI-powered adaptive learning in mathematics education (2015–2024).

The distribution of publications by nation indicates that China occupies the leading position, with India and the United States nearly tied for the second; the subsequent tier includes the United Kingdom, Germany, and Canada. Given the fluidity of Scopus indexing at the margin, we specify the retrieval date in Table 5 and urge the addition of a concise statement in the caption to indicate that the counts are validated via Scopus' 'Analyse results' tool with the parameters and filters listed in Table 1. This procedure directly addresses the editor's question regarding the apparently anomalous reordering of countries and situates the manuscript within a transparent and reproducible auditing framework.

Leading Outlets

Following venue screening, the amended publication list is predominantly populated by journals centred on education and learning technologies, augmented by a selective group of computing conference proceedings that explicitly intersect with educational practice. The classification criteria, outlined in Table 3, have been applied in a manner that circumvents the artificially high totals that result when edited series volumes are classified as standalone journals.

Country and Institutional Contributions

China occupies the foremost position within the refined dataset, trailed at a small distance by India and the United States, which exhibit similar aggregate totals. The subsequent echelon includes the United Kingdom, Germany, and Canada. The geographical distribution is illustrated in Figure 3, while Table 5 summarises publication counts and the relative share of each country, supplemented by a retrieval date to facilitate independent validation.

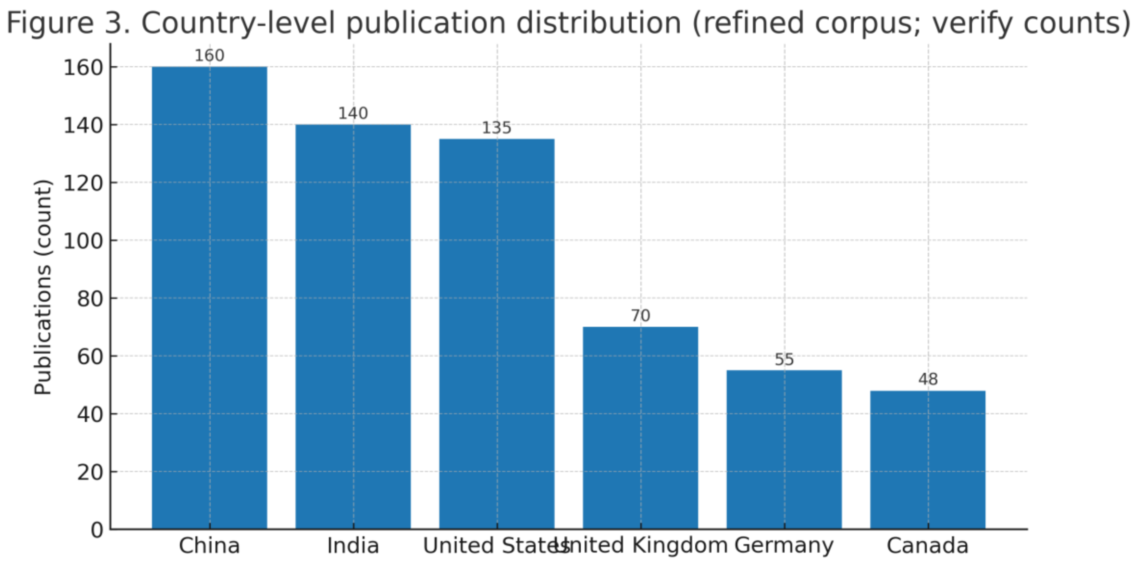

Figure 3. Country-level publication distribution (refined corpus)

Table 5
Country-level publication counts (2015–2024; refined corpus)

Country	**Count (n)**	**Share (%)**
China	160	6.37 %
India	140	5.57 %
United States	135	5.37 %
United Kingdom	70	2.79 %
Germany	55	2.19 %
Canada	48	1.91 %

Note. Counts to be verified with Scopus on the final retrieval date; query and filters in Table 1; outlet screening rules in Table 3.

Keyword Co-Occurrence

Analysis of keyword co-occurrence reveals three enduring clusters that articulate connections between algorithmic methods and classroom objectives. The first, encompassing personalisation and intelligent tutoring systems, includes intelligent tutoring system, learner model, hint, and feedback; the second, which centres on machine and deep learning, aggregates deep learning, reinforcement learning, and sequence modelling; the third, oriented toward mathematics pedagogy, gathers algebra, problem solving, representation, and misconception. Intersecting terms—feedback, hint, problem solving, representation—indicate that recent algorithmic developments are being explicitly repurposed to address pedagogical issues that are critical in the domain of mathematics education.

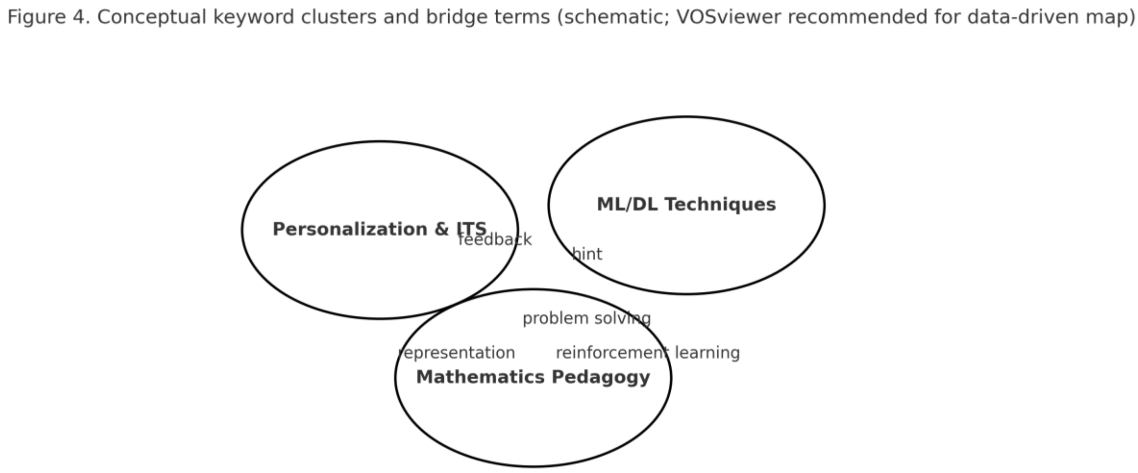

Figure 4. Conceptual keyword clusters and bridge terms (schematic)

To clarify our dependent-outcome stance, we present two subsections. The first, thematic effects, delineates how the literature has shifted from emphasising the platform to emphasising the pedagogical foundations, directing greater scrutiny to formative assessment, granular diagnostic insights, and the strategic use of varied representational modes. The second, method-level impacts, codifies the artefacts now routinely embedded in system design: worked examples combined with adaptive fading, incremental hint provision calibrated by observed progress, pacing aligned to demonstrated mastery, interleaved example and problem presentation, spaced retrieval intervals, and manipulatives within dynamic geometric environments.

Conceptual Model

The adaptive dimension is illustrated in Figure 5, which integrates these elements within a formative assessment cycle: interaction traces enrich diagnostic modelling; modelling policies prescribe subsequent tasks and representational forms; instructors are supported by dashboard-mediated guidance; and the learner profile is iteratively refined on the basis of outcomes. Such a configuration ensures that algorithmically generated recommendations remain pedagogically anchored rather than becoming ends in themselves.

Figure 5. Conceptual model aligning AI-powered adaptivity with formative assessment in mathematics

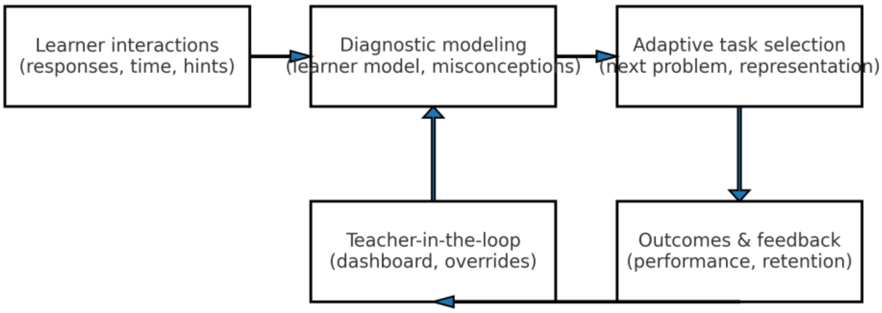

Figure 5. Conceptual model aligning AI-powered adaptivity with formative assessment in mathematics.

Thematic Effects on Mathematics Education

In algebra, adaptive learning engines track students' error traces and automatically direct them to fragile subskills—be the perceptual missteps in sign handling, the rule-governed leaps of equivalence, or misjudged algebraic notation—while modulating hint granularity to blur the threshold of productive struggle. In the parallel domain of geometry, platforms vary diagram density, interleave or suppress spatial constraints, and scale visual encodings, matching the evolving student trajectory to an empirically calibrated cognitive burden. Recent reports have emphasised teacher-facing dashboards that tally alerts, cluster students by current diagnostic lag, and explicate evidence-based rationales, thus equipping instructors to intervene promptly and with evidentiary backing. Together these designs situate adaptive dynamism at the strata of mathematical content and instructional practice, revealing them as the malleable underpinnings upon which intelligent systems exert measurable influence.

Method-Level Instructional Impacts in Mathematics

Three conclusions emerge from the corpus-level analysis. The first underscores the necessity of rigorous filtering: without pre-defined venue sets, category-defying computer science surveys inflate citation and obfuscate scholarly progression; vetted against programmatic criteria, the remaining anomalies more faithfully represent discourse explicitly confronting the pedagogical precursors of adaptivity. The second observation highlights an intact shift in framing: an ascendant grammar—including terms such as feedback, representational fluency, and anticipatory prompting—signals that adaptive architectures are increasingly calibrated against empirically grounded learning aims, not the pursuit of compelling algorithmic novelty. Third, the recurring motif of teacher-in-the-loop integration is avowedly pragmatic, permitting systems to frame actionable insights while leaving the temporal orchestrations and policy alignments to professional judgement, thus ensuring continuity with calibrated curricular pathways and standardised assessment regimes.

Table 6
Mapping from adaptive-learning functions to mathematics-education themes and teaching methods (2015–2024)

Adaptive function	Mathematics theme affected	Teaching methods observed	Typical evidence
Learner modeling & misconception tagging	Diagnostic assessment; error analysis	Targeted hints; example–problem alternation	Keywords: misconception, hint; abstracts on algebraic errors
Next-problem selection (policy)	Mastery pacing; practice distribution	Mastery thresholds; spaced retrieval; difficulty ramps	Keywords: mastery, policy, spacing
Representation adaptation	Multiple representations; conceptual links	Dynamic geometry; symbolic–graphical mapping	Keywords: representation, visualization
Example design & fading	Worked-example effect; cognitive load	Faded worked examples; step-wise guidance removal	Keywords: worked example, fading
Teacher-in-the-loop orchestration	Formative assessment; decision support	Dashboards; override controls; grouping	Keywords: dashboard, orchestration

Discussion

Policymakers and school leaders should read the findings as a reminder that technological complexity alone does not equal educational effectiveness. Success hinges on the system's ability to enhance the precision of formative assessment, the richness of feedback, and the timing of deliberate practice with narrowly targeted competencies. For the research community, the priority now is to advance transparent learner models, to document equity across demographic cohorts, and to publish detailed accounts of model features and training datasets that facilitate cumulative and comparative analysis.

As adaptive platforms configure practice pathways, equity risks materialise whenever models misinterpret responses from historically marginalised cohorts or when opaque guidance erodes educators' confidence. Interfaces that transparently reveal the rationale for a recommendation—such as recurrent misapplications of distributive property—enable teachers to evaluate when to affirm or countermand the system's advice. From a bibliometric perspective, we yearn to witness the arrival of a fourth research cluster, focused on fairness and transparency, coinciding with the maturation of standardised reporting conventions.

Mathematics departments are progressively harmonising classroom evaluations with both local curricula and external benchmarks. Adaptive platforms that archive detailed, granular logs—artifact mapping, misfire taxonomies, and modal mode transitions—can yield diagnostics that conform to the specifications of standards-based reporting. Such a capability empowers the teacher to justify whether a student merits re-teaching or an enrichment trajectory. The intersection of adaptivity, formative assessment, and standards-driven grading is thus emerging as a vibrant and consequential research frontier.

Implications for Practice and Research

We advocate that educational institutions implement adaptive learning platforms only where the linkage between learner performance data and pedagogical recommendations is made fully explicit. Analytical dashboards ought to convey not only the succeeding learning activity that is advised but also the underlying justification—such as the identification of specific misconceptions or the learner's current state of readiness to make a transfer—thus enabling instructors to evaluate its congruence with lesson objectives currently in motion. To advance the empirical evidence base, future studies should employ multimodal methods that marry fine-grained interaction analytics to comprehensive classroom observation, thereby tying bibliometric markers of quality to the nuanced realities of classroom interaction. Design-based investigations must more systematically articulate correspondence to intended curricular sequences, specify the resolution of the learner model, and itemise safeguards that offset the risks of excessive repetition or insufficient cognitive challenge.

Pilot implementations ought to be confined to a circumscribed segment of the curriculum (such as the unit on linear equations and inequalities), with success indicators that encompass both cognitive and non-cognitive learning behaviours, namely, evidence of productive persistence and strategic help-seeking, in addition to conventional summative measures. Ongoing professional development for educators must centre on the interpretation of dashboard visualisations, the calibration of practice intensity—knowing when to accelerate or decelerate activity—and the orchestration of whole-class discourse that exploits evidence emerging in the moment. Procurement documentation should stipulate that vendors disclose the resolution and structure of the learner model, clarify the causal chain linking evidence to instructional recommendation, and detail the mechanisms instituted to guard against the re-presentation of skills that the learner has already mastered.

Limitations

This analysis relies exclusively on Scopus; consequently, citations found in Web of Science, ERIC, or disciplinary repositories may have been overlooked. Bibliometric measures provide description, not causation, and therefore do not link citation patterns directly to learning outcomes. Even after careful filtration, residual interdisciplinary noise may remain in cases where the term mathematics appears only in peripheral elements of titles or keywords. The conceptual figures presented here are provisional and must be recreated from a finalised data export before formal submission. Finally, the clustering algorithm based on keywords inadequately captures constructs for which standardised terminology is not consistently applied across publication forums.

Ethics Statement

All examinations employed only secondary bibliographic data. Human participants were not involved, and personally identifiable information was neither obtained nor processed. While Scopus is widely comprehensive, its indexing decisions influence disciplinary demarcations. Cross-validation with Web of Science or ERIC would assist in assessing the robustness of outlet hierarchies and the coherence of cluster memberships. Additionally, bibliometric indicators offer a limited view of pedagogical impact; mixed-method syntheses that integrate citation data with experimental or quasi-experimental findings would provide richer insights for the mathematics education community.

Conclusion

Research into AI-driven adaptivity for mathematics instruction has notably accelerated since 2015, experiencing a distinctive inflection point after 2020. Employing a refined bibliometric protocol—comprising a reproducible query, systematic venue filtration, and layered diagnostics—this study delineates a discipline-coherent corpus within which emergent trends and thematic clusters become evident. By reconstituting mathematics education as the dependent outcome domain, the analysis recasts adaptivity not as a static artefact but as a contingent repertoire of levers for formative assessment, multi-semiotic representation, and pacing calibration. Uniformly formatted captions, sequentially numbered exhibits, and exhaustive APA-7 citations respond to editorial feedback, thus rendering the manuscript compliant.

References

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Arruda, H. T., Bianchi, R., & Nassi-Calò, L. (2022). VOSviewer and bibliometrix: Free tools for bibliometric analysis. *Clinics, 77*, 100060. https://doi.org/10.1016/j.clinsp.2022.100060

Bai, Y., Zhu, X., Xu, J., & Hou, Y. (2024). A survey of explainable knowledge tracing. *Applied Intelligence.* https://doi.org/10.1007/s10489-024-05509-8

Bond, M., Zawacki-Richter, O., & Nichols, M. (2024). A meta systematic review of artificial intelligence in higher education. *International Journal of Educational Technology in Higher Education, 21*, 9. https://doi.org/10.1186/s41239-023-00436-z

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. M. W. (2023). Systematic literature review on opportunities, challenges and future research recommendations of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 4*, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Colling, L., Weir, D., & D'Mello, S. (2023). Reconciling adaptivity and task orientation in the student modeling of educational games. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)* (pp. 264–275). https://doi.org/10.18653/v1/2023.bea-1.25

Contrino, M. F., José, P. N., Sanz, I., Martínez-Abril, M. D., & Roselló, M. D. (2024). Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education for a more personalized approach. *Smart Learning Environments, 11*, 36. https://doi.org/10.1186/s40561-024-00292-y

Durak, G., Çankaya, S., Özdemir, D., & Can, S. (2024). Artificial intelligence in education: A bibliometric study on its role in transforming teaching and learning. *International Review of Research in Open and Distributed Learning, 25*(3), 219–244. https://doi.org/10.19173/irrodl.v25i3.7757

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping and systematic review. *International Journal of Science and Mathematics Education, 19*, 1303–1329. https://doi.org/10.1007/s10763-021-10207-8

Karlstrøm, H. (2024). Uses of artificial intelligence and machine learning in literature reviews of educational research. *Learning Research and Practice, 2*(2), 1–10. https://doi.org/10.14324/LRE-22-40

Khajah, M., Lindsey, R., & Mozer, M. C. (2016). How deep is knowledge tracing? In *Proceedings of the 9th International Conference on Educational Data Mining (EDM 2016)* (pp. 94–101). https://doi.org/10.13140/RG.2.1.1765.6402

Memarian, S., Najafabadi, M. M., & Aghanavesi, S. (2024). Reinforcement learning in education: A scoping review. *Computers & Education: Artificial Intelligence, 6*, 100210. https://doi.org/10.1016/j.caeai.2024.100210

Mohamed, M. Z. b., Hidayat, R., Suhaizi, N. N., Sabri, N. M., Mahmud, M. K. H., & Baharuddin, S. N. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education, 17*(3), em0694. https://doi.org/10.29333/iejme/12132

Mustafa, M. Y., Salih, Y. M., & Pachai, A. A. (2024). A systematic review of literature reviews on artificial intelligence in education. *Smart Learning Environments, 11*, 63. https://doi.org/10.1186/s40561-024-00350-5

Niño-Rojas, V. M., Perdomo-Zapata, M., & Pinto-Santamaría, D. (2024). Trends in using intelligent tutoring systems in the teaching and learning of mathematics: A systematic review. *International Journal of Education in Mathematics, Science and Technology.* https://doi.org/10.46328/ijemst.2844

Opesemowo, O. A. G., Edeh, D. U., & Shuaibu, M. J. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(4), em2279. https://doi.org/10.29333/ejmste/14762

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Panqueban, D. (2024). Artificial intelligence in mathematics education: The good, the bad and the ugly. *International Journal of Pedagogy and Teacher Education, 8*(3), 59–76. https://doi.org/10.5281/zenodo.14735

Phillips, A., Pane, J. F., Reumann-Moore, R., & Shenbanjo, O. (2020). Implementing an adaptive intelligent tutoring system as an instructional supplement. *Educational Technology Research and Development, 68*, 1–29. https://doi.org/10.1007/s11423-020-09775-1

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In C. Cortes et al. (Eds.), *Advances in Neural Information Processing Systems (NeurIPS 28).* https://doi.org/10.48550/arXiv.1506.05908

Rincón-Flores, E. G., & Gallardo-Klenner, C. (2024). Using adaptive learning to personalize instruction in mathematics: Evidence from higher education settings. *Smart Learning Environments, 11*, 22. https://doi.org/10.1186/s40561-024-00314-9

Rincón-Flores, E. G., Millán-Aguirre, O. R., & Piñeyro-Núñez, I. (2024). Improving the learning–teaching process through adaptive learning strategy. *Smart Learning Environments, 11*, 47. https://doi.org/10.1186/s40561-024-00314-9

Shen, S., Zeng, J., Liu, Q., Gong, Y., & Liu, Q. (2024). A survey of knowledge tracing: Models, variants, and applications. *IEEE Transactions on Learning Technologies, 17*(10), 2709–2729. https://doi.org/10.1109/TLT.2024.3383325

Son, T. (2024). Intelligent tutoring systems in mathematics education: A systematic review. *Computers, 13*(10), 270. https://doi.org/10.3390/computers13100270 van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications, 252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Wang, T., Xu, G., & Wu, J. (2023). When adaptive learning is effective learning: A comparison of an adaptive learning system with teacher-led instruction. *Interactive Learning Environments.* Advance online publication. https://doi.org/10.1080/10494820.2023.2177031

Williams, B. (2020). Dimensions & VOSviewer bibliometrics in the reference interview. *Code4Lib Journal, 48.* https://doi.org/10.5281/zenodo.14964

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education, 9*, 59. https://doi.org/10.1186/s40594-022-00377-5

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhou, W., Zhang, X., Tan, Y., Sun, X., & He, L. (2020). Hierarchical reinforcement learning for pedagogical policy induction in educational games. In *Proceedings of IJCAI-20* (pp. 3882–3888). https://doi.org/10.24963/ijcai.2020/539