JURNAL PEMBELAJARAN DAN MATEMATIKA SIGMA (JPMS)
Vol. 12, No. 2 (2026)
Page : 212 – 220
Doi: https://doi.org/10.36987/jpms.v12i1.10036

ISSN : 2460-593X
E-ISSN : 2685- 5585

DIGITAL ORCHESTRATION OF DIVERSE STUDENT RESPONSES TO OPEN-ENDED MATHEMATICAL TASKS: A SYSTEMATIC REVIEW

VALERIA YEKTI KWASANING GUSTI[1*]

[1*]Program Studi Pendidikan Matematika, FKIP, Universitas Terbuka, Tangerang Selatan, Indonesia

email:
[1*]Corresponding Author: valeria.gusti@eampus.ut.ac.id

Article Info

Article history:
Received Oct 20, 2026
Revised Nov 02, 2026
Accepted Nov 13, 2026

Keywords:
open-ended mathematics
student responses
digital technology
instrumental orchestration
formative assessment

ABSTRACT

This systematic review examines how digital technologies mediate the elicitation and orchestration of diverse student responses to open-ended mathematical tasks, with particular attention to how response diversity becomes visible, organized, selected, compared, connected, and used in mathematical discussion. The review followed PRISMA 2020 reporting principles. An Elicit-assisted search identified 1,000 records, of which 765 passed title-and-abstract screening. Full-text assessment was available for 376 records, resulting in 265 publications in the frozen-protocol corpus. These publications were classified according to their coverage of mathematical openness, digital mediation, and pedagogical orchestration. The principal analytic corpus comprised 29 publications classified as Core and Strong Partial, representing the most relevant evidence for the review's central research questions. Data were extracted and coded according to the functions of the student-response workflow, including collection, access, analysis, categorization, selection, display, comparison, connecting, discussion, justification, and feedback. Discussion/justification was explicitly documented in 28 of the 29 principal analytic publications, while collection/access appeared explicitly in 20, public display in 22, and comparison/connecting in 20. Teacher selection was explicit in eight publications, whereas automated analysis was explicit in five. These findings indicate that digital technologies primarily support the visibility and organization of diverse student responses, while teachers retain a central role in interpreting, selecting, sequencing, and connecting mathematical contributions. The review identifies hybrid teacher-technology orchestration as an important direction for developing mathematically productive classroom discussions.

Published by LPPM Universitas Labuhanbatu. This is an open-access article under the **Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY - NC - SA 4.0)**

1. INTRODUCTION

Open-ended mathematical tasks can reveal a wider range of student thinking than is usually visible in conventional whole-class exchanges [18]. Students may produce different valid answers, solution strategies, representations, examples, constructions, conjectures, or justifications. This diversity is instructionally valuable because it provides material for comparison, generalization, and connection. It also creates an orchestration problem: teachers must notice relevant differences across many responses, decide which contributions warrant public attention, and use them coherently within the limited time available for classroom discussion.

Digital technologies change both the scale and timing of this work. Connected-classroom systems, interactive presentations, shared boards, learning-management environments, dynamic mathematics tools, automated assessment systems, and learning-analytics dashboards can collect student work, preserve it for later inspection, aggregate responses, display them publicly, and, in some cases, organize them according to mathematical characteristics. Research on Padlet-supported discussion,

interactive-presentation environments, connected classrooms, and learning analytics shows that digital tools can make a wider range of student responses available to teachers and peers [5], [8], [10], [14].

Greater visibility, however, is not equivalent to orchestration. In Padlet-supported discussions, the platform preserves student posts and comments, but teacher mediation determines how those contributions enter mathematical discussion [8]. Connected-classroom studies show that teachers can monitor, select, and sequence computer-based responses, although these actions compete with the simultaneous demands of supporting students during exploration [5], [10]. Learning-analytics systems can reduce the burden of scanning a large response set, but the teacher must still judge which differences are mathematically significant and how they should be used.

Computational support is also becoming more sophisticated. STEP learning analytics can analyze example-eliciting tasks automatically and allow teachers to filter submissions according to mathematical characteristics [14], [15]. STACK can automatically evaluate learner-generated mathematical examples at scale [12], while pen-based systems have been used to apply machine analysis to students' hand-drawn mathematical representations [13]. More recent studies of AI-supported mathematics classrooms report orchestration practices such as dashboard monitoring, AI correction, and AI mediation [11], [16]. These developments make it necessary to examine the division of work between technological analysis and teacher judgment.

This review conceptualizes Digital Orchestration as a hybrid process in which teachers and digital technologies perform complementary roles in organizing, interpreting, and utilizing diverse student mathematical responses. Rather than viewing orchestration as an autonomous technological function or an exclusively teacher-directed activity, this review emphasizes the interaction between technological affordances and pedagogical decision-making. Digital technologies support the collection, preservation, analysis, categorization, filtering, and display of student responses, while teachers retain responsibility for interpreting their mathematical significance, selecting relevant contributions, sequencing responses, establishing connections, and facilitating mathematical discussion. Students also contribute to this process by generating, explaining, comparing, and justifying their mathematical ideas. Accordingly, hybrid digital orchestration in this review refers to the coordinated division of labour among technology, teachers, and students in transforming individually generated mathematical responses into collective mathematical reasoning.

The phenomenon can be illustrated with a simple task: "The answer is 9. What mathematical equation can produce 9?" Responses such as 5+4, 3×3, 18÷2, 10−1, and 2^3+1 are all valid, but they differ mathematically. The instructional value lies in what can be done with those differences. A digital environment may make the responses visible, while the teacher may use them to compare operations, representations, efficiency, structure, or generality. The unit of analysis in this review is therefore the relationship among technology, the mathematical task, the response set, and its classroom use.

The review addresses six questions: (1) What are the publication characteristics and study contexts of research on digitally mediated orchestration of diverse mathematical responses? (2) What forms of mathematical openness and response diversity are represented? (3) How do digital technologies mediate the collection, access, recognition, organization, selection, display, and comparison of student responses? (4) How do teachers orchestrate digitally mediated response sets through noticing, selection, sequencing, connecting, questioning, and instructional adaptation? (5) How are digitally mediated responses used to support student mathematical interaction and discussion? (6) What division of labour emerges among technology, teachers, and students, and what research agenda follows from the evidence?

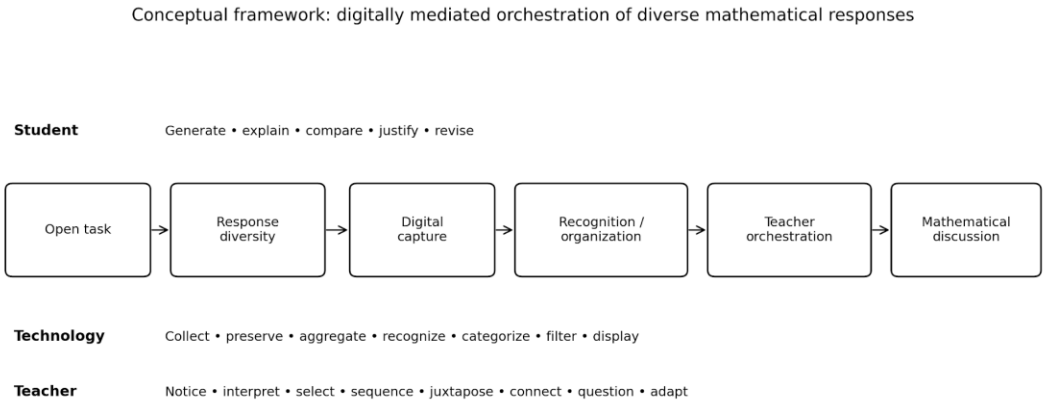


Figure 1. Conceptual framework for digitally mediated orchestration of diverse mathematical responses.

2. RESEARCH METHOD

The review followed PRISMA 2020 reporting principles [17]. The process combined Elicit-assisted retrieval and screening with a frozen full-text protocol and independent verification of studies that were most important to the synthesis. Because the target phenomenon is described differently across mathematics education, educational technology, formative assessment, learning analytics, and classroom-orchestration research, the search and abstract-screening stages were designed to favor recall. Precision was introduced progressively through full-text eligibility criteria and workflow-based coding.

The initial search combined terms referring to digital technology, teachers or instruction, student responses, open-endedness, mathematics, tasks or problems, and elicitation, display, or orchestration. The working Boolean formulation was: (“digital technology” OR “digital technologies” OR technology OR “educational technology”) AND (teacher OR teachers OR educator OR instruction) AND (“student response” OR “student responses” OR “multiple responses”) AND (“open-ended” OR “open ended”) AND (mathematical OR mathematics OR math) AND (task OR tasks OR problem OR problems) AND (elicit OR display OR orchestrate). Screening was based on the underlying phenomenon rather than exact terminology. A study could therefore qualify when it referred to student work, solutions, strategies, representations, examples, constructions, dashboards, sharing, formative assessment, or connected classrooms without using the phrase “student response.” The review focused on literature published from 2016 onward and was updated through 2026.

Eligibility was operationalized across five dimensions. Mathematics education and student-generated mathematical work were required. Response diversity or openness was required for Core and Strong Partial classifications but could be partial or unclear for Fragment evidence. Digital mediation was required and included collecting, accessing, analyzing, categorizing, validating, filtering, selecting, displaying, sharing, or organizing student responses. Pedagogical orchestration was not required for Fragment evidence but was central to Core and Strong Partial classifications. Multiple-choice variation was not treated as response diversity merely because students selected different options.

The Elicit workflow identified 1,000 records. Title-and-abstract screening retained 765 records and excluded 235. Full text was available within the Elicit workflow for 376 retained records; 389 were unavailable at that stage. Of the 376 records assessed at full text, 265 passed the eligibility screen and 111 failed. The 389 records without available full text are therefore reported as retrieval-unavailable rather than as eligibility exclusions. The 265 surviving publications entered the frozen-protocol corpus, after which synthesis-critical publications were independently checked against substantive or full-text sources.

Following the eligibility assessment, the 265 publications in the frozen-protocol corpus were classified into five analytical categories: Core, Strong Partial, Fragment, Adjacent/Contextual, and Exclude. The classification was based on the extent to which each publication addressed five dimensions: (1) mathematics education context, (2) student-generated mathematical work, (3) mathematical openness or response diversity, (4) digital mediation of student responses, and (5) pedagogical orchestration. Core publications provided direct evidence of diverse or open mathematical responses that were digitally mediated and meaningfully incorporated into classroom orchestration. Strong Partial publications addressed several major dimensions but provided incomplete evidence for at least one central component. Fragment publications contributed evidence on specific elements of the response workflow, such as automated recognition, analytics, response collection, or feedback, without sufficiently documenting the complete orchestration process. Adjacent/Contextual publications informed the theoretical and conceptual background but did not provide direct evidence for the central synthesis. Publications classified as Exclude did not satisfy the applicable eligibility requirements.

The principal analytic corpus was subsequently constructed by combining publications classified as Core and Strong Partial, resulting in 29 publications. These publications were selected because they provided the most relevant evidence for examining the relationship between digital mediation, teacher orchestration, student interaction, and the use of diverse mathematical responses in classroom discussions. Fragment and Adjacent/Contextual publications were retained only to support specific aspects of the conceptual discussion and were not counted as primary evidence in the principal analytic corpus.

Data extraction followed the response workflow: task/openness; generation; collection/access; recognition/analysis; categorization/aggregation/filtering; selection; display; sequencing; juxtaposition/comparison; connecting; discussion/justification; and feedback/adaptation. Each function was coded as Yes, Partial, No, or Unclear and, where possible, attributed to the teacher, technology, student, or a teacher-technology hybrid. Openness was coded separately as multiple valid final answers, multiple solution strategies, multiple representations, different valid examples or constructions, or exploratory/conjectural responses.

Verification status was used conservatively. A publication was described as verified only when a substantive or fully accessible article had been independently inspected; records supported only by Elicit screening evidence remained provisional. Publication families were tracked without assuming that papers using the same platform, project, or author group represented the same empirical study. Publications were considered possible duplicates only when there was evidence of overlap in participants, tasks or interventions, data collection, and empirical episodes.

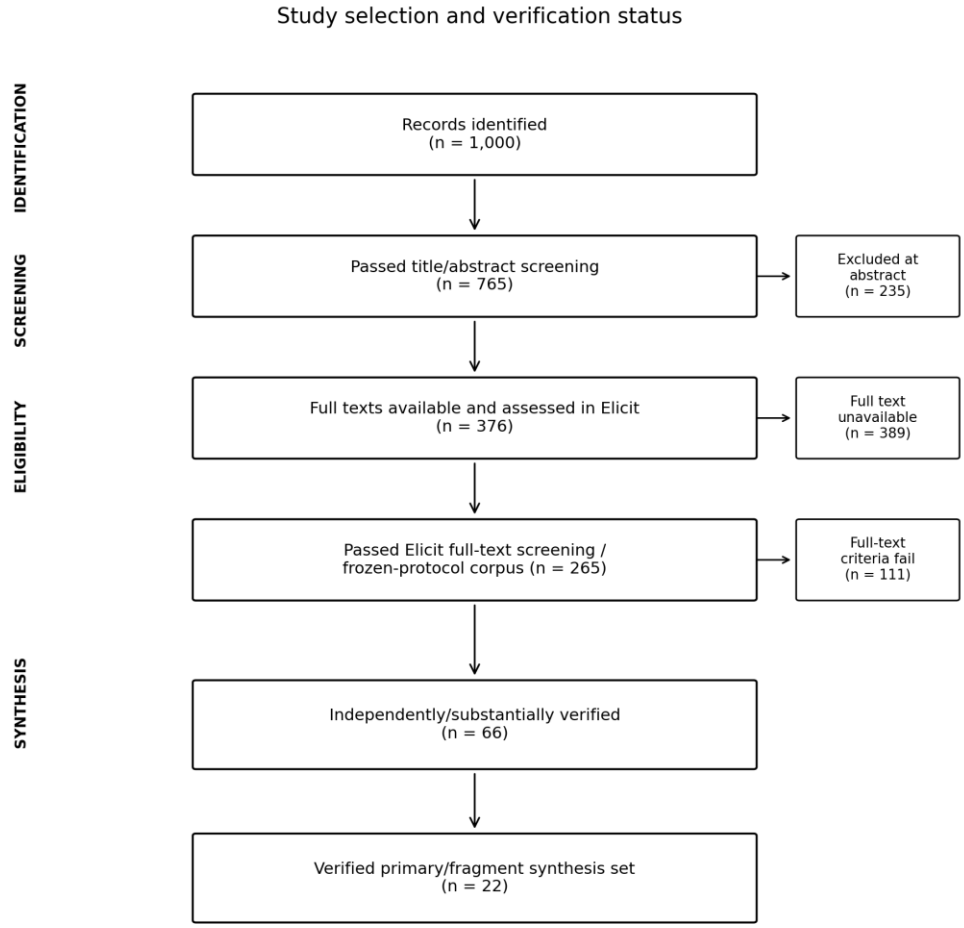


Figure 2. PRISMA-style flow of identification, screening, full-text assessment, and construction of the analytic corpus.

3. RESEARCH RESULTS AND DISCUSSION

3.1 Research result

The principal analytic corpus comprised 29 Core and Strong Partial publications spanning 2016–2026. Publication activity increased after 2021: six publications appeared in 2022, five in 2024, four in 2025, and three in 2026. The studies were distributed across mathematics-education, educational-technology, formative-assessment, learning-environment, and conference outlets. Digital Experiences in Mathematics Education was the most frequent venue, with six publications. Most studies were conducted in secondary-school settings, although elementary, middle-school, and higher-education contexts were also represented.

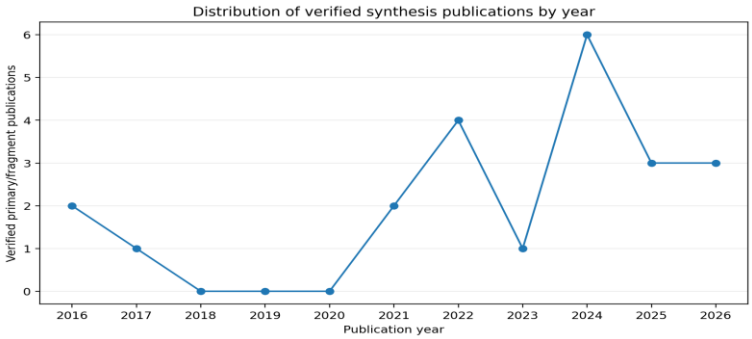


Figure 3. Publication-year distribution of the verified primary and fragment evidence base

The technologies varied substantially and included Padlet, Nearpod, connected-classroom digital worksheets and interactive whiteboards, GeoGebra, Jamboard and Google Slides, online homework environments, the FACT orchestration system, and STEP learning analytics. Because of this heterogeneity, the synthesis compared workflow functions rather than treating the review as a platform comparison.

Forms of mathematical openness

Multiple solution strategies and exploratory or conjectural responses were the most common forms of response diversity, each appearing in 14 publications. Multiple representations appeared in 13 publications, different valid examples or constructions in five, and multiple valid final answers in two. The categories were not mutually exclusive because a single task could support more than one form of openness. These results show that open-ended response should not be treated as a single category. Some tasks invite different correct answers, whereas others elicit different strategies, representations, examples, or conjectures. Each form of openness creates different demands for technological analysis and different opportunities for classroom comparison.

Digital mediation of student responses

Digital mediation was most consistently documented in functions that made student work available for collective use. Collection or access was explicit in 20 of the 29 publications and partial in the remaining nine. Public display was explicit in 22 publications, comparison or connecting in 20, and discussion or justification in 28. Teacher selection was less consistently reported: eight publications provided explicit evidence and six provided partial evidence. Automated analysis was explicit in five publications and partial in six.

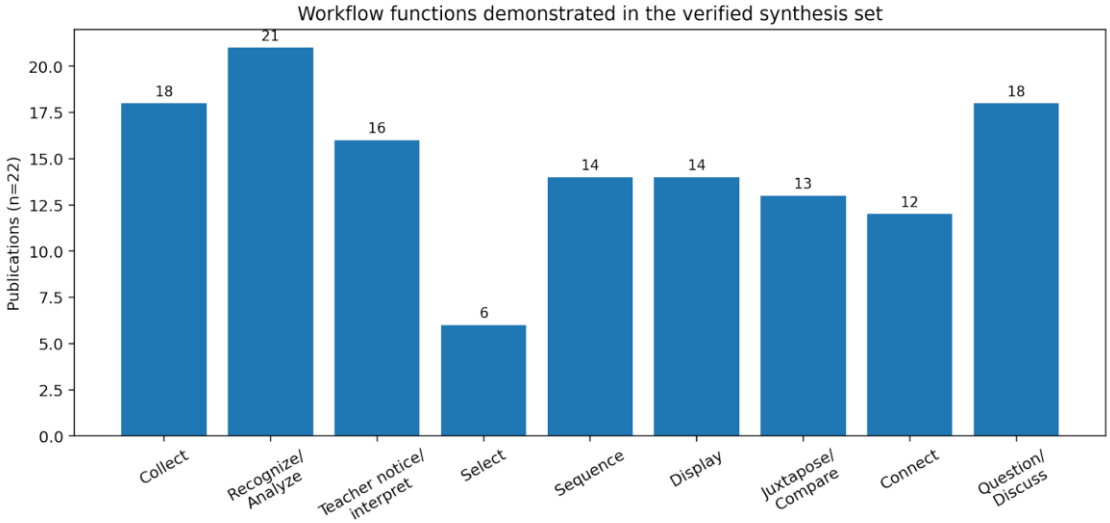


Figure 4. Frequency of digitally mediated response-workflow functions in the verified synthesis set.

Shared platforms often functioned as persistent response spaces. Padlet studies allowed student solutions to remain available for later reading, commenting, retrieval, and comparison [8], [9]. Connected-

classroom and interactive-presentation systems provided similar visibility while giving teachers more direct control over which responses entered public discussion [5], [10].

A smaller body of research addressed computational interpretation of responses. STEP analyzed example-eliciting tasks automatically and represented response characteristics through tables, grids, histograms, and Venn diagrams, which teachers could filter and inspect [14], [15]. Other systems demonstrated automated checking of learner-generated examples or machine analysis of pen-based mathematical representations [12], [13]. These studies establish that open mathematical work can be interpreted computationally, although not all of them document whole-class orchestration.

Teacher orchestration and mathematical discussion

Teacher activity became most visible after a response set was accessible. Across the corpus, teachers monitored student work, interpreted mathematical differences, decided which contributions to foreground, managed transitions between individual, group, and whole-class activity, and used displayed responses to question students or adapt instruction. Instrumental-orchestration studies described practices such as Spot-and-show, Discuss-the-screen, Explain-the-screen, Sherpa-at-work, and Guide-to-investigate [3], [11], [16].

Selection was especially useful for distinguishing visibility from orchestration. Chen et al. [3] described Spot-and-show as the selection and display of examples representing different solution approaches. Other connected-classroom research similarly documented teachers surveying response sets, selecting student work, sequencing selected contributions, and displaying responses for discussion [5], [10]. Sequencing and connecting were harder to establish because some studies documented ordered presentation or transitions without making the mathematical rationale for the order explicit.

Discussion and justification were the most consistently represented downstream functions: 28 of the 29 publications contained explicit evidence. Padlet studies moved from individual posting to peer comments and teacher-mediated classroom talk; FlexTech combined several digital tools to support collective reasoning during pattern generalization [7]; and dialogical environments supported mathematical argumentation [1]. Digital participation nevertheless did not guarantee productive interaction. The FACT orchestration system did not produce greater collaboration or productive struggle than paper-based classes across its development trials, showing that technical orchestration features must be evaluated through the quality of interaction and learning [20].

Division of labour between technology, teachers, and students

Across studies, the most consistent pattern was a division of labour rather than replacement of the teacher. Technology most often captured, preserved, aggregated, recognized, categorized, filtered, or displayed student work. Teachers remained responsible for noticing, interpretation, selection, sequencing, connecting, questioning, and instructional adaptation. Students generated the mathematical response space and participated in explanation, comparison, critique, justification, generalization, and revision.

STEP provided the clearest evidence of hybrid orchestration. The system automatically analyzed responses to example-eliciting tasks and organized them in interactive reports; teachers then inspected distributions, filtered the response set, opened individual submissions, and used patterns or selected examples in formative assessment and classroom discussion [14], [15]. AI-supported mathematics classrooms extended this division of labour through dashboard monitoring, AI correction, and AI mediation while retaining teacher responsibility for interpretation and coordination [11], [16].

3.2 Discussion

The evidence can be interpreted as three related capability layers. The first is visibility: technologies collect, preserve, and display student work. The second is organization: dashboards, filters, grouping mechanisms, and machine analysis make heterogeneous responses easier to inspect. The third is orchestration: teachers use selected or organized responses to develop mathematical ideas through comparison, questioning, connecting, and discussion. Evidence is extensive for visibility, growing for organization, and less consistent for integrated orchestration, particularly when all three layers are examined within the same empirical classroom study.

This distinction clarifies the contribution of the review. The literature does not support a simple claim that digital technologies lack the capacity to select, sequence, compare, or analyze student responses. Connected-classroom technologies, interactive presentations, Padlet environments, learning analytics, and AI-supported tools already provide parts of this functionality [5], [8], [10], [11], [14]. The more defensible gap concerns empirical enactment: there is limited evidence on how teachers use automatically characterized or curated sets of diverse student mathematical responses to enact purposeful selection, sequencing, connecting, and mathematically productive whole-class discourse.

The finding also refines the relationship between technological affordance and teacher agency. Automated checking or characteristic detection can reduce the time required to scan responses [12], [13], [14], but pedagogical importance remains context-dependent. An unusual response may matter because it

reveals a misconception, expresses a sophisticated generalization, provides a productive representation, or bridges two strategies. Computational description therefore becomes instructionally useful only when it is interpreted in relation to the mathematical goals of the lesson.

For technology design, a class response set should be treated as an object for mathematical inquiry rather than as a list of submissions. Useful affordances include rapid overview, mathematically meaningful clustering, transparent descriptors, filters for contrasting cases, previews of individual work, persistence of earlier responses, teacher-controlled public display, flexible reordering, and movement between aggregate patterns and individual submissions. Where systems categorize or recommend responses, the basis for those decisions should remain visible to teachers.

For teacher education, response-space orchestration should be addressed as a specific form of professional competence rather than as generic digital competence. Teachers need opportunities to anticipate likely response diversity, interpret dashboards critically, distinguish system classifications from mathematical judgments, select contrasting cases, plan purposeful sequences, formulate connecting questions, and use errors or unconventional responses productively. Instrumental orchestration and the Five Practices provide complementary structures for this work, while research on expert roles in digitally supported formative-assessment discussions further emphasizes the interpretive work required of teachers [4].

Future research should prioritize hybrid teacher-technology orchestration, explainable mathematical response characterization, pedagogical selection and sequencing, mathematical connecting, discourse quality, teacher noticing, equity and participation, long-term learning outcomes, interoperability across classroom tools, and the emerging role of generative AI in response curation. Stronger studies should connect system logs and screen activity with classroom video, teacher interviews, student work, and learning outcomes so that technological actions can be related to mathematical consequences.

4. CONCLUSION

Digital mathematics classrooms increasingly face an abundance rather than a scarcity of information about student responses. The central instructional problem is therefore not simply how to collect more responses, but how to make mathematically diverse work usable without reducing that diversity to superficial categories. Across the reviewed studies, technology was strongest at making student work visible and persistent; a smaller set of systems also organized or characterized responses computationally. Teachers nevertheless remained central to deciding what was mathematically important and how selected contributions should be connected through discussion.

The review identifies hybrid teacher-technology orchestration as the most productive direction for future work. The strongest emerging model is one in which technology helps organize the response space while teachers retain responsibility for mathematical interpretation, selection, sequencing, connecting, and questioning. Research should now test how these arrangements affect what teachers notice, which student ideas become public, whose contributions are taken up, the quality of mathematical discourse, and what students learn.

For practice, the findings suggest that schools and teachers should evaluate digital tools not only by whether they collect or display student work, but by whether they support movement from response visibility to mathematical comparison and discussion. For research, the priority is to examine complete classroom workflows in which open tasks generate diverse responses, technology organizes those responses, and teachers use them deliberately to develop collective mathematical reasoning.

5. AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Author 1														

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

6. AI Disclosure Statement

During the preparation of this manuscript, the author used ChatGPT (OpenAI) to support brainstorming and the synthesis of ideas and information. The author critically evaluated, reviewed, and revised the AI-assisted outputs and takes full responsibility for the content, interpretation, and conclusions presented in this manuscript.

REFERENCES

[1] G. Albano, M. Mollo, M. Polo, and G. Marsico, "Dialogical interactions mediated by technology in mathematics education," *Dialogic Pedagogy: A Journal for Studies of Dialogic Education*, vol. 10, pp. DT22–DT40, 2022, doi: 10.5195/dpj.2022.517.

[2] P. Black and D. Wiliam, "Developing the theory of formative assessment," *Educational Assessment, Evaluation and Accountability*, vol. 21, no. 1, pp. 5–31, 2009, doi: 10.1007/s11092-008-9068-5.

[3] M. Chen, R. Bos, M. Doorman, and P. Drijvers, "Exploring instrumental orchestration practices for formative assessment through digital technology," *International Journal for Technology in Mathematics Education*, vol. 32, no. 2, pp. 73–84, 2025, doi: 10.1564/tme_v32.2.02.

[4] A. Cusi and F. Morselli, "The key-roles of the expert during classroom discussions aimed at fostering formative assessment processes through the use of digital technologies," *ZDM–Mathematics Education*, vol. 56, no. 4, pp. 741–755, 2024, doi: 10.1007/s11858-024-01572-0.

[5] A. Cusi, F. Morselli, and C. Sabena, "Promoting formative assessment in a connected classroom environment: Design and implementation of digital resources," *ZDM Mathematics Education*, vol. 49, no. 5, pp. 755–767, 2017, doi: 10.1007/s11858-017-0878-0.

[6] P. Drijvers, M. Doorman, P. Boon, H. Reed, and K. Gravemeijer, "The teacher and the tool: Instrumental orchestrations in the technology-rich mathematics classroom," *Educational Studies in Mathematics*, vol. 75, no. 2, pp. 213–234, 2010, doi: 10.1007/s10649-010-9254-5.

[7] A. Eckert and P. Nilsson, "The emergence of the 'FlexTech' orchestration of inferential reasoning on pattern generalization," *Digital Experiences in Mathematics Education*, vol. 8, no. 1, pp. 1–26, 2022, doi: 10.1007/s40751-021-00098-4.

[8] C. Giberti, F. Arzarello, S. Beltramino, and G. Bolondi, "Mathematical discussion in classrooms as a technologically-supported activity fostering participation and inclusion," *Educational Studies in Mathematics*, vol. 118, no. 2, pp. 201–228, 2025, doi: 10.1007/s10649-024-10356-y.

[9] C. Giberti, F. Arzarello, G. Bolondi, and H. Demo, "Exploring students' mathematical discussions in a multi-level hybrid learning environment," *ZDM–Mathematics Education*, vol. 54, no. 2, pp. 403–418, 2022, doi: 10.1007/s11858-022-01364-4.

[10] A. Haj-Yahya and S. Aegbaria, "Using interactive presentations to promote mathematical discourse," *European Journal of Mathematics and Science Education*, vol. 4, no. 1, pp. 1–17, 2023, doi: 10.12973/ejmse.4.1.1.

[11] J. Kim and C. Han, "How are AI-based mathematics classrooms orchestrated? An analysis of instrumental orchestration strategies in elementary lessons," *The Mathematical Education*, vol. 64, no. 4, pp. 671–690, 2025, doi: 10.63311/mathedu.25.64412.

[12] G. Kinnear, P. Iannone, and B. Davies, "Student approaches to generating mathematical examples: Comparing e-assessment and paper-based tasks," *Educational Studies in Mathematics*, vol. 119, no. 2, pp. 179–201, 2025, doi: 10.1007/s10649-024-10361-1.

[13] K. Koile, A. Rubin, S. Chapman, M. Kliman, and L. Ko, "Using machine analysis to make elementary students' mathematical thinking visible," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. Association for Computing Machinery, 2016, pp. 524–525, doi: 10.1145/2883851.2883922.

[14] K. Moed-Abu Raya and S. Olsher, "Teachers' formative assessment practices in their mathematics classroom using learning analytics visualizations," *Digital Experiences in Mathematics Education*, vol. 10, no. 3, pp. 395–417, 2024, doi: 10.1007/s40751-024-00148-7.

[15] K. Moed-Abu Raya and S. Olsher, "From data to insight: Teachers' noticing of mathematical thinking through learning analytics," *Journal of Mathematics Teacher Education*, early access, 2026, doi: 10.1007/s10857-026-09755-0.

[16] Y.-K. Noh and H.-J. Kim, "Exploring teachers' instrumental orchestration and roles in high school mathematics classes using AI-based digital tools," *The Mathematical Education*, vol. 65, no. 2, pp. 309–335, 2026, doi: 10.63311/mathedu.26.6525.

[17] M. J. Page *et al*., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Art. no. n71, 2021, doi: 10.1136/bmj.n71.

[18] M. S. Smith and M. K. Stein, *5 Practices for Orchestrating Productive Mathematics Discussions*. Reston, VA, USA: National Council of Teachers of Mathematics, 2011.

[19] L. Trouche, “Managing the complexity of human/machine interactions in computerized learning environments: Guiding students’ command process through instrumental orchestrations,” *International Journal of Computers for Mathematical Learning*, vol. 9, no. 3, pp. 281–307, 2004, doi: 10.1007/s10758-004-3468-5.

[20] K. VanLehn, H. Burkhardt, S. Cheema, S. Kang, D. Pead, A. Schoenfeld, and J. Wetzel, “Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms?” *Interactive Learning Environments*, vol. 29, no. 6, pp. 987–1005, 2021, doi: 10.1080/10494820.2019.1616567.