Journal of Social Sciences & Humanities
Publisher: Kabul University

Available at https://jssh.edu.af

The Evolution of Formative Assessment Research: A Bibliometric Mapping of Knowledge Structure, Collaboration Networks, and Emerging Trends

Ali Reza Haidari[1*], Reza Naderi [2], Mahdi Nazari
[1, 3]*Bamyan University, Department of Pedagogy, Faculty of Education, Bamyan, AF*
[2]*Bamyan University, Department of English, Faculty of Education, Bamyan, AF*

Received: June 24, 2026
Revised: August12, 2026
Accepted: Oct 13, 2026
Published: Oct 30, 2026

Keywords
- Assessment for learning
- Assessment Literacy
- Bibliometric Analysis
- Formative Assessment
- Self-Regulated Learning

Abstract: Formative assessment has gained extensive academic interest because of its importance for learning, feedback provision, and fostering self-regulated learning. Nevertheless, a framework delineating the intellectual background, field collaboration, and key concepts is needed. This study aimed to visualize the knowledge structure and publication dynamics of formative assessment literature. Data were downloaded from the Web of Science Core Collection, and 425 English-language articles published from 2015 to 2026 were selected using PRISMA-based criteria. Five analyses were conducted using VOSviewer (version 1.6.20): publication trends, citation analysis, coauthorship networks, co-citation analysis, and keyword co-occurrence. Findings showed continuous scholarly production within the research timeframe, with activity concentrated among several authors, institutions, and countries, including the USA, China, and various European countries. Co-citation analysis highlighted educational feedback, assessment for learning, and self-regulated learning as key components of the field's knowledge structure. Keyword analysis uncovered topics such as assessment literacy, learner engagement, learning analytics, gamification, and technology-assisted assessment. Artificial intelligence, personalized feedback, adaptive assessment, and ethical questions in digital assessment require further exploration. However, the mere presence of these topics in the database is not sufficient to determine their development stage or the existence of research gaps. This research gives a structured overview of formative assessment literature in the Web of Science database and suggests directions for further bibliometric and empirical research. Findings should be interpreted in light of the database's coverage and the incomplete information about publications in 2026.

To Cite this Article: Haidari, A. R., Naderi, R., & Nazari, M. (2026). The Evolution of Formative Assessment Research: A Bibliometric Mapping of Knowledge Structure, Collaboration Networks, and Emerging Trends. *Journal of Social Sciences & Humanities 3*(4), 1044-1064. https://doi.org/10.62810/jssh.v3i4.382

Copyright © 2026 Author(s). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

INTRODUCTION

In recent decades, formative assessment has been one of the most crucial elements of educational evaluation. Formative assessment has shifted from measuring learning outcomes

✉ Corresponding author E-mail: arhaidari@bu.edu.af

Available online at https://jssh.edu.af/jssh/article/view/382

to facilitating learning (Brown, 2019). This shift has become possible because learner-centered and constructivist approaches view learning as an active process of building knowledge (Yan & Boud, 2021). From this perspective, assessment is an integral part of the teaching and learning process, not just a tool for evaluating learners' achievements. It enhances learning by improving instruction based on evidence about learners' performance (Acuña & de Dávila, 2021). That is why many researchers and educational policy-makers have studied the role of formative assessment in the educational process.

In the literature on evaluation in education, there are three different concepts of “Assessment of Learning”, “Assessment for Learning” and “Assessment as Learning” that aim at clarifying the varied functions of evaluation in the teaching-learning process (Schellekens et al., 2021). The "Assessment of Learning" concentrates primarily on the measurement of how far the learners have met the expected goals of education; in this case, it evaluates how far the learners have met the set criteria. Conversely, assessment for learning uses evidence to recognize learners’ needs, provide feedback, and inform instruction; therefore, beyond assessing progress, its purpose is to facilitate improvement in the learning process (Brown, 2019). Assessment as learning focuses on the learner's active role in tracking their learning, evaluating themselves, and reconsidering the learning process, and on the idea that assessment is part of self-regulatory development (Yan, 2021; Yan & Boud, 2021). Although different, the three assessment types mentioned above do not necessarily have to be separate and may intersect. In this way, formative evaluation is closely connected to evaluation for learning and may also include assessment as learning, depending on how it is carried out (Brown, 2019; Schellekens et al., 2021).

The first mechanism in formative assessment is feedback that helps learners identify gaps between their current learning and the desired outcomes and take steps to improve their performance (Acuña & de Dávila, 2021). Nevertheless, feedback efficiency depends not only on providing information about the correctness of responses, but also on how students use that information to change their learning strategies and future actions (Yan & Boud, 2021). From this standpoint, formative assessment and self-regulation are linked in the learning process, as feedback collected during evaluation can enable learners to reflect on their learning experience, identify difficulties, and adjust strategies to meet their needs (Ren et al., 2026).

Various studies have investigated the influence of formative assessment on student learning in different aspects. For example, Al Breiki & Yahaya (2026) conducted a meta-analysis on the influence of formative assessment through electronic gamification on students’ academic performance. (Al Breiki & Yahaya, 2026). Moreover, Esomonu & Nduka (2025) looked at the impact of formative assessment on academic performance in science subjects (Esomonu & Nduka, 2025). Moreover, Wu & Yu (2025) explored the part played by teachers’ emotional support in the association between formative assessment and academic performance(Wu & Yu, 2025). Thus, research on formative assessment involves more than analyzing its effect on academic performance; it has also considered other factors, such as

the evaluation method used and the conditions under which it is implemented. However, notable differences in context, evaluation methods, and outcomes analyzed call for periodic review of the research process in this area.

In recent years, formative assessment research has expanded beyond its traditional focus on classroom assessment. Formative assessment research in second language learning has been done to find out how formative assessment contributes to language learning and the development of learning among the learners (Nguyen & Nguyen, 2026; Vassiliou et al., 2022), whereas science education research has emphasized the state of research and evaluation in this field (Zhang et al., 2023). Formative assessment has also been emphasized in higher education and special education (Nguyen & Nguyen, 2026; TING, 2025). In addition, advances in digital technologies and artificial intelligence have opened new possibilities for research on feedback and evaluation. For instance, studies have been conducted in the areas of using learning analytics in formative assessment (Banihashem et al., 2025), providing feedback to students using AI technologies (Burner et al., 2025), as well as the capabilities and issues of formative feedback that is based on generative AI (Kasneci et al., 2023; Webb & Galamba, 2026). This expansion of topics and contexts has made formative evaluation a multidisciplinary field that requires a systematic review of relationships among subjects, researchers, and trends.

Despite the growing body of studies on formative assessment, earlier bibliometric analyses have examined various aspects of the subject but not its publication characteristics, knowledge structure, collaborative networks, and thematic evolution within a single framework. For instance, Karaçöp and İnaltekin (2023) explored the dynamics of formative assessment research in the educational literature from 2000 to 2021 using criteria such as citation analysis, co-authorship, and author keywords (Karaçöp & İnaltekin, 2023). Meanwhile, Sudakova et al. (2022) have focused exclusively on online formative assessment in higher education (Sudakova et al., 2022). Sultan et al. (2024) analyzed formative assessment literature cited in the Scopus database (Sultan et al., 2024). In contrast, Yılmaz (2025) analyzed formative assessment in science education only. Although the above studies provide valuable information on the area in question, it is challenging to form a common view of the field based on either study because their thematic focus, databases, and study periods differ (Yılmaz, 2025). Therefore, the current study seeks to integrate analyses of publication and citation patterns, author collaborations, co-citations, and keyword co-occurrences to create a research map of formative assessment indexed in the Web of Science Core Collection from 2015 to 2026.

This problem becomes especially acute because areas such as learning analytics, generative AI, and intelligent feedback have entered the sphere of formative assessment. For instance, the critical review by Banihashem et al. (2025) has covered the aspects of development, limitations, and future directions of applying learning analytics in formative assessment, while recent studies have examined the use of generative AI for formative feedback and the difficulties associated with connecting technology with educational goals

(Banihashem et al., 2025; Prompiengchai et al., 2025; Webb & Galamba, 2026). The above papers may help with specific areas. However, they cannot explain where new topics fit within the established framework of formative assessment concepts, which research trends have created the strongest connections, and how the field's thematic structure has evolved. In this regard, an integrated approach to analyzing publication trends, collaboration, citation networks, and keyword relationships may help reveal the field's knowledge structure and identify key topics, connections between research areas, and research gaps.

This study investigates the knowledge structure and history of studies on formative evaluation by analyzing scientific articles published in Web of Science from 2015-2026 using a bibliometric approach. In this context, the annual trends in article publication and citation practices are examined to determine the dynamics and scientific influence of research in this field, authors, countries, organizations, and key journals, by means of scientific production and citation indicators; the authorship and citation networks are studied to find out scientific collaboration and intellectual connections that are typical of this research field. Keyword co-occurrence and the dynamics of theme changes are also used to identify pivotal themes, thematic clusters, and emerging research areas. Hence, the present study is intended to answer four main questions:

1. What are the features of the publication and citation process of articles on the topic of formative evaluation over the period from 2015 to 2026?
2. Who are the prominent authors, countries, institutes, and journals in the scientific production and citation process? Which works are the most cited, and who are their authors?
3. What are the characteristics of the author and citation networks? Who are the prominent authors and scientific sources that shape the intellectual and scientific structure of formative evaluation research?
4. What are the prominent concepts and thematic clusters in formative evaluation research? What topical areas should future research highlight based on bibliometric findings?

In addressing the discrepancies between the coverage and analytical emphasis of the previous bibliometric studies, the current research presents a synthesized analysis of the publication and citation structure, the co-authorship networks, the co-citation relationships, and the keyword co-occurrence structures of the bibliographic data of the formative assessment research indexed in the Web of Science Core Collection from 2015 to 2026. Unlike previous studies, the present research does not treat these indices as detached indicators of research activity. The study integrates these indices to understand the volume and visibility of research output, research collaboration trends, the field's theoretical roots, and the research themes stressing the formative assessment discipline within the focused dataset. The study aims to synthesize these dimensions of bibliometric analysis into a single study by intentionally complementing previous studies with different subject orientations, database coverage, or conceptual bases.

RESEARCH METHOD

The study used bibliometric analysis to examine publication trends, citation patterns, conceptual structure, and scientific collaboration networks in formative assessment. Bibliometric analysis is a quantitative methodology for analyzing scientific output, including publication trends, citation relationships, collaboration trends, and thematic structures (Aria & Cuccurullo, 2017; Van Eck & Waltman, 2010). This study used bibliometric analysis to examine the temporal development of formative assessment literature and the relationships among scientific publications, scientists, and topics. The analysis was therefore organized into five headings: publication trends, citations, keyword co-occurrences, co-citations, and co-authorships.

Data Source and Search Strategy

This paper selected the Web of Science Core Collection as the search database. A search was performed for the keywords “FORMATIVE ASSESSMENT”, FORMATIVE EVALUATION”, and “ASSESSMENT FOR LEARNING” for the All-Fields category. The search was limited to 2015–2026, and the results were restricted to Articles, English, and SSCI. The database was accessed on 21 June 2026 at 22:44, and the search history and searched terms were recorded. The search string was based on the search terms and inclusion criteria, but the exact Boolean operators used could not be inferred from the search record, so the search string could not be reconstructed.

Selection Criteria and Screening Process

Preliminary article selection was done in stages. We initially retrieved 3375 records; after screening for English-language inclusion, 2302 records remained. From these 2302 records, we selected 756 documents by limiting the search to journals in the Social Science Citation Index, leaving 576 articles for abstract-level screening. Of these abstracts, 151 were excluded because they did not address the study's scope and research questions on formative assessment, leaving 425 abstracts for full-text screening. Initial screening was done on abstracts, not full-text articles. Interpret the numbers and reasons for exclusion in light of the search and screening histories provided.

Records were included in the analysis based on their relevance to formative assessment and the breadth of research questions. The process involved screening abstracts for studies that listed or discussed formative assessment, formative evaluation, or assessment for learning, and that aligned with the article's focus. If the abstract did not sufficiently demonstrate the study's relevance, the article was not included. Because this was an abstract-level screening and the full text of these studies was not retrieved, the final list of included studies reflects articles accessed through relevant bibliographic databases, in a relevant language, and of the appropriate document type and indexing.

The research focused on journal articles only to ensure consistency in publication type, and only articles were included in the bibliometric database; conference papers and other paper

types were not considered. This limitation should not be mistaken for suggesting that conference papers are in any way inferior to articles published in scientific journals.

Data Preparation and Analysis Software

We created citation, co-citation, co-authorship, and keyword co-occurrence networks by analyzing bibliographic data extracted from the WoS Core Collection using VOSviewer (version 1.6.20). The software used full counting and normalized association strength. Only author-supplied keywords were used in the co-occurrence network, with a minimum of five occurrences for each term. The software performed clustering using the default settings. The analysis identifies the minimum number of occurrences selected for the co-citation and bibliographic coupling networks and the units of analysis applied to each network, and describes them separately.

Methods of Bibliometric Analysis

Trend Analysis of Publication

To investigate the temporal dynamics of the evaluative literature, we examined the number of publications in each year. This approach allowed us to describe trends in publication numbers over time and identify general increases or decreases in scientific output. Because the data were collected on June 21, 2026, the 2026 data are incomplete.

Citation Analysis

To analyze citation patterns, we applied citation analysis to identify articles, authors, journals, institutions, or countries with higher citation counts. We treated citation count as a measure of citation impact within the analyzed sample, not as a measure of the paper's scientific value. We also needed to account for different publication dates and citation potential.

Co-Occurrence Analysis of Keywords

A co-occurrence analysis was performed on keywords to identify associations and connections between concepts and thematic clusters in the evaluations. Keywords with an occurrence rate of at least five were considered. The resulting association and clustering patterns reveal common topics and their relationships. Using this threshold, the results show only the qualifying keywords in the dataset.

Co-Citation Analysis

Co-citation analysis analyzes citations among scientific publications or sources to uncover the intellectual structure of the scientific literature. Co-citation analysis helps identify publications that are cited together in the set of analyzed articles. Interpret the generated network based on the actual citation process, not by attributing a theoretical position to a publication solely because of its position in the network.

Co-Authorship Analysis

Co-authorship analysis has been performed to demonstrate scientific collaboration patterns in the selected papers. This analysis studies collaboration patterns among authors and, when relevant units are considered, can also analyze collaboration among institutions/countries. The findings from this part were used to define the scientific collaboration network and collaboration among units.

Examination of Thematic Evolution and Emerging Topics

The discussion of theme evolution is based on patterns identified in the keyword network and the temporal distribution of topics. In analyzing these patterns, we distinguished between frequent topics, topics appearing at different points in time, and topics with insufficient evidence to determine whether they are emerging. The infrequent occurrence of a keyword, or its complete absence from the keyword network, does not necessarily mean a relevant topic is absent from the research literature; factors such as word choice and the threshold for including keywords in the study can affect this.

Therefore, determining new topics should be based on the analysis of observable data, particularly time and networks. Without data analysis, topics like artificial intelligence, evaluation fairness, and technology ethics cannot be declared new topics in this collection based solely on their relevance to society.

Interpretive considerations and methodological limitations

The study's findings are based on articles from the Web of Science Core Collection, selected based on criteria such as language, SSCI index, article type, time frame (2015 to 2026), and thematic connection to the research topic. Thus, the findings must be analyzed against the background of the selected sample of papers. Papers written in other languages, not covered by the database, or not indexed in other databases may not be included in the final selection. This is especially relevant for analyzing the geographical distribution of scientific publications.

Furthermore, 2026 was not yet complete when the data were collected; therefore, this year's publication count is not appropriate for direct comparison with complete years. The citation number indicates citation power. In addition, because the data were collected before 2026 ended, this year's publication count cannot be used for comparison. The citation number refers to citation strength within the analyzed dataset range and does not alone determine the article's scientific quality. Likewise, Total Link Strength (TLS) reflects a unit's connection strength to others in the network but does not capture all aspects of centrality and scientific influence.

It analyzes the dataset range and does not define an article's scientific quality on its own. Similarly, the Total Link Strength (TLS) value represents a unit's connectivity to other units in the network, but it does not capture all aspects of centrality or scientific influence.

In conclusion, the analysis of themes and emerging trends must account for factors such as search restrictions, selection criteria, analysis thresholds, and database coverage. Hence,

the study's findings reflect only a bibliometric analysis of the articles, not an absolute reflection of all research in formative assessment.

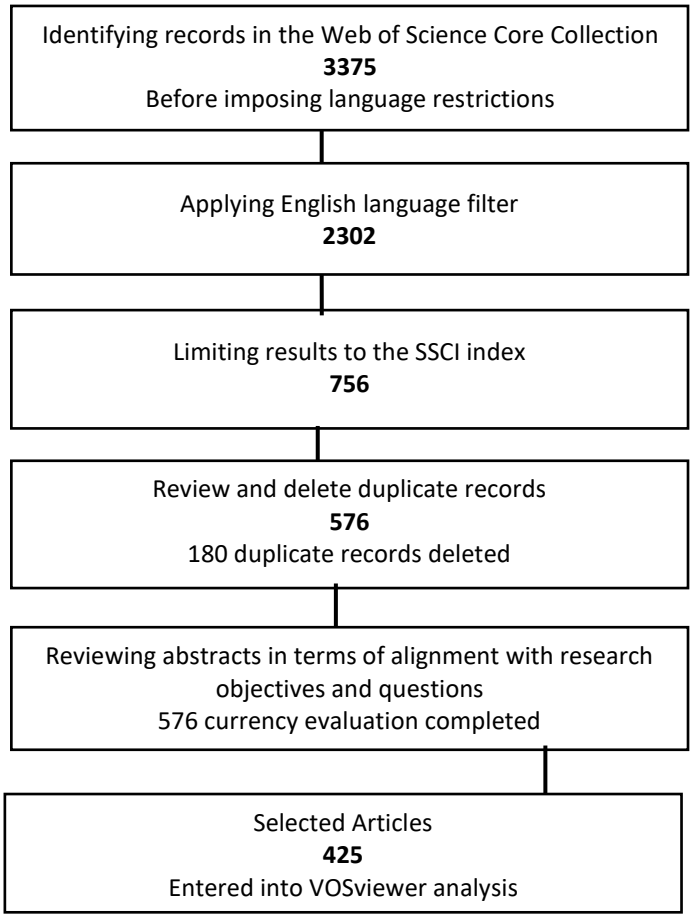


Figure 1. *Flow Diagram of the Article Identification, Screening, and Selection Process*

FINDINGS

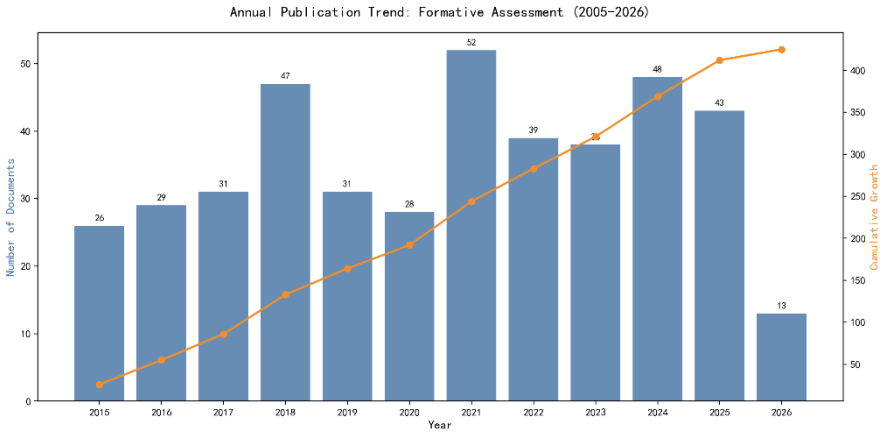


Fig 2. *Annual publication trend of formative assessment articles (2015 – 2026)*

The trend in research publications from 2015-2026 appears positive overall, as most papers were published during this period. The descriptive examination of the publication trends might be summarized into four periods: 2015-2016; 2017-2020; 2021-2023; 2024-

2026. The first period marks the beginning of the trends; the second period highlights moderate development in the research area; the third period (2021-2023) may be more positive, provided the 2021 data are accurate. Approach the last period (2024-2026) with caution, as only 2026 data are fully available, while 2024-2025 data cover only certain months. Overall, the literature database includes 425 articles with 8,411 total citations, averaging 19.8 citations per document. However, it is difficult to argue whether these numbers are high or low because different disciplines use different measures of publication quantity and quality.

Table 1. *Top Authors in This Field - Formative Evaluation*

Rank	Author	TP	TC	ACP	TLS
1	Yan, Z	9	253	28.1	12
2	Palm, T	5	202	40.4	5
3	Furtak, EM	5	143	28.6	12
4	Klieme, E	5	240	48.0	14
5	Andersson, C	4	151	37.8	4
6	Panadero, E	4	351	87.8	10
7	Archila, PA	4	54	13.5	12
8	Gotwals, AW	4	122	30.5	4
9	Hwang, GJ	3	120	40.0	7
10	Clarke-Midura, J	3	18	6.0	8
13	Rakoczy, K	3	129	43.0	8
14	Palmberg, B	3	55	18.3	5

Note. TP = Total Publications; TC = Total Citations; ACP = Average Citations per Publication; TLS = Total Link Strength. TLS represents the cumulative strength of an item's links within the specific bibliometric network analyzed.

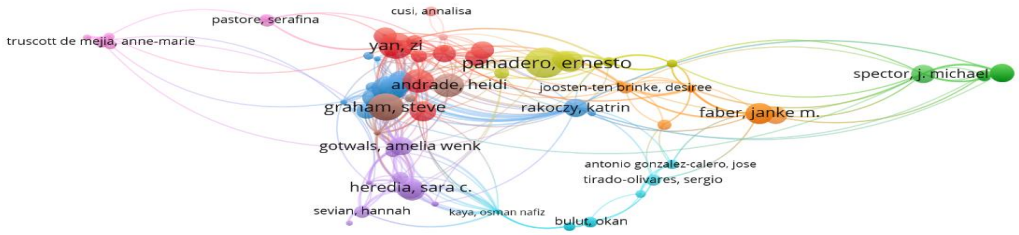


Fig. 3. *Author citation network map*

The table and figure present research results on the most productive authors in assessment studies across three aspects: scientific productivity (TP), citation impact (TC, ACP), and network connectivity (TLS). By publication count, Yan appears to be the most productive author in the chosen field, with nine assessment-related publications. However, by citation impact, Panadero appears to be the most productive researcher, with 351 citations and an ACP of 87.8. Among the remaining researchers, Klieme has the highest TLS (14). Palm, Hwang,

and Rakoczy have higher ACP than the other researchers, suggesting moderate scientific productivity. Based on this analysis, all three authors have a high level of scientific productivity. Yan has the highest TP among the analyzed authors, Panadero has the highest TC and ACP, and Klieme has the highest TLS. Thus, the three authors should be considered the most productive and impactful in the chosen field of assessment. Nevertheless, scientific productivity is not the only determinant of an author's contribution to a topic, and researchers should not be judged solely on those metrics.

Table 2. *Top Institutions in the Field of Formative Assessment*

Rank	Institutions	TP	TC	ACP
1	Education University of Hong Kong (EdUHK)	25	594	23.8
2	Universidad de los Andes (Colombia)	14	154	11.0
3	University of California System	14	432	30.9
4	Michigan State University	11	206	18.7
5	Umea University	11	312	28.4
6	Monash University	10	57	5.7
7	University of Hong Kong	10	498	49.8
8	University of Colorado System	10	430	43.0
9	Maastricht University	9	99	11.0
10	University of Colorado Boulder	9	399	44.3
11	Jilin University	9	51	5.7
12	University of London	9	208	23.1
13	University of Alberta	8	178	22.2
14	State University of New York (SUNY) System	8	283	35.4
15	University of Groningen	8	142	17.8

According to the table above, North America, Europe, and East Asia have the most influential universities in this field. The university with the highest TP is the Education University of Hong Kong, with 25 articles, followed by the University of Colorado Boulder and the University of Colorado, with 22 and 18 articles, respectively. In terms of TC, the University of Hong Kong is also the most productive one with 594 citations. For ACP, the most productive universities are the University of Hong Kong, the University of Colorado Boulder, and the University of Colorado, which have 49.8, 44.3, and 43 citations for each article, respectively, indicating that for the number of published articles and the number of average citations, their tendencies do not coincide.

The leading countries are Hong Kong, the USA, the UK, Sweden, the Netherlands, and Canada. This distribution pattern is explained by the fact that the most productive or most popular (by citations) institutions are concentrated in these regions. The representation of these institutions should not be taken as evidence that the countries not mentioned do not carry out important work in the field of formative assessment. To test this hypothesis, it is necessary to quantitatively compare the number of published articles and citations across countries. The same applies to the role of these organizations in forming theories and methods: testing this hypothesis requires additional studies, such as analyses of co-citations, keywords, or specific content. The productivity and citation patterns of these organizations

differ significantly, providing an opportunity for joint research on the represented organizations.

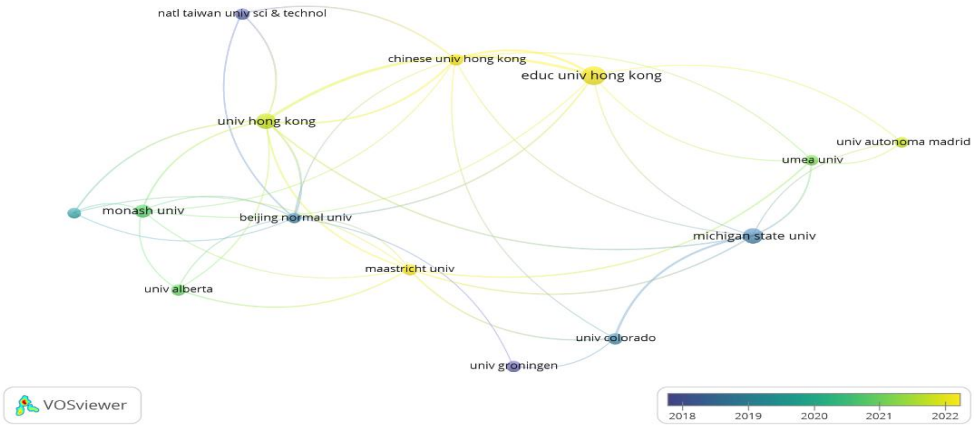


Fig 4. *Citation network map of Institutions*

Table 3. *15 Top Countries in the Field of Formative Assessment*

Rank	Country	TP	TC	ACP
1	United States	132	2855	21.6
2	China	53	1284	24.2
3	Netherlands	32	804	25.1
4	Australia	30	795	26.5
5	Spain	30	816	27.2
6	United Kingdom	29	485	16.7
7	Germany	26	655	25.2
8	Taiwan	17	418	24.6
9	Turkey	13	123	9.5
10	Canada	10	222	22.2
11	Norway	10	118	11.8
12	Switzerland	9	232	25.8
15	South Korea	7	142	20.3

The country analysis shows disparities in productivity, measured by publications and citations in formative assessment research. The United States of America produces the most articles (132 articles and 2,855 citations). It is followed by China (53 articles and 1,284 citations) and the Netherlands, Australia, and Spain, which produce the most publications. However, the country that produces the most articles does not appear to have the highest ACP. Spain has the highest ACP (27.2), followed by Australia (26.5), Switzerland (25.8), Germany (25.2), and the Netherlands (25.1). The list of countries with the highest productivity in quantity does not correspond to the list of countries with the highest ACP. Turkey has the lowest ACP (9.5), followed by Norway (11.8), while the United Kingdom also has a relatively low ACP (16.7).

The national distribution of contributing countries suggests that the most academically productive and cited contributors listed in the table are mainly located in North America, Europe, and East Asia. However, this table shows only a limited range of countries and does

not necessarily imply a geographical gap for formative assessment research between continents. ACP is merely a measure of average citation frequency, not a direct measure of research quality. In sum, the findings differentiate national publication productivity from mean citation impact and show that the two factors need not go hand in hand.

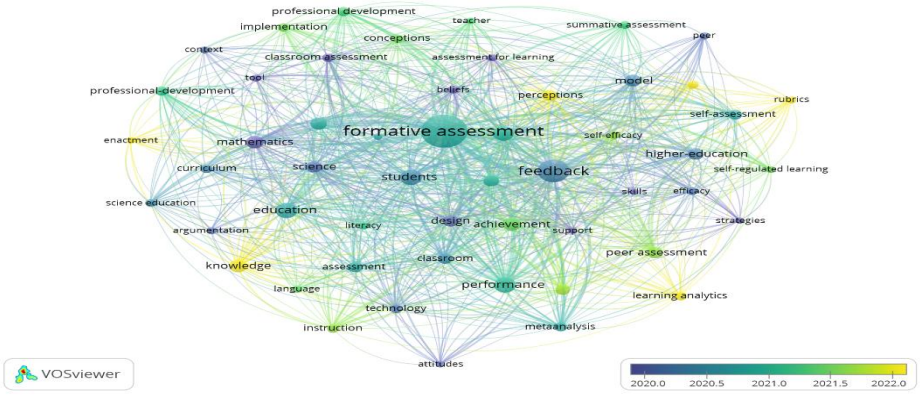


***Figure 5**. International cooperation network of countries*

***Table 4.** Top Co-Cited Authors and Journals in Formative Assessment Research*

Rank	Author	Number of citations	Rank	Journal	Number of citations
1	Anonymous	755	1	ASSESS EDUC	514
2	Black	453	2	ASSESS EVAL HIGH EDU	482
3	Wiliam	184	3	COMPUT EDUC	478
4	Hattie	150	4	REV EDUC RES	377
5	van	149	5	J RES SCI TEACH	371
6	Furtak	137	6	SCI EDUC	288
7	Bennett	136	7	TEACH TEACH EDUC	218
8	Brown	136	8	J EDUC PSYCHOL	180
9	Panadero	126	9	EDUC ASSESS EVAL ACC	169
10	Lee	124	10	INT J SCI EDUC	158

The co-citation analysis highlights several authors and journals worthy of further consideration. As shown in Table 1, Black is the most frequently co-cited author with 453 co-citations, followed by Wiliam (184) and Hattie (150). Other frequently cited authors include van, Furtak, Bennett, Brown, Panadero, and Lee. For each of these individuals, either further study of their work is needed, or their articles are used extensively. In any case, these authors are significant in formative assessment research. However, to fully understand their contributions, readers must consult the published articles, as the current information provides only general descriptions of their work. A content analysis would be required to identify how these authors have contributed specifically to the assessment for learning field and to understand their theoretical backgrounds.

At the journal level, the table shows that the Assessment in Education journal has the highest co-citation frequency (514), followed by Assessment & Evaluation in Higher Education (482) and Computers & Education (478). Other highly co-cited journals include Review of Educational Research, the Journal of Research in Science Teaching, Science Education, and Teaching and Teacher Education. These journals suggest that formative assessment practices draw on literature from educational assessment, educational research, educational

technology, science education, and teacher education. Altogether, the findings of the current study can help identify a set of co-cited journals and authors that can provide a foundation for later-stage comprehensive research in the field. However, the results of this study should not be overinterpreted to suggest that the identified authors and journals are of higher quality than others, and several sources cannot automatically define the field's development or knowledge framework.

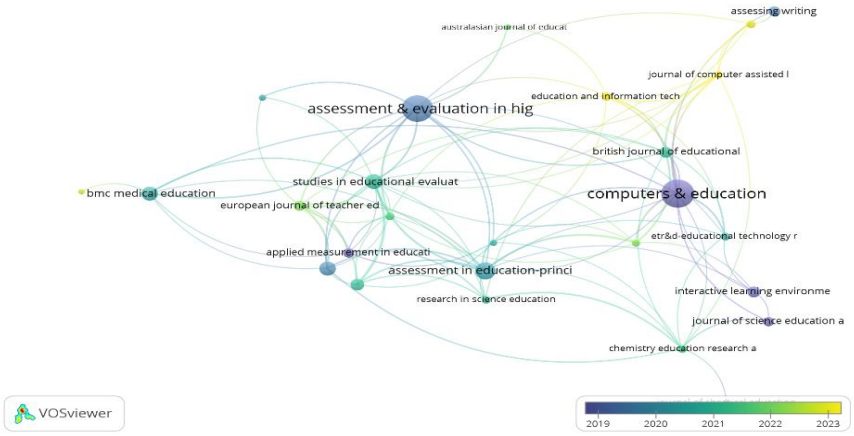


Fig. 6. *Overlay Visualization of Co-Cited Journals in Formative Assessment Research*

Table 5. *Top-Cited Studies and Their Citation Impact in Formative Assessment Research*

Rank	Article Title	Authors	Years	Journal	TC
1	FORMATIVE ASSESSMENT AND WRITING: A Meta-Analysis	Graham, S; Hebert, M; Harris, KR	2015	ELEMENTARY SCHOOL JOURNAL	248
2	Formative and Summative Assessment in the Classroom	Dixson, DD; Worrell, FC	2016	THEORY INTO PRACTICE	217
3	The role of gamified e-quizzes on student learning and engagement	Zainuddin, Z; Shujahat, M; Haruna, H; Chu, SKW	2020	COMPUTERS & EDUCATION	206
4	Fusing self-regulated learning and formative assessment: a review	Panadero, E; Andrade, H; Brookhart, S	2018	AUSTRALIAN EDUCATIONAL RESEARCHER	179
5	The effect of automated feedback on revision behavior and learning	Zhu, MX; Liu, OL; Lee, HS	2020	COMPUTERS & EDUCATION	144
6	The effects of a digital formative assessment tool on mathematics	Faber, JM; Luyten, H; Visscher, AJ	2017	COMPUTERS & EDUCATION	127
7	Formative self-and peer assessment for improved student learning	Wanner, T; Palmer, E	2018	ASSESSMENT & EVALUATION IN HIGHER EDUCATION	127
8	Using Learning and Motivation Theories to Coherently Link Formative Assessment	Shepard, LA; Penuel, WR; Pellegrino, JW	2018	EDUCATIONAL MEASUREMENT-ISSUES AND PRACTICE	113

9	Formative assessment as practice: the role of students' motivation	Leenknecht, M; Wijnia, L; Köhlen, M; Fryer, L; Rikers, R; Loyens, S	2021	ASSESSMENT & EVALUATION IN HIGHER EDUCATION	112
10	Automated writing evaluation for formative assessment of secondary students	Ranalli, J; Link, S; Chukharev-Hudilainen, E	2017	EDUCATIONAL PSYCHOLOGY	107

Citation analysis was used to identify articles on formative assessment. The search was limited to the ten most cited articles. The articles analyzed covered writing assessment, the impact of formative and summative assessment, gamified quizzes, self-regulated learning, automated feedback, electronic-based assessment, peer assessment, student motivation, and automated writing evaluation. The most cited article was 'Formative assessment and writing: A meta-analysis' by Graham, Hebert, and Harris (2015), with 248 citations, and it appeared to be the most influential. It was followed by another article, 'Examining the role of formative and summative assessment in the classroom' by Dixson and Worrell (2016), with 217 citations, and another article on exam evaluation by Zainuddin et al. (2020), cited 206 times. The citation analysis suggests the articles are most frequently cited for their contributions to writing assessment, classroom assessment, and technology-based assessment practices.

The publication years and topicality in the table suggest that the most influential papers cover various aspects of formative assessment, not just one research area. For instance, several articles relate to computers and assessment and automated assessment; they include digital gamified e-quizzing, automated feedback, digital assessment tools, and automated writing evaluation. Other highly cited articles relate to self-regulated learning, peer assessment, student motivation, and learning-assessment connection theory. Computers & Education (3 times) and Assessment & Evaluation in Higher Education (twice) appear more often in the most-cited papers, suggesting they may be among the most influential. In general, the table above also shows that the most influential articles covered various issues in the domain of most cited papers. However, it certainly does not prove the authority and validity of this work, and it does not suggest that these are the fastest (or newest) growing research topics and areas of study.

Table 6. *Emerging Research Trends and Rapidly Growing Topics in Formative Assessment*

Rank	Emerging topics	Course 1	Course 2	Rank	Topics with the greatest growth	Course 1	Course 2	Growth
1	enactment	0	10	1	performance	17	35	+18
2	assessment literacy	0	10	2	knowledge	9	25	+16
3	participation	0	7	3	perceptions	4	19	+15
4	validity	0	6	4	engagement	5	18	+13
5	experiences	0	6	5	learning analytics	2	15	+13
6	challenges	0	5	6	peer assessment	8	19	+11
7	game	0	4	7	instruction	3	14	+11
8	validation	0	4	8	self-efficacy	2	12	+10
9	reliability	0	4	9	gamification	1	9	+8
10	student perceptions	0	4	10	Meta-analysis	6	13	+7

Table 6 below presents the emerging topics and keywords that increased the most in the number of articles in which they appeared over the two periods. With a difference of 10 articles each, enactment and assessment literacy were the most prevalent emerging topics in the second period but were not identified in the first. Other identified high-occurrence keywords are participation (7), validity (6), experiences (6), and challenges (5). In sum, the latter period's literature placed greater emphasis on assessment practices and assessment literacy, learner participation in instruction, and outcome validity. In raw numbers, performance increased the most (17→35; +18), followed by knowledge (9→25; +16) and perceptions (4→19; +15). Engagement (+13), learning analytics (+13), peer assessment (+11), instruction (+11), and self-efficacy (+10) also increased. The increases in learning analytics and gamification indicate growing interest in learning analytics and technology-enhanced assessment. In contrast, increases in peer assessment, engagement, and self-efficacy may reflect growing interest in student agency. We should be cautious about these trends: they describe only changes in how often keywords appeared in the dataset and say nothing about the quality, validity, or depth of conceptual understanding in the work in these areas.

The keyword patterns can also inform unanswered research questions or areas that require deeper investigation. For instance, future studies could explore how assessment literacy, engagement, self-efficacy beliefs, and self-regulated learning interrelate within formative assessment contexts. Learning analytics, gamification, and feedback are also crucial areas to investigate further. For instance, future studies could examine how digital tools support feedback, learning engagement, and the use of assessment evidence for instructional improvement. Artificial intelligence-powered feedback mechanisms, generative AI, fairness, the ethics of digital assessment, and adaptive assessments are also promising areas for further research. However, researchers should not forget that Table 6 represents only formative-assessment-related topics identified in the keyword network. Therefore, further analysis should confirm whether the areas listed above are indeed under-researched. To summarize, Table 6 outlines shifts in trends for certain research topics during the specified timeframe. However, future studies need to examine the theoretical relationships between these topics, their unanswered questions, and their practical implications.

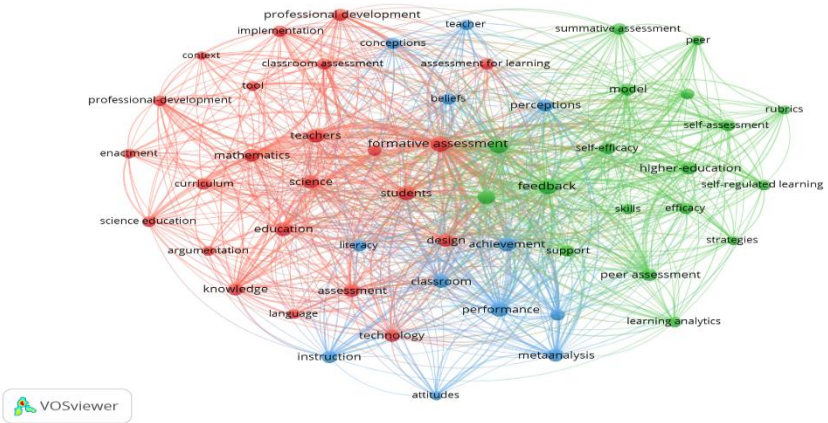


***Fig 7**. Keyword Co-occurrence Network of Research on Formative Assessment and Learning Outcomes*

DISCUSSION

This research examined the distribution of publications, citations, collaboration networks, and article topics on formative assessment published in the Web of Science Core Collection between 2015 and 2026. The results reflect the sample of studies analyzed and indexed in the selected database, not a general statement about the quality and evolution of the entire line of research. The distribution of publications also shows that this line of research is statistically heterogeneous across authors, institutions, and countries. That is, the number and percentage of articles from a country or institution do not necessarily reflect how many studies were conducted on the subject in that country or institution. The present study also includes only articles indexed in the Web of Science database, which may not reflect the global distribution of research on the same topic. Additionally, because all articles analyzed are in English, the results may not apply to research in other languages.

The publication trends may be explained by the previously mentioned articles (Karaçöp & İnaltekin, 2023; Sudakova et al., 2022; Sultan et al., 2024) and other formative assessment and education-related articles. on formative assessment and other education-related articles. However, because other articles differ in selected databases, time period, documents, document selection algorithms, and research problems, a straightforward comparison is not reliable. In particular, the choice of databases, publication years, documents analyzed, and topics presented could significantly influence the results. Therefore, the consistency of the present results with previous studies should be considered in the context of tendencies recognized across different literature collections. Such comparisons should consider the overall number of documents generated each year and their behavior over time. This study shows research behavior on formative assessment within the scope established by the authors. Each study's design methods should explain inconsistencies between the patterns in this paper and those reported by others.

The citation analysis also reveals the intellectual concepts that illustrate how research on formative assessment has evolved from related areas of educational research. The high frequency of feedback, assessment for learning, and self-regulated learning in the analyzed articles may reflect that these closely related concepts are central to discussions of how assessment information supports instruction and learning. Feedback informs learners about their status and provides guidance; assessment for learning underscores the importance of using assessment information to guide further learning; and self-regulated learning emphasizes learners' active role in regulating their learning. Together, these concepts might serve as a theoretical construct for conceptual linkages in the literature. However, the bibliometric linkages should not be understood as reflecting similar processes of using feedback and assessment information across different educational settings, nor as causal effects of feedback or assessment processes on learning. As the research states, the interpretation aligns neatly with the predominant perspectives on assessment and learning strategies espoused by Brown (2019), Yan and Boud (2021), and Schellekens et al. (2021).

Still, to conclude whether the fit is perfect, it is necessary to examine the researcher's facts and statements that justify their position.

Moreover, keyword analysis highlights the range of subjects covered in the sample. As shown in Table 6, assessment literacy increased from 0 to 10, engagement from 5 to 18, and learning analytics from 2 to 15. Peer assessment increased from 8 to 19, self-efficacy from 2 to 12, and gamification from 1 to 9, respectively. In absolute terms, the most-used topics were performance (17-35), knowledge (9-25), and perceptions (4-19). The literature covered aspects ranging from learning results and engagement to assessment literacy, learning analytics, learner engagement, and technology use. This allows researchers to examine how these aspects affect formative assessment goals and practice. This trend, however, does not suggest either a large amount of research or a change in magnitude in education. Again, researchers may have been working in different subjects at different levels, new terms may have been created, and more articles may have been published each year.

These keyword patterns are not, by themselves, a research gap. However, increased focus on assessment literacy and self-efficacy can be used as the basis for questions such as 'what is the relationship between teachers' knowledge of assessment and learners' beliefs about learning and how does this influence formative assessment' or 'what is the relationship between assessment literacy and classroom assessment practices'. Likewise, increased focus on learning analytics and gamification can inform questions such as 'what do learners think about how digital technologies are incorporated into assessment and how do we assess the effectiveness of this incorporation'. Empirical research to be done in the near future may include questions such as 'does feedback support self-regulated learning', 'under what conditions', 'is assessment literacy associated with assessment practice' or 'do digital assessment devices improve timeliness and quality of feedback'. While these questions are suggested based on the bibliometric findings, a list of keyword frequency counts does not allow us to judge their answerability.

Other potential avenues for further work are artificial intelligence-aided assessment, adaptive assessment and personalized feedback. However, their apparent scarcity in these analyses, if taken as evidence of commonality in keyword- and publication-level references to these topics, would not necessarily indicate that these topics are lacking, underdeveloped, or nascent in the literature as a whole. Many relevant works may use different terminology or may not be captured in the database and/or selection criteria used in this study. A complete account could involve temporal keyword analysis, a review of relevant publications, and possibly comparisons with other databases. For these reasons, they remain potential lines of inquiry rather than evidenced gaps.

Using all the indicators presented in the publication, citation, collaboration, and keyword analyses allows for an understanding of trends in research on formative assessment in the examined Web of Science database. While publication and citation indicators describe general trends in volume and impact, collaboration analysis reveals the extent to which researchers in the field cooperate, and co-citation and keyword co-occurrence analyses

identify topics and trends. These measures serve different purposes, however, and are not interchangeable as measures of research quality or impact. For example, citation behavior doesn't necessarily reflect sound assessment behavior, and while keyword relationships indicate the language used in the research, they don't necessarily reveal causal relationships between educational constructs. This paper is a first step toward pulling these perspectives together and providing broad strokes of the structure and themes of the body of literature. This characterization can inform research questions for a subsequent study; however, further qualitative, quantitative, or theoretically based research is needed to explore the process behind this pattern, as well as the educational function of any given assessment technique.

CONCLUSION

This study mapped the development and knowledge structure of formative assessment research published in the Web of Science Core Collection during 2015-2026. Using publication and citation analysis, co-authorship analysis, co-citation analysis, and keyword co-occurrence analysis, the study revealed trends in research output, collaboration, intellectual background, and research emphases in the selected literature. In this way, the results obtained provide an organized view of the contributors to this research area, as well as recurring concepts in the citations and keywords used in the research literature. The novelty of this study lies in applying these perspectives to show how formative assessment research is reflected and structured in the chosen dataset. Rather than proving that the field has matured or that emerging technologies have revolutionized the assessment process, the results provide a basis for bibliometric interpretation of research patterns.

The results open up several opportunities for future research and educational practice. The prevalence of concepts such as feedback, assessment for learning, and self-regulated learning in the selected literature provides an opportunity for future empirical research on the connections among these three components of formative assessment. The presence of assessment literacy and learner engagement in the list of topics identified through keyword analysis will also support further research on teacher assessments and student engagement in the formative assessment process. Furthermore, the presence of digital assessment-related topics in the collected dataset provides an opportunity to examine technologies used to implement formative assessment. However, the effectiveness of these tools must be estimated empirically rather than inferred from publication data. The identified structure of scholarly outputs and cooperation can inform further research on collaborations and the involvement of institutions from different countries in research activities. Based on the bibliometric analysis, these aspects should be considered research priorities, not measures with proven effectiveness.

This research has multiple limitations. First, this paper used data extracted from the Web of Science Core Collection, so publications cited in other databases were not included. Therefore, the results reflect the chosen database, not the full scope of formative assessment research. Second, citation-based indicators can give older publications an advantage because

they have had more time to accumulate citations, while newer studies have not yet received sufficient attention from scholars. Third, limiting the dataset to English-language literature may have omitted studies in other languages. Fourth, the data were gathered on 21 June 2026, so 2026 publications are incomplete and should not be compared with complete annual records. Finally, bibliometric indicators and visualizations provide insights into publication patterns, citation rates, collaborations, and conceptual connections, but they do not prove research quality, assessment efficiency, or the reasons behind trends.

Acknowledgements

The authors acknowledge using ChatGPT (OpenAI, GPT-5.5 version) for assistance with translation and linguistic refinement of the manuscript. The authors also used Grammarly to support grammar, spelling, and language editing. These tools were used solely to enhance the manuscript's clarity, readability, and quality. The authors are responsible for the research design, data analysis, interpretations, and conclusions.

Funding Information

No funding was received for this study.

Statement of Data Availability

Data supporting the findings of this study are available upon request from the corresponding author.

Conflict of Interest Statement

The authors declare that they have no financial, personal, or organizational conflicts of interest related to the subject of this article

REFERENCES

Acuña, A. B., & de Dávila, R. C. (2021). Management of formative assessment and feedback. *Revista Iberoamericana de Educación*. https://doi.org/10.31876/ie.vi.137

Al Breiki, M., & Yahaya, W. A. J. W. (2026). The impact of a gamified formative electronic assessment on students' academic performance: A meta-analysis. *International Journal of Human–Computer Interaction*, *42*(5), 3044–3061. https://doi.org/10.1080/10447318.2025.2531270

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

Banihashem, S. K., Gašević, D., & Noroozi, O. (2025). A critical review of using learning analytics for formative assessment: Progress, pitfalls and path forward. *Journal of Computer Assisted Learning*, *41*(3), e70056. https://doi.org/10.1111/jcal.70056

Brown, G. T. (2019). Is assessment for learning really assessment? *Frontiers in Education*.

Esomonu, N. P.-M., & Nduka, S. V. (2025). META-ANALYSIS OF THE EFFECT OF FORMATIVE ASSESSMENT ON STUDENTS' ACADEMIC ACHIEVEMENT IN SCIENCE SUBJECTS IN NIGERIA. *International Journal of Premium Advanced Educational Research*, *1*(4), 71–81.

Karaçöp, A., & İnaltekin, T. (2023). Formative assessment in educational research published at the beginning of the new millennium: Bibliometric analysis. *Journal of Formative Design in Learning*, *7*(2), 106–125. https://doi.org/10.1007/s41686-023-00081-9

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*. https://doi.org/10.1016/j.lindif.2023.102274

Nguyen, N. Q., & Nguyen, L. P. (2026). Formative assessment for knowledge cultivation in university EFL classrooms: A systematic review from higher education. *Language Testing in Asia*, *16*(1), 22. https://doi.org/10.1186/s40468-026-00430-y

Prompiengchai, S., Narreddy, C., & Joordens, S. (2025). A Practical Guide for Supporting Formative Assessment and Feedback Using Generative AI. *arXiv*, (arXiv:2505.23405). https://doi.org/10.48550/arXiv.2505.23405

Ren, H., Zhang, W., & Cao, H. (2026). Formative assessment as a self-regulation engine in blended language learning: A case study with top-tier engineering talents. *Acta Psychologica*, *264*, 106505. https://doi.org/10.1016/j.actpsy.2026.106505

Schellekens, L. H., Bok, H. G., De Jong, L. H., Van der Schaaf, M. F., Kremer, W. D., & Van der Vleuten, C. P. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, *71*, 101094. https://doi.org/10.1016/j.stueduc.2021.101094

Sudakova, N. E., Savina, T. N., Masalimova, A. R., Mikhaylovsky, M. N., Karandeeva, L. G., & Zhdanov, S. P. (2022). Online formative assessment in higher education: Bibliometric analysis. *Education Sciences*, *12*(3), 209. https://doi.org/10.3390/educsci12030209

Sultan, J., Ulwiyah, S. A., Retnawati, H., & Kastara, R. (2024). *Bibliometric analysis: Scopus indexed formative assessment learning research trends in the last decade*. International Conference on Current Issues in Education.

TING, T. L. (2025). *IMPACT OF FORMATIVE ASSESSMENT ON NURSING STUDENTS' LEARNING OUTCOMES AND CLINICAL SKILLS DEVELOPMENT: A SYSTEMATIC REVIEW AND META-ANALYSIS*.

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

Vassiliou, S., Papadima-Sophocleous, S., & Giannikas, C. N. (2022). *Formative assessment in second language learning: A systematic review and an annotated bibliography*.

Webb, M., & Galamba, A. (2026). Transforming pedagogy with GenAI-supported formative assessment: Challenges for teacher education. *British Journal of Educational Technology*, *57*(3), 690–706. https://doi.org/10.1111/bjet.70051

Wu, J., & Yu, X. (2025). The influence of formative assessment on academic performance: Exploring the role of teachers' emotional support. *Frontiers in Psychology*, *16*, 1567615. https://doi.org/10.3389/fpsyg.2025.1567615

Yan, Z. (2021). Assessment-as-learning in classrooms: The challenges and professional development. *Journal of Education for Teaching*, *47*(2), 293–295. https://doi.org/10.1080/02607476.2021.1885972

Yan, Z., & Boud, D. (2021). Conceptualising assessment-as-learning. In *Assessment as learning* (pp. 11–24). Routledge. https://doi.org/10.4324/9781003052081-2

Yılmaz, S. S. (2025). A Bibliometric Analysis of Formative Assessment Research in Science Education. *International Journal of Education and Artificial Intelligence*, *2*(1), 17–40. https://doi.org/10.5281/zenodo.16615094

Zhang, Y., Wang, W., Xian, Y., Wang, X., & Huang, J. (2023). The Research Status of Formative Assessment in Science Education. *Journal of Baltic Science Education*, *22*(6), 1103–1119.