Jurnal Ilmiah Profesi Pendidikan
Volume 11, Nomor 2, Mei 2026
ISSN (Print): 2502-7069; ISSN (*Online*): 2620-8326

Reconceptualizing Assessment For Learning In Mathematics Education: A Systematic Literature Review Of Digital, Formative, And Student-Centered Practices

Melkior Wewe[1]*, I Gusti L. A Parwata[2], I Dewa A. M Budhyani[2], I Gede Ratnya[2]
[1]Pendidikan Matematika, STKIP Citra Bakti, Indonesia
[2]Universitas Pendidikan Ganesha,Singaraja, Indonesia
*Corresponding Author: melkiorwewe1@gmail.com

Article History
Received : April 07th, 2026
Revised : April 27th, 2026
Accepted : May 16th, 2026

Abstract: Assessment for Learning (AfL) is an important approach in mathematics education because it uses assessment evidence to support learning, not only to measure achievement. In primary and secondary mathematics classrooms, AfL is increasingly connected with digital tools, formative feedback, and student-centered practices. This study aims to reconceptualize AfL in mathematics education at the primary and secondary levels through a systematic literature review focusing on digital, formative, and student-centered practices. This study employed a Systematic Literature Review (SLR) approach guided by PRISMA 2020. A total of 150 articles were identified, and after the selection, screening, and eligibility process, 10 articles were included in the final synthesis. The data were analyzed using thematic synthesis through coding, theme development, and analytical interpretation. The findings show that AfL in mathematics education has developed from conventional formative assessment into a more diagnostic, adaptive, and participatory assessment ecosystem. Three main themes emerged from the synthesis. First, digital assessment supports faster feedback, personalized learning evidence, and better documentation of students' progress. Second, formative assessment emphasizes feedback quality, evidence interpretation, and instructional follow-up. Third, student-centered assessment strengthens AfL when students understand criteria, use feedback, reflect on learning, and regulate their improvement. This study concludes that effective AfL in mathematics education should be pedagogically formative, digitally supported, and activated through students' active participation.

Keywords: assessment for learning; digital assessment; formative assessment; mathematics education; student-centered assessment.

INTRODUCTION

In the Society 5.0 era, assessment is no longer understood merely as an instrument for measuring the final outcomes of a learning process; rather, it is an inseparable part of the learning process itself, supporting continuous pedagogical decision-making. This paradigm shift positions assessment as a means of determining students' position and activity in learning, identifying undesirable problems that may prevent the achievement of learning goals, and determining the follow-up actions needed throughout the learning process based on reflection (Sadler, 1989; Black & Wiliam, 1998; Black & Wiliam, 2009; Bennett, 2011). Within this framework, Assessment for Learning (AfL) has developed as an approach that emphasizes that assessment must function as a mechanism for learning, not merely as a procedure for accountability of outcomes (Sadler, 1989; Black & Wiliam, 1998; Ruiz-Primo, 2011). Therefore, feedback in Assessment for Learning (AfL) can be understood as constructive comments on students' work that help narrow the gap between current performance and the intended learning goals (Hattie & Timperley, 2007; Wisniewski et al., 2020). The strength of Assessment for Learning (AfL) lies in its connection to learning objectives, the quality of the evidence collected, and the capacity of both teachers and students to use that evidence to improve learning (Kingston & Nash, 2011; Bennett, 2011; Schildkamp et al., 2020).

In mathematics learning, Assessment for Learning (AfL) is highly important for both teachers and students to implement effectively, because success in mathematics learning is not limited to determining whether the final answer is right or wrong, but extends to examining the quality of students' thinking processes in solving mathematical problems and critically and consciously reflecting on the strengths and weaknesses of their problem-solving processes (Maskos et al., 2025; Rakoczy et al., 2019).

Mathematics learning requires teachers to continuously capture students' thinking processes so that instruction can be truly responsive to their learning needs. A number of studies show that formative assessment in mathematics can improve the quality of classroom interaction, enhance teachers' responses to students' mathematical thinking, and positively affect learning achievement, learning interest, perceptions of the usefulness of feedback, and students' self-efficacy (van den Berg et al., 2016; Kim, 2019; van den Berg et al., 2018; Rakoczy et al., 2019; Smit et al., 2023). In addition, the positive relationship between formative assessment and summative assessment in mathematics learning indicates that AfL is not a peripheral additional element, but an essential part of an effective mathematics learning ecosystem (Gezer et al., 2021; Kyaruzi et al., 2019). Thus, AfL deserves to be positioned as a core instructional strategy that integrally connects learning goals, learning activities, and evidence of student learning.

The development of digital technology has further expanded and transformed AfL practices in mathematics learning. Assessment, which was previously conducted predominantly manually, has shifted toward models that are faster, more adaptive, interactive, and documented in real time (Faber et al., 2017; Barana et al., 2021). Through digital technology, teachers can obtain diagnostic information more quickly, while students can receive immediate feedback and tasks tailored to their individual learning needs (Faber et al., 2017; Rodríguez-Martínez et al., 2023). In addition, digitalization has changed how students understand instructions, revise strategies, and navigate the process of solving mathematical problems more dynamically (Barana et al., 2021; Söderström, 2024). The presence of online formative peer assessment, web-based portfolio assessment, and the integration of teacher-assessment, self-assessment, and peer-assessment indicates that digitalization not only replaces the medium of assessment, but also reconfigures the logic of AfL into one that is more data-driven, personalized, and interactive (Alemdag & Yildirim, 2022; Chang et al., 2012). Therefore, the reconceptualization of AfL in the digital era needs to be understood not merely as a technical change, but as a pedagogical transformation in how evidence of learning is collected, interpreted, and used in mathematics learning.

Beyond digitalization, the development of AfL increasingly underscores the importance of student-centered learning orientation. The literature shows that self-assessment, peer assessment, self-reflection, and students' use of learning data provide broader opportunities for learners to understand learning goals, monitor progress, and take responsibility for their own learning processes (Topping, 2009; Panadero et al., 2017; Double et al., 2020). In mathematics learning, peer assessment is known to strengthen mathematical communication and help students assess the quality of mathematical arguments more explicitly (Patchan et al., 2022). Meta-analyses also show that self-assessment and peer assessment produce meaningful effects on academic performance when positioned as part of the learning process, rather than merely as administrative evaluation procedures (Panadero et al., 2017; Yan et al., 2022; Double et al., 2020). In this context, AfL no longer places students as objects of evaluation, but as active subjects involved in reading evidence of learning, interpreting feedback, and reflectively adjusting their mathematical learning strategies.

Nevertheless, the effectiveness of AfL does not occur automatically. Its implementation depends greatly on teachers' capacity to design meaningful learning, interpret evidence of learning, provide actionable feedback, and adjust instruction based on students' responses (Schildkamp et al., 2020; Kim, 2019). Systematic reviews show that assessment literacy, data literacy, pedagogical competence, and readiness to use evidence of learning are important prerequisites to ensure that AfL does not stop at mere technical procedures (Schildkamp et al., 2020; Lei & Lei, 2026). Teacher professional development programs based on AfL have also been shown to contribute to improved student achievement and metacognition (de Vries et al., 2022; de Vries et al., 2023). In mathematics learning, teachers who are increasingly responsive to students' mathematical thinking tend to be better able to use assessment as an inherent part of instruction, rather than as an activity separate from the learning process (Kim, 2019; Maskos et al., 2025). This indicates that the success of AfL is largely determined by the quality of its implementation, not merely by the availability of instruments or technology.

At the same time, studies on AfL in mathematics learning still show fragmentation, both theoretically and empirically. Many studies discuss AfL only from certain perspectives, such as feedback, self-assessment, peer-assessment, digital assessment, large-scale assessment, or teacher assessment literacy, but few have synthesized AfL as a practice that is simultaneously digital, formative, and student-centered, particularly at the primary and secondary education levels (Maskos et al., 2025; Schildkamp et al., 2020; Puspitaningtyas et al., 2025). Empirically, studies on digital assessment still focus largely on the effectiveness of particular tools and media, while studies on student-centered assessment tend to separate discussions of peer assessment, self-assessment, self-regulation, and formative feedback. In the context of Indonesia and Southeast Asia, research has developed toward the reconstruction of 21st-century classroom assessment strategies, the development of HOTS instruments based on numeracy literacy, gamification, automatic formative assessment, AI integration, and the mapping of assessment for learning, assessment as learning, and assessment of learning; however, this landscape still tends to be fragmented and has not been widely summarized in a single systematic synthesis that fully integrates these three dimensions (Rosidin et al., 2020; Wardani & Siregar, 2023; Barana & Marchisio Conte, 2024; Liu et al., 2024; Ong & Toh, 2024; Baiduri, 2022; Sunarto, 2025; Puspitaningtyas et al., 2025). This condition indicates a strong academic need to build a more integrative, up-to-date, and relevant conceptual framework of AfL for contemporary mathematics learning.

Based on the above description, this study proceeds from the assumption that the reconceptualization of Assessment for Learning in mathematics learning needs to be carried out by bringing together three main dimensions that have often been studied separately, namely digital practices, the formative nature of assessment, and student-centered orientation. The novelty of this study lies in its effort to construct a conceptual synthesis that explicitly brings these three dimensions together within a single, coherent analytical framework, particularly in the context of primary and secondary education. Thus, this study not only maps the development of the AfL literature in mathematics learning, but also offers a more comprehensive conceptual basis for understanding how technology, evidence of learning, student participation, and teachers' pedagogical responses intersect in shaping contemporary mathematics assessment practices.

In line with this, the study is directed toward answering several main questions, namely how the conceptual development of Assessment for Learning in mathematics learning is explained in the literature, how digital AfL practices are implemented in primary and secondary education, how the formative and student-centered characteristics of AfL are constructed in mathematics learning, what factors support and hinder its implementation, and how previous research findings can be synthesized to produce a reconceptualization of AfL that is relevant to the needs of contemporary education. Based on these issues, this study aims to describe the conceptual development of AfL in mathematics learning, analyze forms of digital AfL practices, identify the characteristics of formative and student-centered AfL, explain the supporting and inhibiting factors in its implementation, and construct a conceptual synthesis as a form of reconceptualizing Assessment for Learning in mathematics learning at the primary and secondary education levels. To achieve these aims, this study employs a systematic literature review approach to trace, select, analyze, and synthesize relevant research findings in a structured manner. Through this approach, the study is expected to produce a comprehensive picture of the conceptual development and practice of AfL in mathematics learning.

METHODS

Research Design

This study employed a *Systematic Literature Review (SLR)* design to examine the conceptual development and empirical findings of Assessment for Learning (AfL) in mathematics education. A systematic literature review was considered appropriate because it allows researchers to identify, select, evaluate, and synthesize previous studies in a structured and transparent way (Petticrew & Roberts, 2006; Xiao & Watson, 2019). In this study, the review process was guided by the PRISMA 2020 framework, which provides systematic procedures for reporting the processes of identification, screening, eligibility assessment, and inclusion of studies (Page et al., 2021). The findings from the selected studies were analyzed using thematic synthesis, which is suitable for identifying patterns, developing descriptive themes, and constructing analytical interpretations across reviewed studies (Thomas & Harden, 2008).

Time and Place of the Study

This study was conducted from 2015-2025. Since this research was literature-based, the research setting was not a physical school or classroom, but academic publication databases. The literature search was conducted through Scopus, Web of Science, ERIC, and Dimensions, with additional searching through Google Scholar to broaden the coverage of relevant studies. The use of multiple databases is recommended in systematic reviews to increase the comprehensiveness of the search and reduce the possibility of selection bias (Gusenbauer & Haddaway, 2020; Xiao & Watson, 2019).

Population and Sample

The population of this study consisted of peer-reviewed journal articles related to Assessment for Learning in mathematics education published between 2015 and 2025. In a systematic literature review, the unit of analysis is not individual respondents, but relevant documents or studies that meet predetermined criteria (Snyder, 2019). Therefore, the sample of this study consisted of 10 selected articles that met the inclusion criteria and were included in the final synthesis. The sampling technique used in this study was purposive sampling, because the article. were selected based on specific criteria relevant to the research focus. Purposive selection is appropriate in literature-based studies when researchers need to include only sources that directly address the topic, context, and research objectives (Petticrew & Roberts, 2006). The inclusion criteria were: peer-reviewed journal articles; published between 2015 and 2025; written in English or Indonesian; focused on mathematics education at the primary or secondary level; and discussed Assessment for Learning, formative assessment, digital assessment, feedback, self-assessment, peer assessment, or student-centered assessment. Articles were excluded if they were not related to mathematics education, did not focus on AfL or formative assessment, were not peer-reviewed, were inaccessible in full text, or were not relevant to primary and secondary education.

Research Procedure

The research procedure consisted of three main stages: literature identification, study selection, and data analysis. First, the literature was identified by searching academic databases using combinations of keywords such as Assessment for Learning, formative assessment, mathematics education, digital assessment, feedback," peer assessment, self-assessment, and student-centered assessment**.** Keyword combination and database searching are important steps in systematic reviews because they determine the breadth and relevance of the studies retrieved (Xiao & Watson, 2019). The initial search identified 150 articles from the selected databases.

Second, the study selection process was conducted following the PRISMA 2020 stages. All retrieved articles were compiled and checked for duplication. After duplicate removal, the titles and abstracts were screened to determine their relevance to the research focus. The remaining articles were then assessed through full-text eligibility review based on the inclusion and exclusion criteria. This process is consistent with PRISMA 2020, which emphasizes transparent reporting of how studies are identified, screened, assessed for eligibility, and included in the final review (Page et al., 2021). After the selection, screening, and eligibility assessment process, 10 articles were included in the final review.

Third, the selected articles were analyzed using thematic synthesis**.** The analysis began by reading each article carefully to understand its focus, method, findings, and contribution to AfL in mathematics education. Relevant data were then coded based on recurring ideas related to digital AfL practices, formative assessment characteristics, feedback use, student participation, and implementation factors. The codes were grouped into descriptive themes and further interpreted into analytical themes. This procedure follows the thematic synthesis approach, which involves coding findings, developing descriptive themes, and generating analytical themes to produce a broader conceptual understanding of the reviewed literature (Thomas & Harden, 2008). Through this process, the study produced a conceptual synthesis of AfL as a pedagogically formative, digitally supported, and student-centered assessment practice in mathematics education. records were removed. titles and abstracts were screened to identify potentially relevant studies.

Fourth, full-text articles were assessed based on the inclusion and exclusion criteria, including publication type, educational context, topical relevance, and full-text availability. Only studies that fully met these criteria were retained for the final synthesis.

Data Analysis

Data from the selected articles were extracted systematically using a structured review matrix. The extracted information included author(s), publication year, country/context, educational level, research design, AfL focus, digital dimension, formative elements, student-centered orientation, major findings, implementation challenges, and conceptual contributions. The data were then analyzed using thematic synthesis. This process involved three stages: open coding of key findings from each study, categorization of similar codes into broader groups, and development of analytical themes**.** The final synthesis was organized into three major themes: Assessment for Learning (AfL) as a digital practice, AfL as a formative practice, and AfL as a student-centered practice**.** All stages of searching, selecting, extracting, and analyzing data were documented systematically to ensure the consistency and trustworthiness of the review findings.

FINDINGS AND DISCUSSION

Table 1. Synthesied Article

No	Penulis/ Tahun	Judul Artikel	Jurnal, Volume (Nomor), Halaman	Negara/ Konteks	Detail Studi	Temuan Utama
1	Faber, Luyten, & Visscher (2017)	*The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment*	*Computers & Education,* 106, 83–96	Belanda; siswa kelas 3 sekolah dasar	Eksperimen acak pada 79 sekolah dan 1.808 siswa; menggunakan alat asesmen formatif digital **Snappet**	Alat asesmen formatif digital memberi efek positif pada capaian matematika dan motivasi siswa; efek capaian lebih kuat pada siswa berprestasi tinggi
2	van den Berg, Bosker, & Suhre (2018)	*Testing the effectiveness of classroom formative assessment in Dutch primary mathematics education*	*School Effectiveness and School Improvement,* 29(3), 339–361	Belanda; pembelajaran matematika sekolah dasar	Kuasi-eksperimen; 17 guru pada kelompok perlakuan dan 17 guru pada kelompok kontrol	Perbedaan hasil pascates antarkelompok tidak signifikan setelah kontrol variabel, tetapi **kualitas implementasi** classroom formative assessment berhubungan positif dengan performa matematika siswa kelas 5
3	Rakoczy, Pinger, Hochweber, Klieme, Schütze, & Besser (2019)	*Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy*	*Learning and Instruction,* 60, 154–165	Jerman; kelas matematika tingkat menengah (kelas 9)	Studi kuantitatif pada kelas matematika sekolah menengah	Menunjukkan bahwa **kebermanfaatan umpan balik yang dirasakan** dan **self-efficacy siswa** merupakan mekanisme penting dalam menjelaskan pengaruh AfL pada pembelajaran matematika
4	Kyaruzi, Strijbos, Ufer, & Brown (2019)	*Students' formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania*	*Assessment in Education: Principles, Policy & Practice,* 26(3), 278–302	Tanzania; 48 sekolah menengah, 2.767 siswa Form 3/Grade 11	Survei dan *focus group discussion*; analisis SEM	Persepsi siswa terhadap kualitas umpan balik dan *scaffolding* guru memprediksi penggunaan umpan balik; penggunaan umpan balik kemudian berhubungan positif dengan performa matematika
5	Patchan, Rambo-Hernandez, Deitz,	*Using peer assessment to improve middle school mathematical*	*The Journal of Educational Research,*	Amerika Serikat; siswa kelas 6/middle	Membandingkan umpan balik dari empat teman	Siswa sekolah menengah pertama dapat memberi umpan balik yang **sama bermanfaatnya** dengan guru

DOI: https://doi.org/10.29303/jipp.v11i2.4867

No	Penulis/ Tahun	Judul Artikel	Jurnal, Volume (Nomor), Halaman	Negara/ Konteks	Detail Studi	Temuan Utama
	& McNeill (2022)	*communication*	115(2), 146–160	school	sebaya dengan umpan balik guru dalam tugas matematika	untuk meningkatkan komunikasi matematis dan kemampuan pemecahan masalah
6	Zulliger, Buholzer, & Ruelmann (2022)	*Observed quality of formative peer and self-assessment in everyday mathematics teaching and its effects on student performance*	*European Journal of Educational Research*, 11(2), 663–680	Swiss; pembelajaran matematika sekolah dasar	Observasi kualitas *peer-* dan *self-assessment* dalam pembelajaran matematika; 634 siswa	Tidak ditemukan efek umum yang besar terhadap performa, tetapi *self-assessment* berkualitas tinggi bermanfaat bagi siswa berprestasi rendah karena memperluas repertoar strategi berhitung mereka
7	Rodríguez-Martínez, González-Calero, del Olmo-Muñoz, Arnau, & Tirado-Olivares (2023)	*Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions*	*British Journal of Educational Technology*, 54(1), 76–97	Spanyol; siswa kelas 5 sekolah dasar	AfL berbasis *learning analytics* untuk menyusun pekerjaan rumah yang dipersonalisasi pada topik pecahan	Personalisasi tugas berdasarkan data asesmen formatif membantu meningkatkan pemahaman siswa tentang pecahan dan menunjukkan potensi kuat AfL digital berbasis data
8	de Vries, Dimosthenous, Schildkamp, & Visscher (2023)	*The impact of an assessment for learning teacher professional development program on students' metacognition*	*School Effectiveness and School Improvement*, 34(1), 109–129	Belanda; guru matematika sekolah menengah	Studi eksperimen pada 47 guru dan 803 siswa; program pengembangan profesional guru berbasis AfL	Program pengembangan profesional guru berdampak positif pada kemampuan metakognitif siswa untuk **memprediksi** dan **mengevaluasi**, meskipun tidak signifikan pada aspek perencanaan
9	Smit, Dober, Hess, Bachmann, & Birri (2023)	*Supporting primary students' mathematical reasoning practice: The effects of formative feedback and the mediating role of self-efficacy*	*Research in Mathematics Education*, 25(3), 277–300	Swiss; 71 kelas sekolah dasar, 1.261 siswa	Program 10 minggu; analisis multilevel tentang umpan balik formatif, penalaran, dan self-efficacy	Pada level kelas, umpan balik formatif memprediksi penalaran matematis melalui mediasi self-efficacy; pada level individu, umpan balik terutama memprediksi self-efficacy
10	Barana & Marchisio Conte (2024)	*Promoting socioeconomic equity through automatic formative assessment*	*Journal on Mathematics Education*, 15(1), 227–252	Italia; siswa kelas 8/lower secondary	Studi eksperimen selama satu tahun; 299 siswa kelompok perlakuan dan 257 siswa kontrol	*Automatic formative assessment* memberi dampak positif pada hasil belajar matematika, dengan efek lebih kuat pada siswa dari latar SES rendah dan kecenderungan mengurangi ketimpangan hasil belajar

Based on Table 1, the ten selected articles indicate that Assessment for Learning in mathematics education has developed in three major directions. First, the digital dimension is strongly evident in the studies by Faber et al. (2017), Rodríguez-Martínez et al. (2023), and Barana and Marchisio Conte (2024), which show that technology enables real-time feedback, task personalization, and the strengthening of equity in learning outcomes. Second, the formative dimension is reflected in the studies by van den Berg et al. (2018), Rakoczy et al. (2019), and de

Vries et al. (2023), which emphasize the importance of implementation quality, the perceived usefulness of feedback, and teachers' capacity to enact AfL. Third, the student-centered dimension appears in the studies by Kyaruzi et al. (2019), Patchan et al. (2022), Zulliger et al. (2022), and Smit et al. (2023), which demonstrate that the use of feedback, peer assessment, self-assessment, and self-efficacy plays an important role in strengthening communication, reasoning, and engagement in mathematics learning. Based on the PRISMA 2020 selection process, 10 articles were identified that met the inclusion criteria and were considered the most relevant to the focus of this study, namely Assessment for Learning (AfL) in mathematics education at the primary and secondary levels, with an emphasis on digital practices, formative characteristics, and student-centered orientation. These ten studies came from diverse contexts, namely the Netherlands, Germany, Tanzania, the United States, Switzerland, Spain, and Italy. In terms of educational level, the selected studies covered primary school, junior secondary school, and secondary school. Meanwhile, in terms of research design, the reviewed articles showed methodological diversity, including randomized experiments, quasi-experiments, surveys, classroom observations, and teacher professional development studies. This diversity indicates that AfL in mathematics education has been examined from broad perspectives, both as a learning intervention, as a classroom practice, and as part of teacher capacity development.

Substantively, the selected articles show that AfL in mathematics education is developing in increasingly complex ways. Some studies position AfL as a digital practice that enables rapid feedback, adaptive tasks, and personalized learning, as seen in the studies by Faber et al. (2017), Rodríguez-Martínez et al. (2023), and Barana and Marchisio Conte (2024). Others emphasize the formative function of AfL in improving the quality of learning interactions, the usefulness of feedback, self-efficacy, metacognition, and pedagogical decision-making, as reflected in the studies by van den Berg et al. (2018), Rakoczy et al. (2019), de Vries et al. (2023), and Smit et al. (2023). Meanwhile, another group of studies highlights the strengthening of students' roles through peer assessment, self-assessment, the use of feedback, and mathematical communication, as seen in the studies by Kyaruzi et al. (2019), Patchan et al. (2022), and Zulliger et al. (2022). Overall, the description of the selected studies shows that AfL is no longer understood merely as a classroom evaluation technique, but as a pedagogical ecosystem that connects evidence of learning, teachers' actions, student participation, and technological support

Discussion

The review findings show that Assessment for Learning (AfL) in mathematics education at the primary and secondary levels can no longer be understood merely as formative assessment in a narrow sense, but rather as a pedagogical system that develops through the integration of formative functions, digital technology support, and students' active engagement. This finding is important because it demonstrates that the development of AfL in the mathematics literature has shifted from a teacher-centered model and conventional feedback toward a more dialogic, adaptive, and student-agency-oriented model. Accordingly, the following discussion is organized based on the research questions so that the relationship between the synthesized findings and the focus of the study can be presented systematically.

The conceptual development of Assessment for Learning in mathematics education

The first research question asks how the conceptual development of Assessment for Learning in mathematics education is explained in the literature. Based on the synthesis of the review findings, it can be understood that conceptually AfL remains grounded in the core idea that assessment is used to improve learning during the learning process, rather than merely to evaluate final outcomes (Sadler, 1989; Black & Wiliam, 1998; Black & Wiliam, 2009). However, in the context of mathematics education, this concept has developed more specifically because the nature of mathematics requires assessment that is able to capture students' thinking processes, solution strategies, misconceptions, and the quality of their reasoning, rather than merely the accuracy of answers (Kim, 2019; Rakoczy et al., 2019; Maskos et al., 2025).

The review findings also show that the conceptual development of AfL in mathematics is moving toward a stronger responsiveness to

DOI: https://doi.org/10.29303/jipp.v11i2.2172

students' mathematical thinking. Kim's (2019) findings emphasize that the implementation of formative assessment lessons creates opportunities for teachers to learn how to become more responsive to students' mathematical thinking. This indicates that AfL in mathematics is not merely an evaluative instrument, but also a pedagogical means of understanding how students think. This finding strengthens the view that the concept of AfL in mathematics should be understood as an interactive process involving teachers, students, and learning evidence.

Thus, the conceptual development of AfL in mathematics education can be seen as a shift from an outcome-oriented assessment model toward a model oriented to diagnosis, pedagogical response, and continuous improvement. The concept of AfL in mathematics becomes stronger when it is positioned not merely as a classroom evaluation strategy, but as a pedagogical mechanism to support mathematical reasoning, learning reflection, and instructional adaptation.

Digital practices of Assessment for Learning in mathematics education

The second research question focuses on how digital Assessment for Learning practices are implemented in mathematics education. The review findings show that the digital dimension is one of the most prominent developments in AfL. Faber et al. (2017) found that the use of digital formative assessment tools can improve students' mathematics achievement and motivation. Rodríguez-Martínez et al. (2023) showed that learning analytics-based assessment can be used to design personalized homework, while Söderström (2024) demonstrated that computer-based formative assessment can support mathematical problem solving through heuristic and metacognitive feedback.

These findings indicate that digital technology expands the function of AfL in at least three ways. First, digitalization accelerates the process of collecting learning evidence. Second, digitalization enables the personalization of interventions based on student learning data. Third, digitalization enriches the forms of feedback that can be provided to students. In the context of mathematics education, these three aspects are highly important because students often need rapid responses to procedural errors, conceptual misconceptions, or inappropriate problem-solving strategies.

However, the synthesis also shows that technology does not automatically make assessment formative. Its formative function still depends on how the digital information is used by teachers and students to improve learning. In other words, the digital dimension in mathematics AfL must be understood not merely as the use of technological devices, but as a transformation in the way learning evidence is collected, interpreted, and used pedagogically. Therefore, effective digital AfL practices are those that preserve the core principle of formative assessment, namely the use of information for learning improvement.

The characteristics of formative and student-centered AfL practices

The third research question asks how the characteristics of formative and student-centered AfL practices appear in mathematics education. The synthesis shows that the formative characteristics of AfL are evident through a clear cycle involving learning goals, evidence collection, feedback, and follow-up instructional action. van den Berg et al. (2018) found that the impact of classroom formative assessment does not always emerge automatically, which means that the quality of implementation is the primary determining factor. Rakoczy et al. (2019), Kyaruzi et al. (2019), and Smit et al. (2023) consistently showed that the usefulness of feedback and students' self-efficacy are important mechanisms explaining the influence of AfL on mathematics learning outcomes.

From a student-centered perspective, the review findings show that AfL becomes more effective when students are positioned as active users of assessment information. Patchan et al. (2022) showed that peer assessment can improve students' mathematical communication. Kyaruzi et al. (2019) emphasized that what matters is not only the quality of the feedback provided by teachers, but also whether students actually use that feedback. Gezer et al. (2021) also showed that formative assessment can be especially helpful in lower-performing classrooms, indicating the importance of student engagement in using assessment data for learning improvement.

These findings suggest that the main characteristics of formative and student-centered AfL practices in mathematics education include several elements. First, assessment must be designed to make students' thinking processes visible. Second, feedback must be clear, actionable, and relevant to the learning goals. Third, students must be involved in understanding criteria, reflecting on their work, using feedback, and improving their learning strategies. Thus, ideal AfL practice is not only formative in function, but also participatory in pedagogical terms.

Supporting and inhibiting factors in the implementation of digital, formative, and student-centered AfL

The fourth research question highlights the factors that support and hinder the implementation of digital, formative, and student-centered AfL in mathematics education. Based on the synthesis, the main supporting factors lie in teachers' competence, especially in assessment literacy, the ability to interpret learning evidence, and the capacity to adjust instruction responsively (Kim, 2019; Schildkamp et al., 2020). In addition, adequate technological support, well-designed mathematics tasks, and high-quality feedback are also important factors that strengthen the successful implementation of AfL (Faber et al., 2017; Rodríguez-Martínez et al., 2023; Söderström, 2024).

From the student perspective, supporting factors for AfL implementation include self-efficacy, positive perceptions of the usefulness of feedback, and readiness to engage in self- or peer-assessment (Rakoczy et al., 2019; Kyaruzi et al., 2019; Smit et al., 2023). This indicates that the success of AfL depends not only on teacher design, but also on students' readiness to act as active learning agents.

Conversely, the inhibiting factors identified in the literature include uneven quality of implementation, limited teacher capacity in managing formative assessment, the potential use of technology for merely administrative purposes, and the suboptimal use of feedback by students. The findings of van den Berg et al. (2018) are important in this regard, as they show that the implementation of AfL without strong quality does not automatically produce significant impact. Thus, the main challenge is not only bringing AfL into the classroom, but ensuring that it is implemented in a high-quality, meaningful way that is aligned with students' mathematics learning needs.

The reconceptualization of Assessment for Learning in mathematics education

The fifth research question asks how previous research findings can be synthesized to produce a reconceptualization of AfL in mathematics education. Based on the overall findings, the reconceptualization of AfL in this study leads to the understanding that effective AfL is pedagogically formative, digitally strengthened, and activated through student participation. The formative dimension provides the core function of using evidence to improve learning. The digital dimension expands the speed, accuracy, documentation, and personalization of assessment. The student-centered dimension ensures that assessment is not only done to students, but also with students.

Accordingly, AfL in mathematics education at the primary and secondary levels needs to be understood as an ecosystem of pedagogical assessment. In this ecosystem, teachers act as designers and directors of the use of learning evidence, technology functions as a mediator for the collection and use of data, and students become active agents in reading, interpreting, and using assessment information to improve their own learning. This understanding goes beyond the classical view that sees AfL only as a teacher feedback technique, because the synthesis results show that the success of AfL depends greatly on the interconnectedness between feedback, teacher response, technological support, and student agency.

This reconceptualization also carries a theoretical implication that AfL in mathematics can no longer be positioned as a small subsection of classroom assessment, but rather as a broader framework that connects assessment, learning, technology, and student self-regulation. Practically, these results indicate that the development of AfL in schools needs to be directed toward teacher training, strengthening assessment literacy, meaningful use of technology, and the creation of classroom cultures that support students' active engagement.

Overall, this discussion shows that the five research questions are interconnected and form one coherent picture. The conceptual development of AfL in mathematics points toward an assessment model that is increasingly diagnostic and responsive. Digital practices expand the capacity of AfL to function rapidly and adaptively. Formative and student-centered characteristics clarify that the quality of feedback and student participation are at the core of AfL's success. Supporting and inhibiting factors emphasize that AfL implementation is strongly influenced by teacher quality, technology, and student readiness. All of these findings ultimately point to the need to reconceptualize AfL as an integrative assessment system in mathematics education

This study aimed to reconceptualize Assessment for Learning (AfL) in mathematics education at the primary and secondary levels through a systematic literature synthesis focusing on three main dimensions: digital practices, formative nature, and student-centered orientation. Based on the review of 10 selected articles, it can be concluded that AfL in mathematics education has evolved from merely a classroom assessment technique into a broader and more complex pedagogical framework. Assessment for Learning (AfL) can no longer be adequately understood simply as the provision of feedback during instruction, but rather as a system that continuously connects the collection of learning evidence, the interpretation of data, students' active engagement, and teachers' instructional follow-up.

From the perspective of digital practice, the findings show that technology has expanded the capacity of Assessment for Learning (AfL) through real-time feedback, automated assessment, learning analytics, and the personalization of mathematics learning tasks. These findings confirm that digitalization not only changes the medium of assessment, but also reconfigures the logic of learning by enabling evidence of learning to be collected, interpreted, and acted upon more quickly and more precisely. From the formative perspective, this study shows that the main strength of Assessment for Learning (AfL) lies in its ability to alter the course of learning through the use of learning evidence and meaningful feedback. AfL becomes truly formative when the information obtained does not remain merely as data, but is used to improve students' learning strategies and teachers' pedagogical decisions. From the student-centered perspective, the synthesis shows that effective AfL creates space for students to be actively involved in understanding learning goals, using feedback, engaging in reflection, and participating in self-assessment as well as peer assessment.

This study also concludes that the effectiveness of Assessment for Learning (AfL) is strongly influenced by the quality of its implementation. Teachers continue to play a central role in designing assessment pathways, interpreting learning evidence, building structures for reflection, and orchestrating instructional follow-up. Therefore, the success of Assessment for Learning (AfL) depends not only on the availability of technology or assessment instruments, but more importantly on teachers' pedagogical capacity and support for student agency in the mathematics learning process.

Conceptually, the main contribution of this study is to affirm that AfL relevant to contemporary mathematics education needs to be understood as a practice that is simultaneously digital, formative, and student-centered. This reconceptualization provides a more integrative foundation for the development of both theory and practice in mathematics assessment, while also addressing the fragmentation in the literature, which has tended to separate these three dimensions. Thus, Assessment for Learning (AfL) in mathematics education should no longer be positioned as a supplementary element of instruction, but rather as the core of a learning ecosystem that is responsive, reflective, and adaptive to students' needs.

This study has several limitations that need to be acknowledged. As a systematic literature review, the synthesis results depend greatly on the quality and scope of the articles retrieved, so the choice of databases, keywords, publication year range, and language directly affects the breadth of the findings. In addition, the reviewed studies come from diverse national contexts, educational levels, research designs, and forms of Assessment for Learning (AfL) implementation, so the generalization of the findings to all educational contexts should be made with caution. This study also focuses on three main dimensions—digital, formative, and student-centered—so other factors such as curriculum, assessment policy, school culture, instructional leadership, and socio-economic conditions have not been discussed in depth. Furthermore, because this study is conceptual in nature, it has not yet empirically tested the resulting reconceptualized model of Assessment for Learning (AfL) in specific mathematics classroom contexts.

Based on these limitations, future research is recommended to empirically test an

Assessment for Learning (AfL) model that integrates digital, formative, and student-centered dimensions within a comprehensive mathematics instructional design. Future studies also need to expand the research context, particularly in developing countries and Southeast Asia, so that the findings become more contextualized to local conditions. In addition, research should give greater attention to implementation factors, especially teacher competence in assessment literacy, data literacy, the interpretation of mathematical thinking, and the facilitation of self-assessment and peer assessment. Going forward, a combination of systematic review, meta-analysis, and longitudinal studies is also important in order to develop a more comprehensive understanding of AfL, including its impact not only on academic achievement, but also on mathematical communication, higher-order thinking, self-regulation, motivation, self-efficacy, and equity in learning.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to STKIP Citra Bakti, Indonesia, and Universitas Pendidikan Ganesha, Singaraja, Indonesia, for the academic and institutional support provided throughout the completion of this manuscript. The authors also extend their appreciation to colleagues and all parties who contributed, directly or indirectly, to the development and refinement of this study. Their support, encouragement, and constructive input have been invaluable to the completion of this work

REFERENCES

Alemdag, E., & Yildirim, Z. (2022). Design and development of an online formative peer assessment environment with instructional scaffolds. *Educational Technology Research and Development, 70*(4), 1359–1389. https://doi.org/10.1007/s11423-022-10115-x

Baiduri. (2022). Effect of self and peer assessments on mathematics learning achievement. *Al-Jabar: Jurnal Pendidikan Matematika, 13*(1), 13–21. https://doi.org/10.24042/ajpm.v13i1.10731

Barana, A., & Marchisio Conte, M. (2024). Promoting socioeconomic equity through automatic formative assessment. *Journal on Mathematics Education, 15*(1), 227–252. https://doi.org/10.22342/jme.v15i1.pp227-252

Barana, A., Marchisio, M., & Sacchet, M. (2021). Interactive feedback for learning mathematics in a digital learning environment. *Education Sciences, 11*(6), 279. https://doi.org/10.3390/educsci11060279

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Chang, C.-C., Tseng, K.-H., & Lou, S.-J. (2012). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a web-based portfolio assessment environment for high school students. *Computers & Education, 58*(1), 303–320. https://doi.org/10.1016/j.compedu.2011.08.005

de Vries, J. A., Dimosthenous, A., Schildkamp, K., & Visscher, A. J. (2022). The impact on student achievement of an assessment for learning teacher professional development program. *Studies in Educational Evaluation, 74*, 101184. https://doi.org/10.1016/j.stueduc.2022.101184

de Vries, J. A., Dimosthenous, A., Schildkamp, K., & Visscher, A. J. (2023). The impact of an assessment for learning teacher professional development program on students' metacognition. *School Effectiveness and School Improvement, 34*(1), 109–129.

https://doi.org/10.1080/09243453.2022.2116461

Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review, 32*, 481–509. https://doi.org/10.1007/s10648-019-09510-3

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education, 106*, 83–96. https://doi.org/10.1016/j.compedu.2016.12.001

Gezer, T., Wang, C., Polly, A., Martin, C., Pugalee, D. K., & Lambert, R. (2021). The relationship between formative assessment and summative assessment in primary grade mathematics classrooms. *International Electronic Journal of Elementary Education, 13*(5), 673–685. https://doi.org/10.26822/iejee.2021.220

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods, 11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Kim, H.-J. (2019). Teacher learning opportunities provided by implementing formative assessment lessons: Becoming responsive to student mathematical thinking. *International Journal of Science and Mathematics Education, 17*(2), 341–363. https://doi.org/10.1007/s10763-017-9866-7

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*(4), 28–37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

Kyaruzi, F., Strijbos, J.-W., Ufer, S., & Brown, G. T. L. (2019). Students' formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania. *Assessment in Education: Principles, Policy & Practice, 26*(3), 278–302. https://doi.org/10.1080/0969594X.2019.1593103

Lei, W., & Lei, Z. (2026). Formative assessment literacy: A systematic review. *Language Testing in Asia, 16*, 7. https://doi.org/10.1186/s40468-025-00418-0

Liu, J., Zhou, L., Bharaj, P. K., Zhou, D., & Lo, J.-J. (2024). Using international large-scale assessment for learning: Analyzing U.S. students' geometry performance in TIMSS. *Journal of Research and Advances in Mathematics Education, 9*(3), 126–143. https://doi.org/10.23917/jramathedu.v9i3.4019

Maskos, K., Schulz, A., Oeksuez, S. S., & Rakoczy, K. (2025). Formative assessment in mathematics education: A systematic review. *ZDM–Mathematics Education, 57*, 679–693. https://doi.org/10.1007/s11858-025-01696-x

Ong, E. T. J., & Toh, T. L. (2024). Self-directed learning through gamification of mathematics lessons: Literature review and a framework for enactment. *Journal of Research and Advances in Mathematics Education,* 9(4), 205–221. https://doi.org/10.23917/jramathedu.v9i4.4729

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review,* 22, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Patchan, M. M., Rambo-Hernandez, K. E., Deitz, B. N., & McNeill, J. (2022). Using peer assessment to improve middle school mathematical communication. *The Journal of Educational Research, 115*(2), 146–160. https://doi.org/10.1080/00220671.2022.2074948

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing.

Puspitaningtyas, A. R., Nurhasanah, F., & Indriati, D. (2025). Emerging trends in mathematics assessment practices in Southeast Asia. *Journal of Research and Advances in Mathematics Education, 10*(3), 167–181. https://doi.org/10.23917/jramathedu.v10i3.8868

Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction, 60*, 154–165. https://doi.org/10.1016/j.learninstruc.2018.01.004

Rodríguez-Martínez, J. A., González-Calero, J. A., del Olmo-Muñoz, J., Arnau, D., & Tirado-Olivares, S. (2023). Building personalised homework from a learning analytics based formative assessment: Effect on fifth-grade students' understanding of fractions. *British Journal of Educational Technology, 54*(1), 76–97. https://doi.org/10.1111/bjet.13292

Rosidin, U., Haryanti, N., Lora, H. A., & Viyanti. (2020). Reconstruct the class assessment strategy: Promoting the 21st-century learning. *Indonesian Journal of Science and Mathematics Education, 3*(1), 27–36. https://doi.org/10.24042/ijsme.v3i1.6056

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation, 37*(1), 15–24. https://doi.org/10.1016/j.stueduc.2011.04.003

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144. https://doi.org/10.1007/BF00117714

Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research, 103*, 101602. https://doi.org/10.1016/j.ijer.2020.101602

Smit, R., Dober, H., Hess, K., Bachmann, P., & Birri, T. (2023). Supporting primary students' mathematical reasoning practice: The effects of formative feedback and the mediating role of self-efficacy. *Research in Mathematics Education, 25*(3), 277–300. https://doi.org/10.1080/14794802.2022.2062780

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Söderström, S. (2024). Computer-based formative assessment for problem solving. *International Journal of Mathematical Education in Science and Technology, 55*(10), 2645–2669. https://doi.org/10.1080/0020739X.2023.2178982

Sunarto, M. J. D. (2025). Integrating AI in individualized mathematics learning: A teacher-centered approach. *Indonesian Journal of Science and Mathematics Education, 8*(3), 531–548. https://doi.org/10.24042/ijsme.v8i3.26795

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*, Article 45. https://doi.org/10.1186/1471-2288-8-45

Topping, K. J. (2009). Peer assessment. *Theory Into Practice, 48*(1), 20–27. https://doi.org/10.1080/00405840802577569

van den Berg, M., Bosker, R. J., & Suhre, C. J. M. (2018). Testing the effectiveness of classroom formative assessment in Dutch primary mathematics education. *School Effectiveness and School Improvement, 29*(3), 339–361. https://doi.org/10.1080/09243453.2017.1406376

van den Berg, M., Harskamp, E. G., & Suhre, C. J. M. (2016). Developing classroom formative assessment in Dutch primary mathematics education. *Educational Studies, 42*(4), 305–322. https://doi.org/10.1080/03055698.2016.1193475

Wardani, N., & Siregar, T. J. (2023). The development of HOTS (higher-order thinking skills) assessment instrument based on numeration literacy in mathematics for junior high school. *Indonesian Journal of Science and Mathematics Education, 6*(2), 238–252. https://doi.org/10.24042/ijsme.v6i2.17433

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research, 39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Yan, Z., Lao, H., Panadero, E., Fernández-Castilla, B., Yang, L., & Yang, M. (2022). Effects of self-assessment and peer-assessment interventions on academic performance: A meta-analysis. *Educational Research Review, 37*, 100484. https://doi.org/10.1016/j.edurev.2022.100484

Zulliger, S., Buholzer, A., & Ruelmann, M. (2022). Observed quality of formative peer and self-assessment in everyday mathematics teaching and its effects on student performance. *European Journal of Educational Research, 11*(2), 663–680. https://doi.org/10.12973/eu-jer.11.2.663