



Asya Studies

Akademik Sosyal Araştırmalar / Academic Social Studies

Year: 10, Number: 37, p. 1-35, Autumn 2026

Assessment and Evaluation in Maths Education: The Current Scientific Productivity, Conceptual and Thematic Networks, and Trends*

Matematik Eğitiminde Ölçme ve Değerlendirme: Alandaki Mevcut Bilimsel Verimlilik, Kavramsal ve Tematik Ağlar ile Eğilimler

ISSN: 2602-2877 / E-ISSN: 2602-263X

Research Article
Araştırma Makalesi

Article Arrival Date
Makale Geliş Tarihi
11/03/2026

Article Accepted Date
Makale Kabul Tarihi
11/09/2026

Article Publication Date
Makale Yayın Tarihi
30/09/2026

Asya Studies

Assoc. Prof. Dr. Ahmet Bozak
Hatay Mustafa Kemal University, Faculty of
Education, Department of Educational
Sciences, Hatay-TÜRKİYE
abozok77@hotmail.com
ORCID: [0000-0001-5275-0514](https://orcid.org/0000-0001-5275-0514)

Assoc. Prof. Dr. Abdulkadir Özkaya
Hatay Mustafa Kemal University, Faculty of
Education, Department of Sciences
Education, Hatay-TÜRKİYE
kayakadir78@gmail.com
ORCID: [0000-0002-6962-4597](https://orcid.org/0000-0002-6962-4597)

Eray Selçuk
Teacher / Ministry of National Education,
Hatay-TÜRKİYE
erayselcuk84@gmail.com
ORCID: [0000-0003-4033-4219](https://orcid.org/0000-0003-4033-4219)

Abstract

The study aims to reveal the current scientific productivity, conceptual and thematic networks, and trends in educational research on Assessment and Evaluation of Maths Education (AEoME) through a bibliometric analysis of 1303 articles from the Web of Science and Scopus databases. The findings reveal that the AEoME field showed consistent growth in publications from the 2000s onwards. However, an international standard for the AEoME has yet to be established, as ongoing discussions reflect a shift from purely summative to more formative approaches based on conceptual understanding, reasoning, and problem-solving. The conceptual networks on teacher competencies remain crucial, and the field focuses on integrating technology into innovative approaches, such as flipped classrooms and new digital methods. While exams like PISA and TIMSS on mathematical literacy introduce a global perspective, the AEoME are also seeking a new process, shifting from a traditional assessment to a modern digital perspective. Recent theoretical discussions highlight that such a transformation should also be restructured towards artificial intelligence (AI) and further digital technologies. Addressing the geographical gap in AEoME publications requires broader cooperation between developed and underrepresented countries in Africa and Central Asia.

Keywords: Maths Education, Assessment and Evaluation, Bibliometrics, Conceptual Networks and Trends, Scientific Trends

Öz

Bu çalışma, Web of Science (WoS) ve Scopus veritabanlarından elde edilen 1303 makalenin bibliyometrik analizi yoluyla Matematik Eğitimi'nde Ölçme ve Değerlendirme (MEÖD) alanındaki güncel bilimsel üretkenliği, kavramsal ve tematik ağları ve eğitim araştırmalarındaki eğilimleri ortaya koymayı amaçlamaktadır. Bulgular, MEÖD alanının 2000'lerden itibaren yayınlarda sürekli bir büyüme gösterdiğini ortaya koymaktadır. Temel kavramlar küresel sınavlara odaklansa da literatür, MEÖD için uluslararası ortak bir standardın henüz tam olarak oluşturulmadığına işaret etmektedir; zira devam eden tartışmalar, tamamen özetleyici yaklaşımlardan kavramsal anlayış, akıl yürütme ve problem çözmeye dayalı daha biçimlendirici yaklaşımlara doğru bir kaymayı yansıtmaktadır. Öğretmen yeterliliklerine ilişkin kavramsal ağlar kritik önem taşımaktadır ve alan, tersine çevrilmiş sınıflar ve yeni dijital yöntemler gibi yenilikçi yaklaşımlara teknolojiyi entegre etmeye odaklanmaktadır. PISA, TIMSS gibi matematiksel okuryazarlık sınavları küresel bir bakış açısı sunarken, MEÖD'de geleneksel değerlendirmeden modern dijital bir bakış açısına geçmeye odaklı yeni bir süreç arayışındadır. Güncel teorik tartışmalar, bu tür bir dönüşümün yapay zekâ (YZ) ve diğer dijital teknolojilere yönelik olarak da yeniden yapılandırılması gerektiğine dikkat çekmektedir. MEÖD yayınlarındaki coğrafi dengesizliği gidermek için, gelişmiş ülkeler ile Afrika ve Orta Asya'daki az temsil edilen ülkeler arasında daha yaygın bir iş birliğine ihtiyaç duyulmaktadır.

Anahtar Kelimeler: Matematik Eğitimi, Ölçme ve Değerlendirme, Bibliyometrik, Kavramsal Eğilimler ve Ağlar, Bilimsel Eğilimler

* Ethics committee approval is not required for this research.

Citation Information / Atıf Bilgisi

Bozak, A., Özkaya, A. & Selçuk, E. (2026). Assessment and Evaluation in Maths Education: The Current Scientific Productivity, Conceptual and Thematic Networks, and Trends. *Asya Studies*, 10(37), 1-35. <https://doi.org/10.31455/asya.1919608>

INTRODUCTION

Measurement and evaluation in education, especially in mathematics, are often regarded as important subjects by researchers, not only for field research but also for accurately assessing learning outcomes, identifying potential strengths and weaknesses, and designing developmental activities based on these findings. This section presents the conceptual and theoretical framework of measurement and evaluation, issues related to mathematics, new perspectives, and the problem statement and research questions of the current study.

The Conceptual and Theoretical Background

The concepts of “assessment and evaluation” are often treated as synonyms, as distinct concepts, or as twin concepts. However, in education, they are used to check the quality of teaching and learning processes. In other words, seek what can be done to improve the educational outcomes in a specific field. Assessment includes various instructional measures to determine individuals' attributes. Various information about students' learning levels and the identification of strengths and weaknesses to support students' academic progress. On the other hand, evaluation is the process of judging or determining value and comparing it with a certain standard; in other words, it is more focused on grading and the final determination of quality and quantity in a process. In addition, assessment is a diagnostic and formative process that identifies potential areas for improvement by providing feedback on performance throughout the learning process, while evaluation is a more judgmental and summative process that focuses on outcomes, provides an overall grade, and ascertains whether standards are met (Howard & Donaghue, 2014; Kellaghan & Stufflebean, 2003; Yambi & Yambi, 2018).

In an efficient assessment and evaluation process, other key concepts and considerations include measurement quality, testing, achievement, validity, and reliability. Measurement is the process of determining the quantitative or qualitative academic outputs of an individual or a group. A test, through written, observation or oral means, is assessed by measuring the knowledge or ability of an individual via certain appropriate criteria. In such a process, those means, criteria, and standards are considered complementary concepts for assessment and evaluation. Furthermore, the concepts of validity and reliability also entail several issues in the true assessment and evaluation of outcomes, including objectivity, comprehensiveness, practicality, comparability, and utility as a measurement or testing tool, which should be both valid and reliable (Callingham et al., 2009; Kubiszyn & Borich, 2024; Weir, 2005; Yambi & Yambi, 2018).

In AEoME literature, there are several types of assessment, such as formative vs summative, informal vs formal, continuous vs final, process vs product, and divergent vs convergent. However, two central approaches are commonly used in the literature. The first approach focuses on assessing learning through summative judgements, formal examinations and tests, using reliable and objective measures, while the second approach emphasises a formative perspective through feedback to improve performance (Callingham et al., 2009; McAlpine, 2002; Kubiszyn & Borich, 2024).

Assessment and Evaluation in Math Education

The problems affecting students' achievement and performance in math education are deemed a critical issue in many countries. These problems are often correlated with socioeconomic factors as well. Therefore, it should be revealed how mathematics can be taught more effectively across diverse settings and circumstances through authentic and equitable teaching approaches (Boaler, 2002; Boaler & Staples, 2008; Moses & Cobb, 2001). AEoME, which has primarily concentrated on achievement results, has shifted towards enhancing learning outcomes, national and international tests, and comparative studies (Black & Wiliam, 2012; Nortvedt & Buchholtz, 2018). The assessment perspective has been criticised for many years,

prompting a focus on alternative methods to highlight the field. Ruthven (1994) argued that math assessment should focus more on emerging trends such as the authenticity and realism of assessment tasks; the interpretive quality of assessment information; and the integration of the processes of teaching, learning, and assessment at the beginning of the 1990s. More recently, Radmehr & Vos (2020) suggest a new 20th-century perspective on math assessment, focusing on higher-order thinking rather than the traditional approach that relies on knowledge, low cognitive demands, and tests students on memorised calculation procedures.

In addition, with increasing digitalisation, AEoME needs to evolve to reflect this trend, as well as new digital measurement instruments and assignments (Drijvers, P. (2018); Radmehr & Vos, 2020). Artificial Intelligence (AI) is increasingly utilised in educational assessment, transforming traditional paradigms through automation. This also necessitates educators to adopt and become familiar with more effective digital pedagogical approaches and practices (Owan et al., 2023; Meylani, 2025). Furthermore, AI-powered tools, such as Intelligent Tutoring Systems and adaptive feedback mechanisms, are likely to improve the objectivity and responsiveness of AEoME, fostering a dynamic learning environment, engagement, and critical thinking (Bayrak & Aslanci, 2022).

Measuring learning outcomes in the mathematics curriculum is also crucial for many countries, both in response to changing needs and demands at the national and international levels and in line with assessment norms such as PISA (Program for International Student Assessment). Furthermore, to keep up with global education standards for the math curriculum. This necessitates the active involvement of researchers and educational stakeholders with respect to changing global needs (Addey, 2017; Borch et al., 2020; Eryani et al., 2024). Therefore, the current broader perspective, informed by systematic reviews in the field, is thought to reveal potential gaps and contribute to math education at the national and international levels.

The Problem Situation

Bibliometric analyses are considered an efficient method for examining scientific productivity, citation relationships, academic collaborations, and methodological trends in a specific research area (Aria & Cuccurullo, 2017). Despite a few bibliometric studies on more specific topics in math education (Hasumi & Chiu, 2022; Toker & Aksoy, 2024), there is a lack of systematic reviews, particularly on AEoME. Bibliometric analyses are of great importance in identifying gaps in the research on AEoME and guiding future studies. Furthermore, bibliometric analyses can help identify which topics receive more attention in educational research on AEoME and which areas require further research, enabling researchers to better understand current trends in AEoME and to improve educational policies (Bayrak & Aslanci, 2022).

While previous systematic studies have generally focused on specific themes, this research holistically analyses educational research on AEoME by conducting a multidimensional assessment of a broad dataset. Therefore, the current study aims to provide a roadmap for future research and contribute to the development of educational policies on AEoME. Although significant progress has been made in educational research on AEoME in recent years, our study reveals that substantial gaps remain in the literature. Moreover, these gaps are likely to hinder the development of assessment and evaluation policies and practices in math education. Therefore, we argue that addressing these gaps is critical for a more comprehensive understanding of the educational processes in AEoME.

Therefore, the study addresses the following research questions:

1. How have the publication volume and citations in AEoME research evolved over the years?

2. What are the most influential sources, countries, and collaboration networks shaping the AEoME field?
3. What are the current conceptual and thematic changes and trends in AEoME research?

METHOD

The following section details the study's research type, data collection methods, PRISMA flowchart, and bibliometric analysis process.

Literature is deemed to be the accumulation of studies conducted in a particular field over time. Examining the data in a portion or all these studies is known as bibliometric analysis (İnceoğlu, 2014). Bibliometrics is also deemed to be a powerful tool for reducing the complexity of the scientific landscape and evaluating the impact of scientific studies. Such analyses are quantitative examinations of publications produced by individuals or institutions in a specific field, period, and region, and of the relationships among these publications. These analyses also address the structural developments and impacts of scientific literature (Zupic & Cater, 2015).

The current research employs the Biblioshiny application and the R-based bibliometrix package (Aria & Cuccurullo, 2017) for data analysis. Bibliometric analysis techniques are mainly used for performance analysis and science mapping. While performance analysis focuses on evaluating the research elements, science mapping reveals the relationships among research data. The study preferred R programming as it provides a flexible and robust syntax (Hackenberger, 2020; Tjøstheim et al., 2022). Furthermore, R programming supports a variety of analyses, ranging from classical statistics to science mapping, machine learning, and data visualisation (Tjøstheim et al., 2022; Cobo et al., 2011). In addition, such an analysis is highly effective in demonstrating the research dynamics, changes, innovations, and trends within a field, thereby helping to objectify the situation under investigation (Glanzel, 2003).

Donthu et al. (2021), summarise the following steps for conducting a bibliometric analysis: (1) Defining the aims and scope of the study around bibliometric data; (2) Choosing an appropriate bibliometric analysis technique; (3) Collecting the data in the literature for the bibliometric analysis; (4) Conducting the bibliometric analysis and evaluating the results through visual elements. Accordingly, we conducted the current research through the following steps

Data Collection, the research keywords, and the inclusion & exclusion criteria

The study's data collection involves several steps using the Scopus and WoS platforms. First, the study analysed the titles, abstracts, and keywords of articles within the scope of Math education and included the targeted articles in the corpus by filtering for specific keywords relevant to the assessment and evaluation of math education. The study deems that the detailed keywords in the search below (student assessment; student evaluation; student achievement; student performance; student measurement; measurement of student learning; student learning measurement; learning outcomes measurement; achievement measurement; performance measurement; classroom assessment; classroom evaluation; classroom-based assessment; classroom assessment practices; classroom measurement; classroom-based measurement; educational measurement; learning measurement; mathematics education; math education; geometry; algebra; calculus) reflect central and fundamental concepts, as well as to elucidate the core themes and major research directions in the assessment and evaluation of math education. Such a comprehensive analysis within the targeted sample emphasises the standardised, globally recognised terminology in the context of math education, ensuring consistency and interoperability in bibliometric analysis as well.

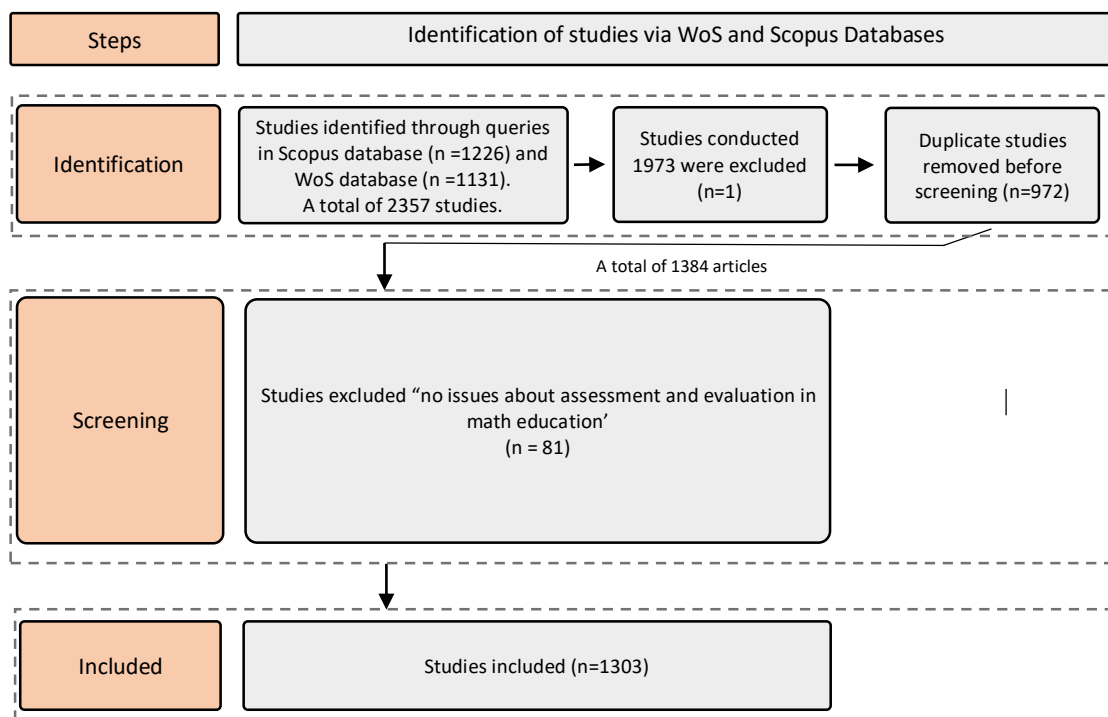
Accordingly, the study has sought the following terms in scholarly articles through Scopus and WoS databases: "teaching; learning; math; mathematics, math education; geometry; algebra; calculus; student assessment; student evaluation; student achievement; student performance; student measurement; measurement of student learning; student learning measurement; learning outcomes measurement; achievement measurement; performance measurement; classroom assessment; classroom evaluation; classroom-based assessment; classroom assessment practices; classroom measurement; classroom-based measurement; educational measurement; learning measurement; scale validation; scale development; instrument development; test validation; test construction; test development"

The literature search was conducted in January 2026, covering articles from 1983 to 2026. Although we collected data in the first month of 2026, we also intentionally included an additional 17 articles already indexed in the databases as "Early Access" or published in the early January 2026 issues to capture the most up-to-date trends in the field. The study considered eligibility and selection criteria, including inclusion and exclusion parameters, within the scope of the research. While articles on AEoME using the relevant keywords were included in the sample, other themes, apart from AEoME, were excluded from the dataset.

Data Processing and Analysis

The study included 1131 articles from the WoS database and 1226 from the Scopus database. The total number of articles was 1303 across both databases after removing duplicates. The study also followed the PRISMA (Moher et al., 2009) flowchart criteria presented in Figure 1. Accordingly, after the initial literature review, a total of 1054 studies were excluded from the initial data pool (972 duplicates and 82 based on eligibility criteria). In the final evaluation phase, 1303 studies were included in the research review.

Figure 1: Data Selection Criteria through the PRISMA Flowchart



Bibliometric analyses are conducted using R Studio software, based on the bibliometrix tool and the `biblioshiny()` function. This function enables data extraction from various sources, including the WoS and Scopus datasets, and offers a comprehensive set of features for bibliometric analysis. This approach allows the research to generate various bibliometric indicators, including publication and citation counts and the H-index. The `biblioshiny()` function facilitates comparative analyses and visualisation of results through interactive graphs, such as network maps, diagrams, and geographic maps, and supports exporting these results.

The analyses in the study are based on visualisations of the distribution of sources according to Bradford's Law and of researchers' publication rates according to Lotka's Law. In addition, keyword word clouds, trend graphs of research topics by year, network graphs revealing collaborations among authors, sources, and countries, and three-field graphs of the relationships among keywords, sources, and authors were used.

FINDINGS

The current research offers a comprehensive evaluation of scientific publications in the WoS and Scopus databases using keywords and themes in AEoME. The targeted data include parameters such as article titles, author information, citation counts, publication dates, and abstracts, obtained through big-data analysis. Consequently, the findings are considered to influence future research directions for educators. The study also employs various statistical methods, including quantitative data on the number of articles, publication dates, and specific keywords, to assess and evaluate mathematics education.

1. The Findings on RQ1 (How have the publication volume and citations in AEoME research evolved over the years?)

a) The Distribution of Citations by Year

Table 1 displays the citation performance and publication volume of the studies from 1983 to 2026. Accordingly, the number of publications was quite limited in the early years (1983–1995) (N=1–4), but the average citation values per article were relatively high in subsequent years. Specifically, studies published in 1986 (MeanTCperArt=66.00) and 1989 (MeanTCperArt=53.50) seem to have had a noticeable impact on the literature. Notably, in 1996, the average total citation was 112.60, and the average annual citation was 3.63, indicating a strong influence of studies published during this period on the AEoME.

Our analysis reveals a gradual increase in publications from the 2000s onwards, with a more noticeable rise after 2010. For example, there were 63 publications in 2015, compared to 74 in 2017, 98 in 2022, and 118 in 2024. However, our study notes a relative decline in the average number of citations per article in recent years. This is likely because the time required for newly published studies to accrue citations has not yet elapsed. Indeed, the low average citation values in 2025 and 2026 (0.83 and 0.35, respectively) are due to the shorter citable-year periods in those years. Furthermore, the 2026 data represent a partial year, consisting of only 17 early-access or January publications available at the time of data retrieval. Nonetheless, we observed that the AEoME research field has expanded rapidly, especially since the 2000s.

Table 1: The Distribution of Citations by Year (1983-2026)

Year	MeanTCperArt	N	MeanTCperYear	CitableYears
1983	12,00	1	0,27	44
1984	0,00	1	0,00	43
1986	66,00	1	1,61	41
1989	53,50	2	1,41	38
1990	8,50	2	0,23	37
1992	6,50	4	0,19	35
1993	33,00	2	0,97	34
1994	12,67	3	0,38	33
1995	32,00	1	1,00	32
1996	112,60	5	3,63	31
1997	38,50	6	1,28	30
1998	20,00	2	0,69	29
1999	32,67	6	1,17	28
2000	25,75	8	0,95	27
2001	49,45	11	1,90	26
2002	30,00	9	1,20	25
2003	27,60	5	1,15	24
2004	28,38	8	1,23	23
2005	43,08	12	1,96	22
2006	22,43	14	1,07	21
2007	28,73	22	1,44	20
2008	35,57	30	1,87	19
2009	35,11	27	1,95	18
2010	23,61	18	1,39	17
2011	30,29	34	1,89	16
2012	40,46	37	2,70	15
2013	21,32	40	1,52	14
2014	23,57	60	1,81	13
2015	26,73	63	2,23	12
2016	16,36	55	1,49	11
2017	23,03	74	2,30	10
2018	18,40	63	2,04	9
2019	10,69	71	1,34	8
2020	13,32	71	1,90	7
2021	12,21	66	2,04	6
2022	8,55	98	1,71	5
2023	6,09	93	1,52	4
2024	3,91	118	1,30	3
2025	0,83	143	0,42	2
2026	0,35	17	0,35	1

N =number of publications

MeanTCperArt = average total citations per article

MeanTCperYear = average total citations per year

We observed that the publication output has increased significantly since 2010, and early studies tend to have a higher citation impact due to their longer circulation periods. In this context, Figure 2 depicts the distribution of articles published by year.

Figure 2: Total Number of Articles Published Annually (1983-2026)

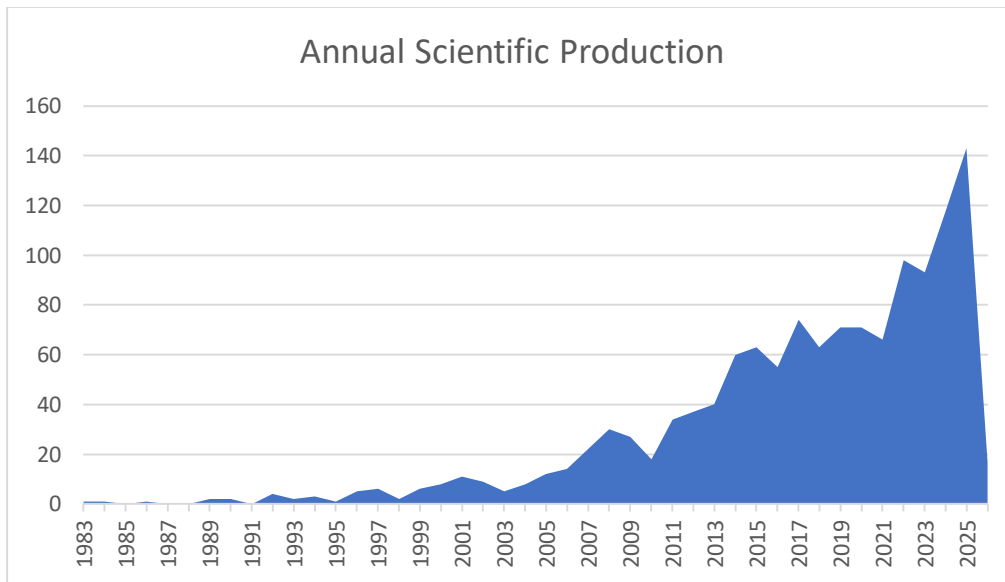


Figure 2 shows that the number of publications was quite low between the 1980s and 1990s. Our study found that scientific output gradually grew from the 2000s, with a more noticeable rise after 2010, and the number of publications increased rapidly in the 2020s, and the field expanded considerably.

b) Scholarly Outputs of the Articles

As shown in Table 2, our dataset comprises 1303 scholarly articles published across 501 sources. This annual growth rate of 6.81% indicates a steady increase in scientific output in the research field. The average document age is 8.71 years, while the mean number of citations per article is 16.44. Additionally, the dataset contains 1411 Keywords Plus and 3181 author keywords, reflecting the conceptual diversity of the topics.

Our study found that 3251 authors contributed to the research field, with 237 studies being single-authored and an average of 3.01 co-authors per article. Furthermore, the international co-authorship rate was calculated at 8.67%. However, we limited the document types solely to articles (n=1,247), including 14 early-access articles and 42 review articles.

Table 2: Scientific Output from 1983 to 2026.

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	1983:2026
Sources (Journals, Books, etc)	501
Documents	1303
Annual Growth Rate %	6,81
Document Average Age	8,71
Average citations per doc	16,44
DOCUMENT CONTENTS	
Keywords Plus (ID)	1411
Author's Keywords (DE)	3181
AUTHORS	
Authors	3251
Authors of single-authored docs	221
AUTHORS COLLABORATION	
Single-authored docs	237
Co-Authors per Doc	3,01
International co-authorships %	8,672
DOCUMENT TYPES	
Article	1247
Article; early access	14
Review article	42

c) The most Prolific Authors in the Field

Analysing the most prolific authors in the field not only identifies scientific contributions but also reveals research trends, academic collaboration networks, and the evolution of scientific performance over time. Such analyses identify researchers who have shaped the field's development, highlighting research areas of focus and potential opportunities for academic collaboration. Furthermore, this allows for a clearer understanding of the developmental dynamics of scientific output in the field of AEoME. Accordingly, the analysis of the most prolific authors on AEoME is presented in Figure 3 and Table 3.

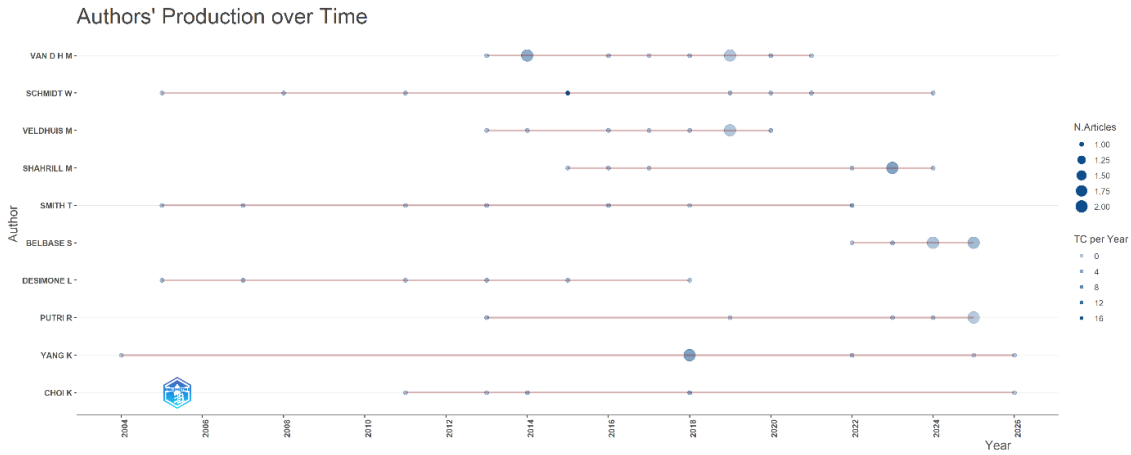
Figure 3: Scientific Output of the Most Prolific Authors by Year (1983–2026)

Figure 3 illustrates the leading authors in the field of AEoME over the years and the annual citation impact of their publications. In the graph, the horizontal axis shows the years, while the vertical axis displays the authors; additionally, the size of the points reflects the number of articles published each year, and the colour intensity indicates the annual citation rate. The results reveal that some authors maintain a consistent publication output over many years, whereas others tend to concentrate their output in specific periods. Notably, researchers such as Schmidt W., Smith T., and Yang K. have made long-term academic contributions, while some authors, such as Belbase S., Putri R., and Shahrill M., have more recent publication records.

Further information on the authors' publication output is provided in Table 3, which includes the number of articles published by each author in specific years (freq), the total number of citations received by these publications (TC), and the average annual citation rate (TCpY).

Table 3: Total Number of Papers and Citations of the Most Prolific Authors.

Author	year	freq	TC	TCpY
BELBASE S	2022	1	1	0,2
BELBASE S	2023	1	9	2,25
BELBASE S	2024	2	3	1
BELBASE S	2025	2	4	2
CHOI K	2011	1	20	1,25
CHOI K	2013	1	9	0,643
CHOI K	2014	1	38	2,923
CHOI K	2018	1	35	3,889
CHOI K	2026	1	0	0
DESIMONE L	2005	1	36	1,636
DESIMONE L	2007	1	44	2,2
DESIMONE L	2011	1	19	1,188
DESIMONE L	2013	1	32	2,286
DESIMONE L	2015	1	21	1,75
DESIMONE L	2018	1	2	0,222

PUTRI R	2013	1	36	2,571
PUTRI R	2019	1	2	0,25
PUTRI R	2023	1	3	0,75
PUTRI R	2024	1	3	1
PUTRI R	2025	2	0	0
SCHMIDT W	2005	1	16	0,727
SCHMIDT W	2008	1	33	1,737
SCHMIDT W	2011	1	27	1,688
SCHMIDT W	2015	1	199	16,583
SCHMIDT W	2019	1	13	1,625
SCHMIDT W	2020	1	8	1,143
SCHMIDT W	2021	1	12	2
SCHMIDT W	2024	1	3	1
SHAHRIILL M	2015	1	11	0,917
SHAHRIILL M	2016	1	14	1,273
SHAHRIILL M	2017	1	7	0,7
SHAHRIILL M	2022	1	5	1
SHAHRIILL M	2023	2	21	5,25
SHAHRIILL M	2024	1	3	1
SMITH T	2005	1	36	1,636
SMITH T	2007	1	44	2,2
SMITH T	2011	1	19	1,188
SMITH T	2013	1	32	2,286
SMITH T	2016	1	33	3
SMITH T	2018	1	2	0,222
SMITH T	2022	1	26	5,2
VAN D H M	2013	1	13	0,929
VAN D H M	2014	2	52	4
VAN D H M	2016	1	17	1,545
VAN D H M	2017	1	6	0,6
VAN D H M	2018	1	14	1,556
VAN D H M	2019	2	3	0,375
VAN D H M	2020	1	19	2,714
VAN D H M	2021	1	4	0,667

Freq: Frequency (Number of publications by the author in a specific year)

TC: Total Citations (Total number of citations by the author)

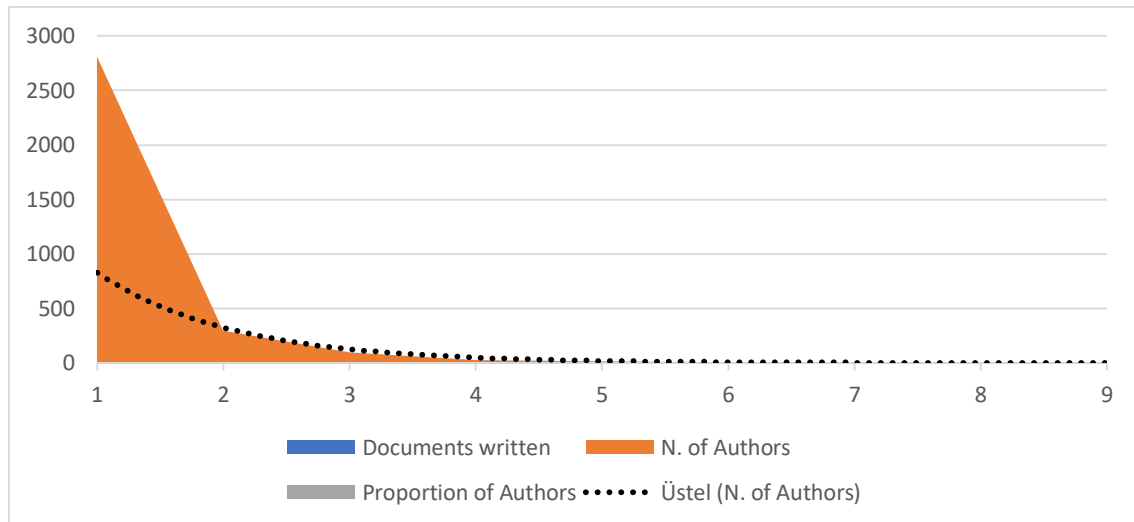
TCpY: Average Citations Per Year (Average number of citations received by the author per year).

The findings in Table 3 reveal that some authors exhibited high output and achieved high citation impact during certain periods. For example, Schmidt W. consistently produced publications between 2005 and 2024, and his work published in 2015, with a total of 199 citations and an average annual citation value of 16,583, is one of the publications showing the highest

impact in the table. Similarly, Choi K. stands out with his publications between 2011 and 2018, with his 2018 publication receiving 35 citations and an average annual citation value of 3,889. The similar citation values of the studies published by Smith T. and Desimone L. between 2005 and 2013 suggest that these authors may share common research lines or collaborative relationships within the field. In more recent years, researchers such as Belbase S., Putri R., and Shahrill M. have increased their publication output and made new contributions to the field, especially in the period after 2022. Furthermore, it is noteworthy that Van D. H. M. consistently produced publications between 2013 and 2021, and in 2014, with two publications totalling 52 citations, he made a significant impact.

The findings also disclose that scientific output in the field is concentrated among a few core researchers, and some studies achieve high citation impact, particularly in certain years. The frequency distribution (Lotka's law) of the scientific productivity of the most prolific authors is presented in Figure 4.

Figure 4: Distribution of Scientific Productivity by Number of Authors (Lotka's Law)



Lotka is one of the essential bibliometric models that reveal the frequency distribution of authors' scientific productivity. Figure 4 shows the distribution of scientific productivity in AEoME according to Lotka's law. This method also explains the common productivity pattern, where many authors publish only a few papers, while a small number produce many publications. Lotka's law is considered an important model for explaining the distribution of phenomena and scientific productivity. Accordingly, findings based on Lotka's law are presented in Table 4, which systematically displays authors' productivity.

Table 4: Authors' Output through Lotka's Law.

Documents written	N. of Authors	Proportion of Authors
1	2811	0,865
2	299	0,092
3	95	0,029
4	24	0,007
5	13	0,004
6	4	0,001
7	2	0,001
8	2	0,001
10	1	0

Table 4 reveals that many authors in the field have published only one article. Indeed, 86.5% of the total 2811 authors produced only one publication. On the other hand, 299 authors (9.2%) published two articles, and 95 (2.9%) published three. Our findings reveal that the number of authors decreases rapidly as the number of publications increases. For example, authors who published four articles accounted for 0.7%, those who published five articles accounted for 0.4%, and those who published six or more articles accounted for an even lower share.

Surprisingly, our study observed that only one author published ten articles. Such a distribution shows that scientific production is carried out by a small number of researchers, while other researchers' contributions were relatively limited to the field. The findings reveal that author productivity in the field of AEoME largely conforms to the inverse-proportional distribution predicted by Lotka's law. The h-index values for the authors are given in Figure 5.

Figure 5: Academic Impact of Authors through H-index Scores.

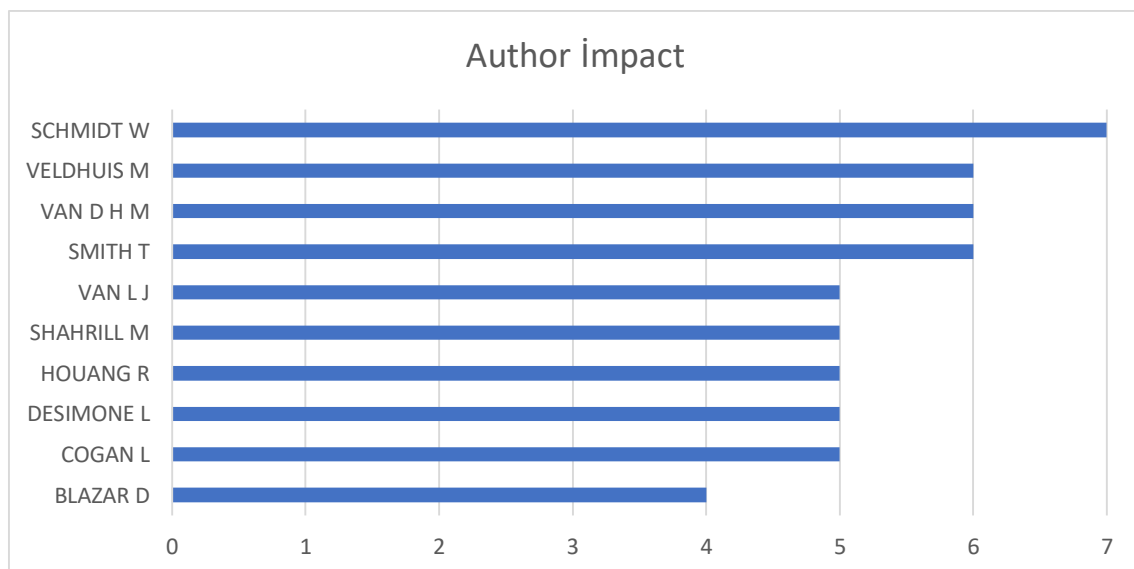


Figure 5 shows that the author, Schmidt W., tops the list with a score of 7, making him the researcher with the highest academic impact; this score is followed by Veldhuis M., Van D.

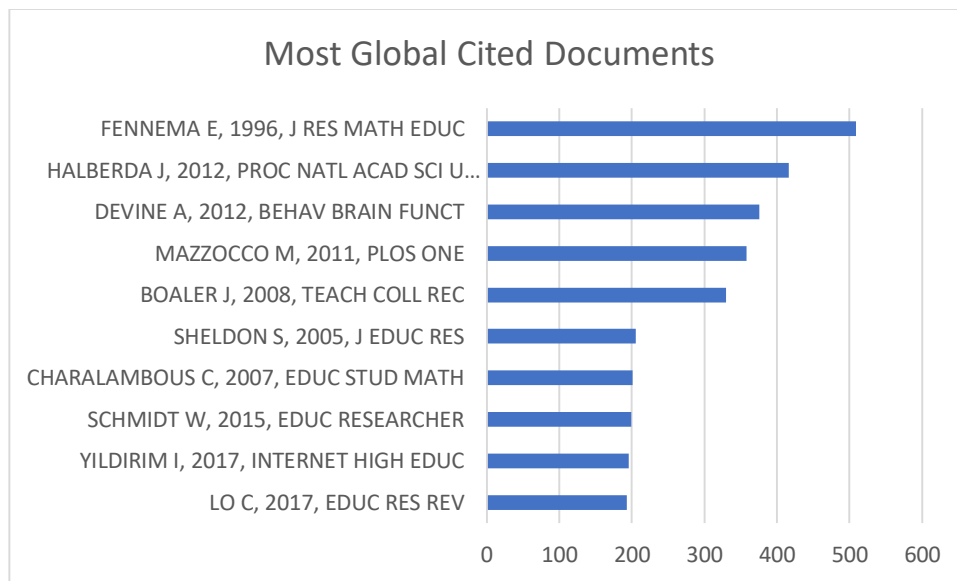
H. M., and Smith T., each with 6 points. Van L. J., Shahrill M., Houang R., Desimone L., and Cogan L., who make up most of the researchers, cluster in the middle group with 5 points each, while Blazar D. has 4 points. Overall, the data shows that the h-index values range from 4 to 7.

2) Findings on the RQ2 (What are the most influential sources, countries, and collaboration networks shaping the AEoME field?)

a) The articles with the highest global citations

The impact level of the articles is determined by the number of citations and local and global citation rates, determining the importance of the relevant studies in the field. The studies with the highest citation counts are displayed in Figure 6.

Figure 6: The total Number of the Most Cited Articles between 1983 and 2026

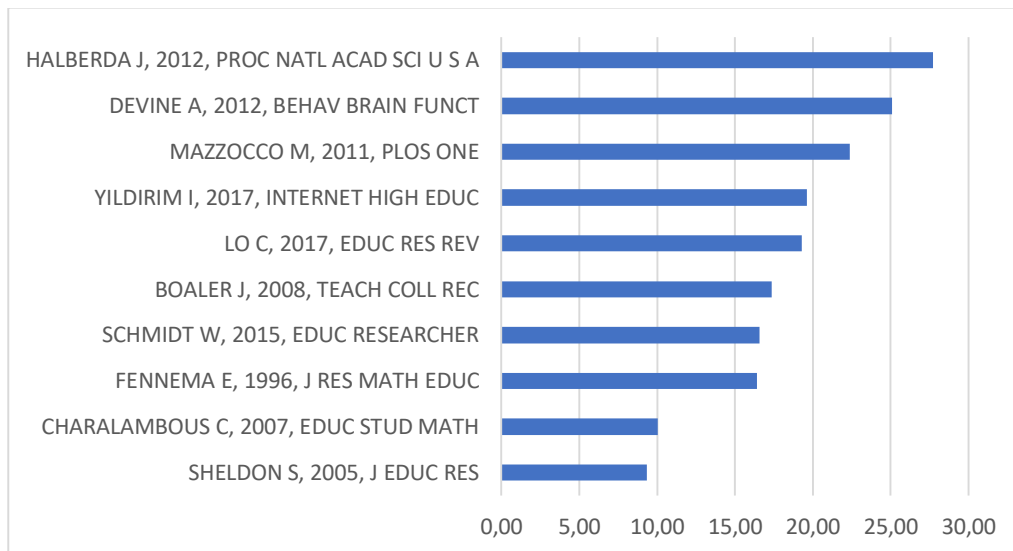


The international citation count data presented in Figure 6 reveal that the study by Fennema E. (1996) has the most dominant impact in the literature, with over 500 citations. This is followed by Halberda J. (2012), which exceeds the 400 citations, and Devine A. (2012), Mazzocco M. (2011), and Boaler J. (2008), which are in the 300-400 range, constituting the high-impact studies in the field. In addition, the works of authors such as Sheldon S. (2005), Charalambous C. (2007), Schmidt W. (2015), Yildirim I. (2017), and Lo C. (2017) are closely clustered, with approximately 200 citations. The fact that recent studies, particularly those from 2017, have quickly reached this level of citation reflects the popularity of these topics in the current literature, while the 1996 study, which maintains its top position, demonstrates continuity in the field's theoretical foundations.

Nonetheless, our study observed that the articles in Figure 7 exhibit a different distribution of average annual citation rates than the total citation ranking in Figure 6. E.g. Fennema E. (1996), which is at the top in total citation count, falls behind in annual average (approximately 16 citations) due to its long presence in the literature; while Halberda J. (2012) with approximately 28 citations per year and Devine A. (2012) with 25 citations per year stand out as the studies with the highest momentum in the current literature. In particular, the studies by Yildirim I. and Lo C., published in 2017, with an average of 20 citations per year, indicate that these articles have been cited directly since publication and reflect current trends in the field. On

the other hand, studies such as those by Sheldon S. (2005) and Charalambous C. (2007) exhibit more stable navigation, with annual citation rates of 10 or fewer.

Figure 7: The Annual Citing Rates of the Articles between 1983 and 2026.



b) The Publication Performance of Scientific Journals over the Years

Understanding the dynamics of research resources and references is crucial for grasping the development of the relevant literature and changes in publications over time, as it provides strategic data on current academic knowledge and informs emerging research areas. Accordingly, the findings on the resource-based distribution of publications in AEoME, as well as the development of these resources, are presented in Table 5.

Table 5. The Distribution of Publications According to Journals and Resources

Year	International Journal of Mathematical Education in Science and Technology	International Journal of Science and Mathematics Education	Journal for Research in Mathematics Education	Education Sciences	Education Studies in Mathematics	Eurasia Journal of Mathematics, Science and Technology Education
1983	0	0	0	0	1	0
1984	1	0	0	0	1	0
1985	1	0	0	0	1	0
1986	1	0	0	0	1	0
1987	1	0	0	0	1	0
1988	1	0	0	0	1	0
1989	1	0	0	0	1	0
1990	1	0	0	0	1	0
1991	1	0	0	0	1	0

1992	1	0	0	0	1	0
1993	1	0	1	0	1	0
1994	1	0	1	0	1	0
1995	1	0	1	0	1	0
1996	1	0	2	0	1	0
1997	1	0	3	0	1	0
1998	1	0	3	0	1	0
1999	2	0	3	0	1	0
2000	2	0	4	0	1	0
2001	2	0	6	0	1	0
2002	2	0	7	0	1	0
2003	2	0	8	0	1	0
2004	2	0	8	0	1	0
2005	2	0	9	0	2	0
2006	2	1	13	0	2	0
2007	2	2	16	0	3	1
2008	4	2	20	0	3	3
2009	7	3	20	0	3	4
2010	8	6	20	0	3	4
2011	9	8	20	0	4	4
2012	10	8	20	0	5	4
2013	12	9	21	0	5	4
2014	15	11	21	0	9	4
2015	16	16	21	0	12	4
2016	21	18	22	0	12	4
2017	23	20	22	0	13	4
2018	24	24	22	2	13	5
2019	25	24	22	4	13	5
2020	27	26	22	4	13	6
2021	29	28	23	7	13	6
2022	34	30	23	9	15	9
2023	38	32	23	12	16	12
2024	40	32	23	17	18	16
2025	44	35	24	22	19	20
2026	44	35	24	22	20	20

Table 5 reveals that the main journals contributing to the mathematics education literature, AEoME, have diversified over time. In the 1980s and 90s, journals such as ‘Educational Studies in Mathematics’ and ‘International Journal of Mathematical Education in Science and Technology’ were pioneers, while from the early 2000s onwards, ‘Journal for Research in Mathematics Education’ saw a significant increase in cumulative publication volume. After 2010, we observed a general acceleration in the literature; in particular, ‘International Journal of Mathematical Education in Science and Technology’ and ‘International Journal of Science and Mathematics Education’ stand out as the journals with the highest production volume, with 44 and 35 publications, respectively, projected for 2026.

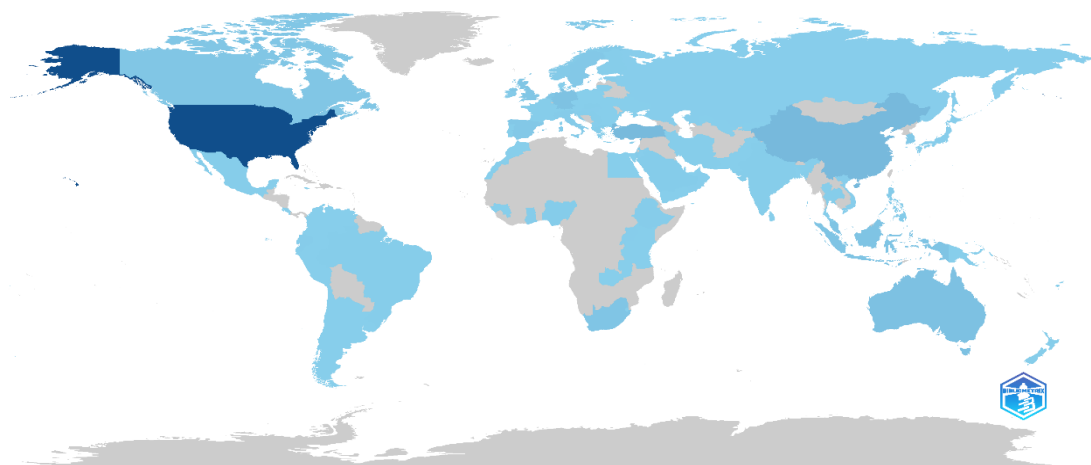
On the other hand, the rapid rise of journals like ‘Education Sciences’, which were included in the literature later, and their reaching a significant volume after 2018, reflects the impact of open access and current publishing trends on the field. Overall, the data show that research on AEoME has undergone a remarkable expansion in both source diversity and publication density over the last 15 years.

c) Scientific Productivity in Publication by Countries

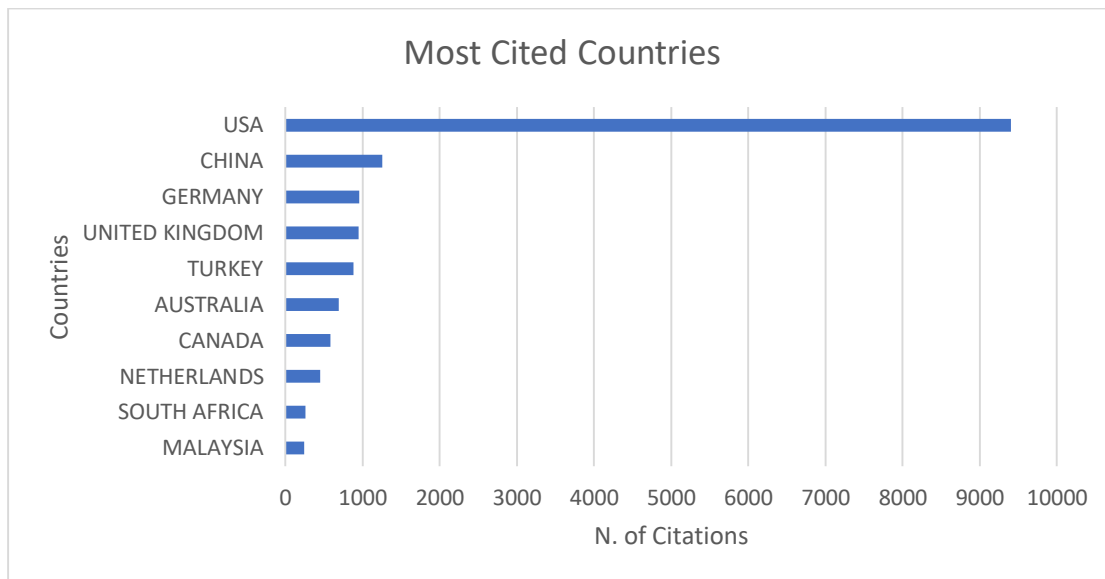
Analysing the distribution of publications across continents and countries can also make significant contributions to interpreting the findings related to the subject under investigation. The total output of countries in AEoME is displayed in Figure 8.

Figure 8: The Scientific Outputs in Publications by Countries (1983-2026)

Country Scientific Production



Our study has observed that the USA ($f=1065$) has been the most productive, with absolute dominance, publishing approximately 6.5 times as many publications as its closest followers. This is followed by China ($f=164$) and Turkey ($f=159$). Surprisingly, despite Turkey ranking third, this surpasses that of developed countries such as Germany, Canada, and the Netherlands. The colour distribution also reveals strong clustering along the North American, Asia-Pacific, and European lines, while many regions in Africa and Central Asia (grey areas) are underrepresented in the literature, pointing to significant geographical polarisation worldwide. Findings on the countries with the most citations are shown in Figure 9.

Figure 9: Most Cited Publications by Countries.

Our findings in Figure 9 reveal that the USA, 9,000 citations, has been the leading country in total citations by countries as well, while Turkey, previously ranked third in total output, has remained among the top five in citation rankings; however, this points out a relatively lower impact for international citation counts than for total publications.

These findings in Figures 8 and 9 reveal that countries such as the United States, China, and Germany have been dominant and central in the field, and also indicate a certain geographical clustering that parallels their output power.

d) The Global Publication Cooperation across the Countries

International collaborations are likely to enhance academic research by increasing diversity in specific areas and providing a broader, multicultural perspective across fields of study. The current study has also examined the global collaborations in the AEoME field and reached the following findings.

Figure 10: Countries' Collaboration Map

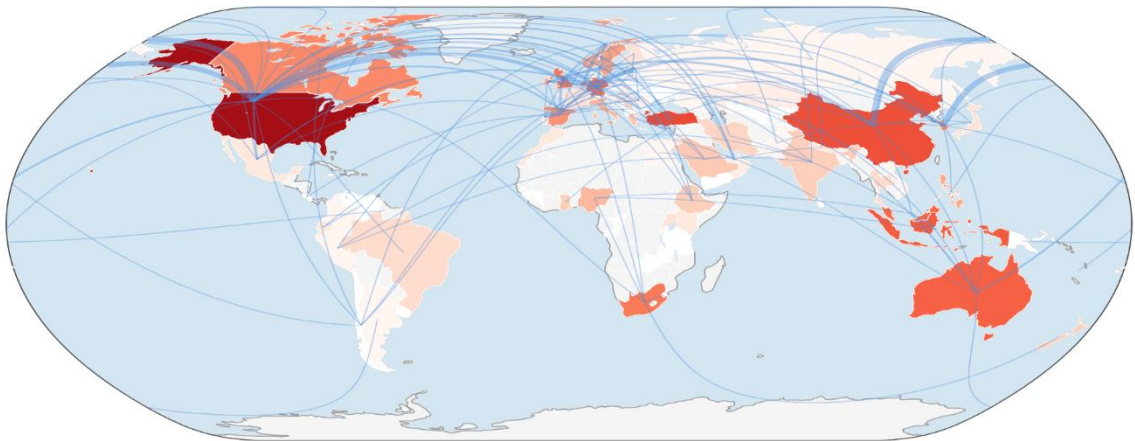


Table 6: Countries with the Most Publication Cooperation in the AEoME Field.

From	To	Frequency
USA	CHINA	13
USA	KOREA	10
USA	UNITED KINGDOM	6
CHINA	UNITED KINGDOM	3
GERMANY	USA	3
NETHERLANDS	NORWAY	3
USA	AUSTRALIA	2
USA	FRANCE	2
USA	MEXICO	2
USA	TURKEY	2

The findings in Table 6 on international collaboration networks signal the existence of a strategic global axis; in particular, the high intensity of cooperation between the United States and China ($f=13$) and South Korea ($f=10$) on modern assessment approaches and evaluation standards, as well as technologically focused applications. The intense red tones and central connecting lines in the map indicate that the critical themes, such as validity, rubric design, and formative assessment in mathematics education, have been fostering regional and intercontinental

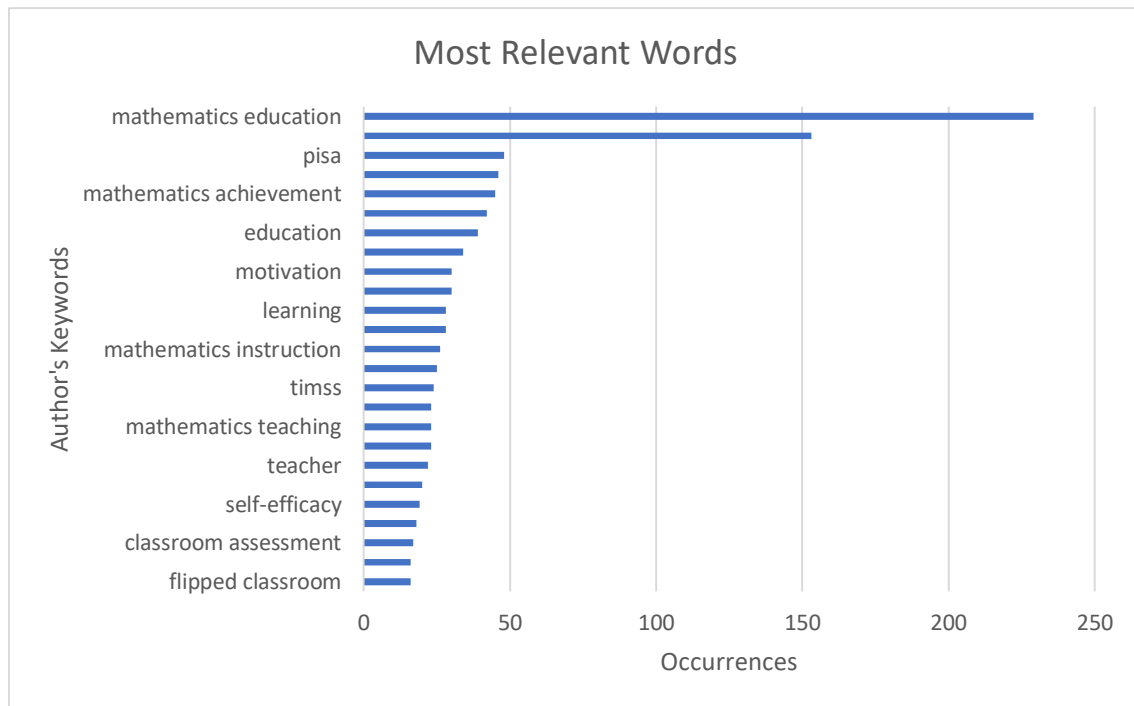
intellectual focus, especially along the North America-Asia-Europe axis. Accordingly, our study reveals that innovative assessment and evaluation models and statistical processes in the field have become a global standard in the literature through close academic networks and co-authored studies among these leading countries.

3) Findings on RQ3 (What are the Current Conceptual and Thematic Changes and Trends in AEoME research?)

a) The Current Key Concepts of AEoME Research

Keyword analysis is used to identify the conceptual focus in AEoME and to understand current conceptual and theoretical changes and trends in the literature, as well as across interrelated sub-disciplines. Such an analysis identifies the processes under investigation, such as the transition from traditional measurements to modern digital assessment technologies and highlights potential research gaps through niche themes that have not received sufficient attention in the literature.

Figure 11: Most Frequently Used Author Keywords in Articles (1983-2026)



Our findings in Figure 11 reveal the crucial points clustered around the term "mathematics education". When analysed from an assessment and evaluation perspective, the high ranking underscores that international exams such as "PISA" and "TIMSS" reflect global achievement standards and enable comparative analyses. In addition, keywords such as "motivation", "self-efficacy", and "flipped classroom", along with the concept of "mathematics achievement", reveal a resilient, evolving, multidimensional structure encompassing student motivation, perceptions of self-efficacy, technology-supported next-generation approaches to digital pedagogy, and academic achievement outcomes.

Figure 13: Tree Map of the Keywords in AEoME (1983-2026)

Our findings in Figure 13 reveal the proportional weights and hierarchical distribution of the key concepts in literature between 1983 and 2026. The frequencies and percentages within the dataset disclose the main focuses of research in the following headings:

1. Center and Core Terms: "Mathematics education" (9%) and "mathematics" (8%) are the most frequently used terms. This confirms that the publications are directly structured around the core concepts of the mathematics education discipline.

2. Measurement, Evaluation, and Success: The terms such as "achievement" (5%), "performance" (3%), and "student achievement" (2%) show that the main motivation of the literature is to measure student outcomes and analyse academic success in the field.

3. International Standards and Literacy: The concepts such as "PISA" (2%), "TIMSS" (1%), and "mathematical literacy" (1%) prove that research is structured around international achievement criteria and the ability to apply mathematics to daily life (Math literacy) rather than local achievement.

4. Pedagogical Approaches and Interventions: The terms "instruction" (2%), "curriculum" (2%), and "teaching" (1%) indicate that the transfer of mathematical knowledge, the structure of curricula, and classroom practices are crucial research themes.

5. Affective and Technological Dimensions: Keywords such as "motivation" (2%), "self-efficacy" (1%), and "technology" (1%) indicate that modern mathematics education focuses not only on cognitive skills but also on student psychology and the role of digital tools in the learning process. Such a finding demonstrates that the research on AEoME continues to evolve across a wide range of themes.

b) The Contextual Outlook of AEoME Research around the Current Networks, Themes, and Trends

Identifying trending topics in a specific academic research field not only advances research in that field but also significantly improves its academic productivity by guiding researchers towards more niche, research-worthy topics. So, our study presents a trend analysis through keywords in AEoMe. The relevant data is shown in Figure 14.

Figure 14: Trending Topics in AEoME (2001-2025)

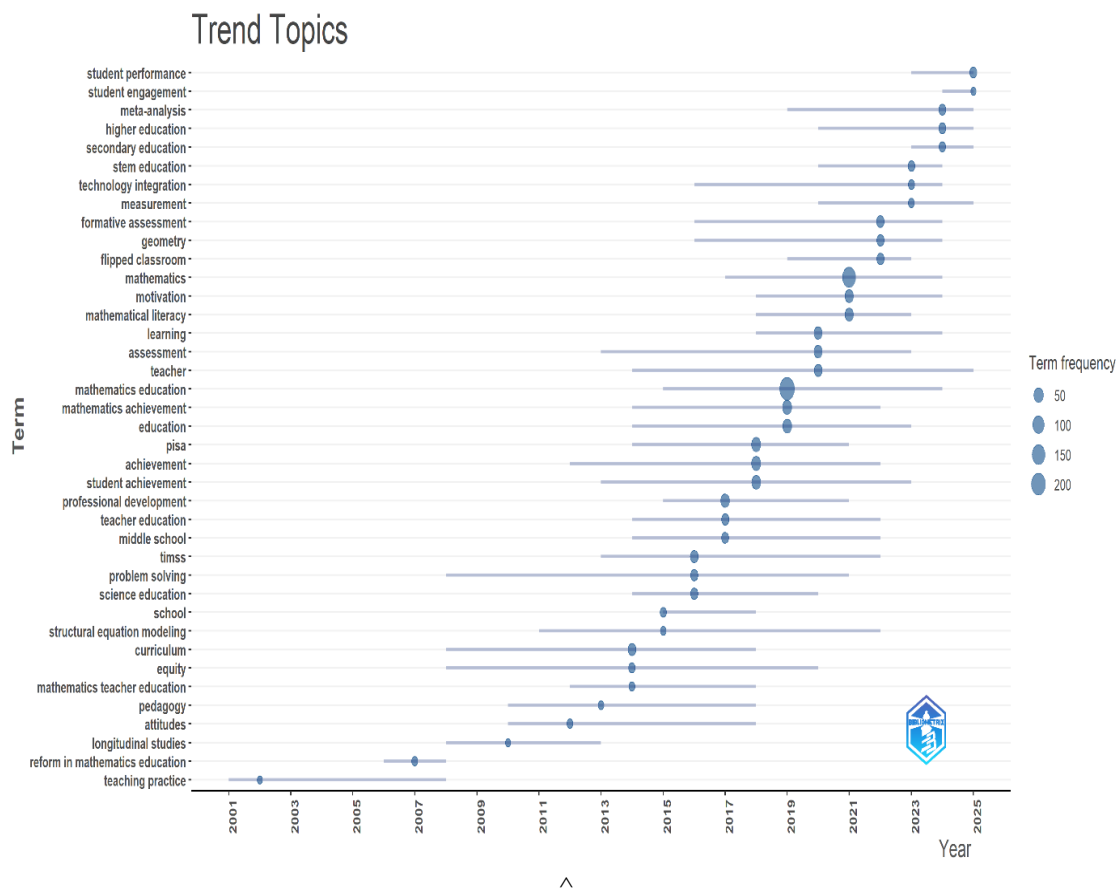
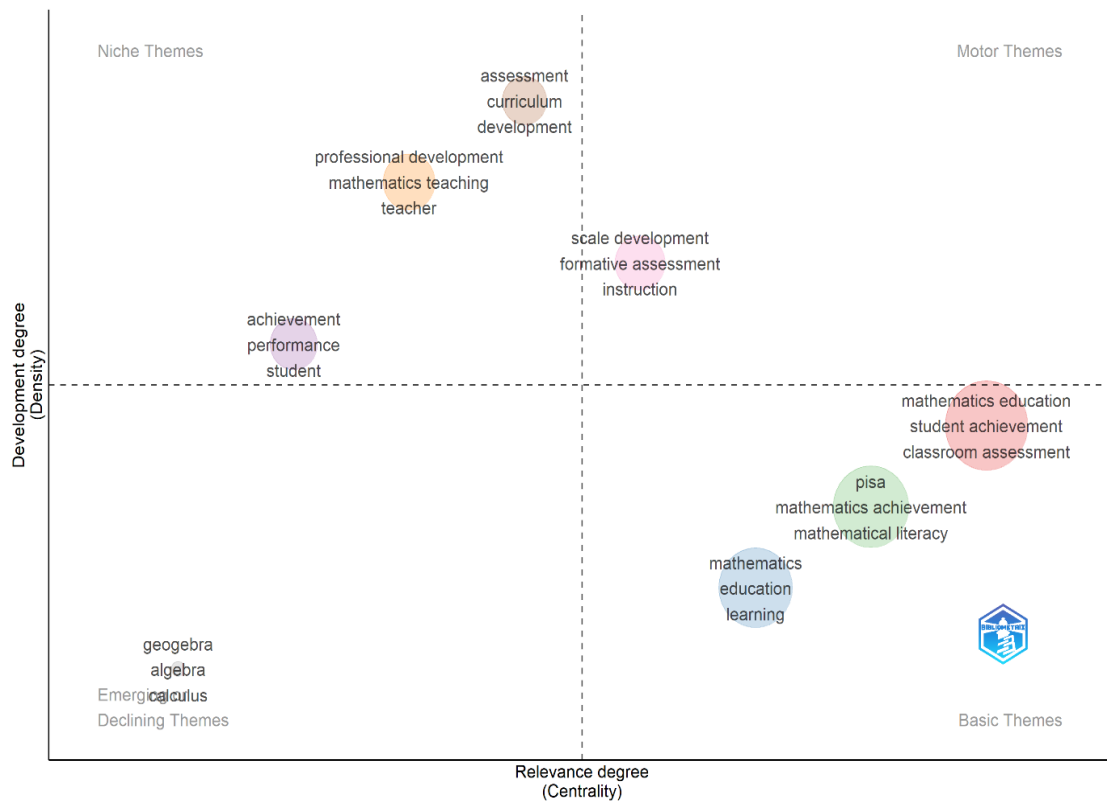


Figure 14 provides a broader perspective on trending topics, including paradigm shifts and directions from 2001 to 2025. Accordingly, we observed that the field has been influenced by international standards like PISA and TIMSS, through a more professionalisation phase focused on "achievement" and "problem-solving" since the mid-2010s; however, the evaluation processes used to be addressed mostly through structural elements such as "teaching practices" and "reform" at the beginning of the millennium.

The concept of "assessment," with the widest horizontal line in the graph, indicates a timeless and central position in the literature; the clustered blue dots in recent years point out that assessment and evaluation are no longer solely about results, but are also integrated with process-oriented digital tools such as "student engagement," "formative assessment," and "technology integration." This finding demonstrates how research on AeOME has evolved from a classical achievement-focused assessment to a modern data-mining perspective that considers technology and student performance in a multidimensional way.

Figure 15: Thematic Terminology Map of AEoME.



Our findings, the themes in Figure 15, reveal four subthemes along the centrality and density axes. Accordingly, "mathematics education," "student achievement," and "PISA" represent the "Basic Themes" with high centrality values, forming the main academic dome of the field; while topics such as "assessment," "curriculum," and "professional development" represent the "Niche Themes" with high density values, focusing on a specific framework.

The proximity of "scale development" and "formative assessment" represents the "Motor Themes" region, indicating that these areas are driving forces in the literature. However, the fundamental mathematical content, "algebra," "calculus," and "GeoGebra", within the "Emerging or Declining Themes" represent a strategic shift in research focusing from classical instruction topics towards assessment, pedagogical development, and international achievement standards.

In a specific research field, strategically diverse diagrams can be created to understand the conceptual framework, and the resulting themes can be classified into subheadings.

Accordingly, our bibliometric analysis findings are presented in a strategic diversity diagram (Figure 15) and classified into four main quadrants based on the centrality and density axes.

Figure 16: The Grouping of Keywords on AEoME.

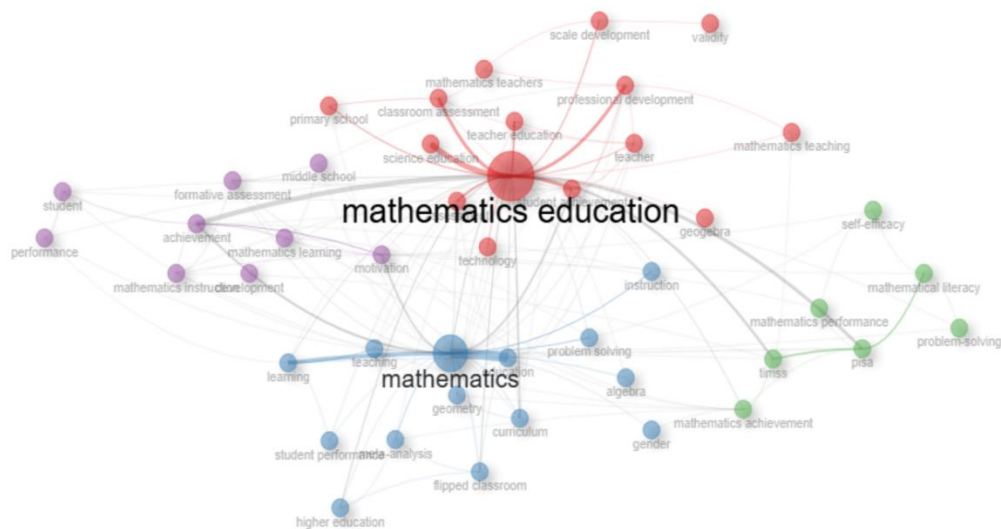


Figure 16 illustrates the conceptual network structure (co-occurrence network) of keywords in AEoME. This network analysis reveals the relational distances and thematic associations of concepts in the literature. Such an analysis provides a strategic map for understanding which contextual frameworks certain terms cluster around and which academic discussions they converge on. In the network map, node size represents word frequency, and line thickness represents the frequency with which these concepts co-occur (link strength). As a result of the analysis, it was determined that the field clusters around 4 main intellectual clusters:

Cluster 1: Teacher Education and Professional Development (Red): The cluster, concentrated at the top, includes the concepts of teacher education, professional development, and mathematics teachers. Such a finding in our study suggests that teacher competencies and continuous professional development are key to success in mathematics education.

Cluster 2: Educational Technologies and Methodological Synthesis (Blue): This cluster represents the integration of technology into education and innovative approaches such as flipped classrooms. Furthermore, the central position of the meta-analysis node in this cluster demonstrates that the effectiveness of technological interventions in the field is frequently synthesised with evidence-based methods.

Cluster 3: Psychological Factors and Academic Achievement (Purple): This student-centred cluster highlights motivation, achievement, and student performance. Our study found that mathematical achievement is not only a cognitive process but is also directly related to motivational factors.

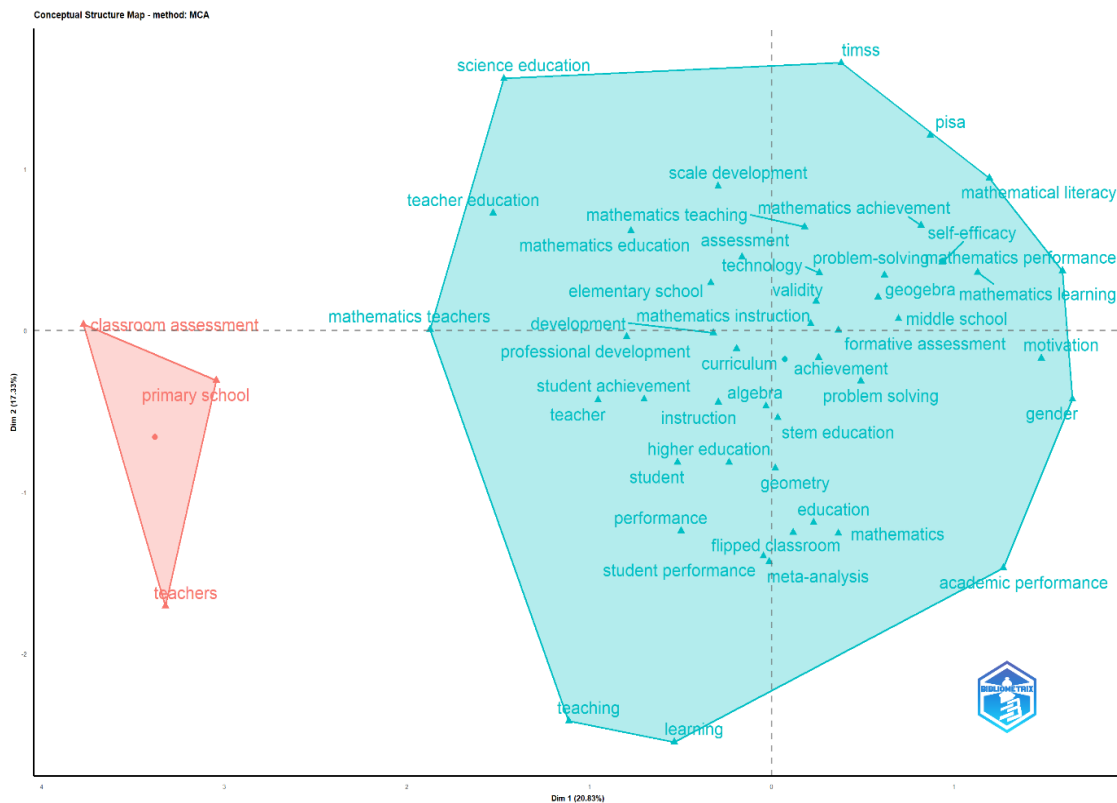
Cluster 4: International Assessments and Literacy (Green): This cluster comprises PISA, TIMSS, and mathematical literacy, which represent global standards for mathematics education. The emphasis on problem-solving skills in this cluster reflects the tendency of international exams to measure students' higher-order thinking skills, as well as their proximity to life and progressivism.

The bibliometric findings above demonstrate that AEoME research focuses on a multidimensional, integrated ecosystem encompassing teacher competence, technological innovation, affective characteristics, and global standards, rather than a single perspective.

c) The Conceptual Structure in AEoME Research

Factorial analysis is a statistical technique used to examine the structure of relationships among observed variables; in our study, we examined keywords associated with the field of AEoME in relation to the findings in Figure 17.

Figure 17: The Findings on the Authors' Keywords through Factorial Analysis



This conceptual structure map, generated using Multiple Correspondence Analysis (MCA), displays the distribution of terms in two dimensions. The Dim 1 and Dim 2 axes in the graph reveal the proximity of the concepts to one another and their thematic groupings in the literature.

1. Axes and Variance Explained Ratios

Dim 1 (X-Axis - 20.83%): This represents the distinction between "Assessment and Evaluation Approaches" and "Learning Processes" in the literature. While more

traditional/classroom assessment-focused concepts are placed on the left, performance and international standards (PISA, TIMSS) are on the right.

Dim 2 (Y-Axis - 17.33%): This reveals the difference between "Institutional/System-Oriented" (upper part: Teacher Education, TIMSS, PISA) and "Process/Application-Oriented" (lower part: Learning, Teaching, Performance).

2. Concept Groups and Clusters

A. Blue Cluster (Large Group: Mathematics Education Ecosystem)

This massive cluster forms the main body of the literature and is further divided into several sub-foci:

International Assessments: In the upper right corner, the concepts PISA, TIMSS, Mathematical Literacy, and Mathematical Achievement are grouped together. This shows that a big part of the literature focuses on international achievement tests and literacy.

Technological Applications and Teaching Methods: In the middle section, the terms such as GeoGebra, Technology, Flipped Classroom, and Problem Solving, reveals that modern teaching tools and methods attract a central importance.

Psychological Factors: On the right side, concepts such as Self-efficacy, Motivation, and Gender reveal that mathematics achievement is addressed not only in academic but also in affective and demographic dimensions.

Educational Phases: Terms such as Middle School, Elementary School, and Higher Education are placed around the centre, indicating that the studies span a wide range of educational levels.

B. Red Cluster (Small Group: Classroom Practice and Teachers)

This cluster, which is more isolated on the left, includes the concepts of "Classroom Assessment," "Primary School," and "Teachers." This separation indicates that classroom assessment and teacher-focused studies in primary schools differ methodologically or in content from the general mathematics education literature (the larger cluster focused on PISA/TIMSS)

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

Discussion

The current study provides important findings and recommendations on the AEoME field through bibliometric analyses. The most remarkable results, discussions and recommendations are as follows:

General Outlook of the AEoME Field

Our study revealed the AEoME field experienced a steady rise in publications from the 2000s onwards, with a more pronounced increase after 2010. Additionally, the average number of citations per article has been relatively high in recent years (Finding 1/a). While our bibliometric findings highlight the prominence of international exams (e.g., PISA, TIMSS), the broader literature indicates that a unified international standard for the AEoME based on a common understanding has not yet been reached on mathematics assessment, particularly on a shift from purely summative testing toward a formative approach for educators that emphasises conceptual understanding, reasoning, and problem-solving (Kubiszyn & Borich, 2024; Radmehr

& Vos, 2020). International perspectives and exams like TIMSS and PISA are seen as a means of ensuring these standards (Addey, 2017; Borch et al., 2020; Eryani et al., 2024; Lamichhane, 2018; Manizade & Orrill, 2020); Furthermore, the keywords we found in our study (through RQ3/a), such as international exams "PISA" and "TIMSS", reflect global achievement standards, which also enable comparative analyses for the field. Keywords such as "motivation", "self-efficacy", and "flipped classroom", along with the concept of "mathematics achievement", reveal a resilient, evolving, multidimensional structure encompassing student motivation, perceptions of self-efficacy, technology-supported, next-generation approaches to digital pedagogy, and academic achievement outcomes in AEoME.

Likewise, the prominent terms in our study, such as "assessment," "scale development," and "validity", indicate that studies also focus on developing reliable measurement tools and seeking methodological validity, while concepts such as "PISA," "TIMSS," and "mathematical literacy" point to research around international measurement standards and higher-level skills. However, building a shared understanding is not quite easy, as such implementation requires substantial changes and a common understanding of both assessment and the curriculum worldwide. Furthermore, AEoME faces new, diverse challenges introduced by artificial intelligence (AI) and massive digitisation worldwide (Engelbrecht & Borba, 2024; Meylani, 2025).

Scientific Productivity across Countries and Continents

Our study has revealed that the USA has been the most productive country in the world in terms of publications. This is followed by China and Turkey. Furthermore, we observed significant geographical polarisation in publication rates worldwide, with high rates in North America, Asia-Pacific, and Europe, while many countries in Africa and Central Asia are underrepresented in the literature (Finding 1/c-d). International research on AEoME is expected to contribute to the development of an international understanding of AEoME. In addition, we observed a high level of cooperation among the United States, China, and South Korea. Nonetheless, such a corporation should also be widespread, particularly in those countries across Africa and Central Asia, to lessen the significant geographical polarisation observed worldwide in our study. We suggest that international exams like TIMSS and PISA, as well as the United Nations' educational SDGs, could be used more effectively in an international perspective and understanding towards this goal.

The Common Key Concepts in the AEoME Field

Our study found that terms such as "formative assessment," "professional development," and "self-efficacy" in the AEoME field have been framed within a rather multidimensional, multidisciplinary framework that includes teacher development, student motivation, and process-oriented improvements (Finding 3/a; Figures 11-12). Such a finding is deemed to provide a clearer understanding of the conceptual relationships within specific study topics, such as assessment techniques, global exams, and affective characteristics in the literature. The positive impact of formative assessment on students' learning outcomes has often been accepted by researchers (Abell & Siegel, 2011; Black & Wiliam, 2006; Vogelzanga & Admiraal, 2017). However, it is difficult for teachers to shift their traditional perceptions and practices to this new assessment perspective as their assessment practices resist change due to various personal and contextual

factors, on the other hand, they are expected to use assessment processes to engage all students and provide more individualised assessment practices for diverse learners in the class (Bozkurt & Izci, 2020; Izci, 2018). Therefore, teacher education at universities and further training to develop the teachers' knowledge on their own professional performance and student achievement are deemed to play useful roles in shaping the AEoME field for new conditions and applications (Blömeke & Delaney, 2012; CBMC, 2001).

Our findings on the key concepts within the studies reveal that, in addition to the most commonly used concepts like "Mathematics education" and "mathematics", the publications on AEoME are directly structured around core concepts such as "Measurement, evaluation, success, achievement, performance, student achievement" (Finding 3/a; Figure 13). This result underscores the importance of measuring student outcomes and analysing academic success in the field, thereby motivating many studies. Furthermore, other key concepts, such as international standards, literacy, PISA, TIMSS, and mathematical literacy, demonstrate that international achievement criteria and the ability to apply mathematics to daily life (Mathematical literacy) have been highlighted in the literature. The key concepts around pedagogical approaches, such as "instruction, curriculum and teaching," also indicate that the teaching of mathematical knowledge, the structure of curricula, and classroom practices have attracted considerable attention in the AEoME field.

The results above (Finding 3/a; Figure 13) also align with previous research in the AEoME literature, as assessment and evaluation, student achievement and performance, particularly mathematical literacy and assessments focusing on higher-order thinking skills in mathematics learning, have been widely highlighted by several research (Craig & Guzmán, 2018; Gradini, 2022; Guskey & Brookhart, 2019; Gökçe & Güner, 2021). In addition, the importance of internationally standardised exams, math curricula, and the quality of math instruction have been the subject of several studies in the AEoME field (Demir, 2025; Hasanah et al., 2024; Johnson & Fitzmaurice, 2025).

The Current Contextual Outlook of AEoME Field around Themes, Networks and Trends

Our study has revealed that the AEoME field has been influenced by international standards like PISA and TIMSS, through a more professionalisation phase through the themes such as "achievement" and "problem-solving since the mid-2010s; however, this tendency has particularly focused on structural elements such as "teaching practices" and "reform" (Finding 3/b; Figure 14). The central themes around "assessment" has also been integrated with process-oriented digital tools such as "student engagement," "formative assessment," and "technology integration." These results suggest that research on AeOME is shifting from a classical achievement-focused assessment to a modern digital or data-mining perspective that considers technology and student performance in a multidimensional way (Hongli et al., 2024; Luo et al., 2024). However, our study argues that such a transformation should also be questioned in light of newly emerging conditions. According to several studies (Engelbrecht & Borba, 2024; Meylani, 2025), maths education is expected to adapt AI technologies to AEoME applications.

The results of our study have also revealed diverse thematic dimensions in the field, such as basic, niche, motor, and emerging or declining themes. Accordingly, topics such as

mathematics education, student achievement, and PISA still form the main academic roof of the field as "Basic Themes," while topics such as assessment, curriculum, and professional development represent the "Niche Themes" through focusing on a specific framework. On the other hand, we observed that topics such as scale development and formative assessment have been driving forces in the field as "Motor Themes," while the fundamental mathematical topics, such as algebra, calculus, and GeoGebra, turn out to be rather "emerging themes", shifting from standard instructional topics towards assessment, pedagogical development, and international achievement standards. This result also indicates that the AEoME field retains its dynamic nature through evolving trends in fundamental topics of algebra, calculus, and GeoGebra, as well as the development of scales for up-to-date measurement and formative assessment, with a focus on student achievement at both national and international levels.

Our study revealed the crucial conceptual network structure of the AEoME field within the broader academic landscape. Accordingly, the academic landscape and discussion in the field clustered within 4 subtitles: (1) Teacher Education and Professional Development; (2) Educational Technologies and Methodological Synthesis; (3) Psychological Factors and Academic Achievement; and (4) International Assessments and Literacy. This result indicates that teacher competencies and continuous professional development (CPD) are considered key to success in mathematics education, and that the AEoME field focuses on integrating technology into education and on innovative approaches, such as flipped classrooms, technological interventions, and new digital methods. Furthermore, mathematical achievement in the field has been considered not only a cognitive process but also directly related to motivational factors for students' achievement and performance. Terms such as PISA, TIMSS, and mathematical literacy exemplify global research standards and perspectives. The focus on problem-solving skills highlights how international exams tend to assess students' higher-order thinking in real-world contexts through the progressivist paradigm in AEoME.

The Conceptual Structure of AEoME Research

Our study has also found important results on the conceptual structure of the AEoME field through factorial analysis. Accordingly, we observed a distinction between "Assessment and Evaluation Approaches" and "Learning Processes" in the literature. The traditional/classroom assessments are confronting performance and international standards (PISA, TIMSS). In addition, the research field of AEoME has been dominated by two subgroups: mathematics education ecosystem versus classroom practice and teachers. The mathematics education ecosystem highlights concepts such as "international assessments, technological applications and teaching methods, psychological factors, and educational phases"; on the other hand, classroom practice and teachers' subgroup cover rather distinct concepts such as classroom assessment and teacher-focused studies, differing methodologically or in content from the mathematics education ecosystem subgroup.

These results indicate that the AEoME field is evolving from a traditional/classroom-assessments perspective to a performance- and international-standards-based perspective. In addition, the concepts focusing on classroom assessment and teachers, as more specific research fields, differ methodologically or in content from other dominant concepts in the mathematics education ecosystem.

CONCLUSION

The AEoME field showed consistent growth in publications from the 2000s onwards. Although core concepts focus on global exams, the literature indicates that a unified international standard for the AEoME has yet to be fully established, as ongoing discussions reflect a shift from purely summative to more formative approaches based on conceptual understanding, reasoning, and problem-solving. There is significant geographical polarisation in AEoME publications worldwide, with many countries in Africa and Central Asia underrepresented in the literature, while there is a high level of cooperation among the United States, China, and South Korea. Nonetheless, such cooperation should be widespread across other continents to lessen this geographical polarisation.

The conceptual networks and structure of the AEoME reveal that teacher competencies remain crucial, and the field focuses on integrating technology into education and on innovative approaches, such as flipped classrooms, technological interventions, and new digital methods. Furthermore, mathematical achievement depends not only on cognitive processes but also on motivational factors. International exams such as PISA, TIMSS, and assessments of mathematical literacy introduce a global perspective to AEoME by evaluating students' problem-solving and higher-order thinking in real-world contexts. The central themes around "assessment" in AEoME are seeking new process-oriented digital tools, such as "student engagement," "formative assessment," and "technology integration," shifting from a traditional assessment to a modern digital perspective that considers technology and student performance in a multidimensional way. However, such a transformation should also be enhanced by emerging conditions, such as AI and other digital technologies.

Recommendations

Through the results in the current study, we offer the following recommendations for policymakers, educators and researchers:

1. For Policymakers: The international standard in national math curricula should be supported in more efficient ways and with sustainable policies.
2. For Teacher Training: The new pre-service and in-service programs towards the AEoME field, focusing on formative approaches based on conceptual understanding, reasoning, and problem-solving, and also supported by digital pedagogy and AI applications, should be provided for teachers.
3. For Academic research: To lessen the geographical polarisation in AEoME publications among underrepresented countries in Africa and Central Asia, the cooperation should be widespread between developed countries and underrepresented countries across continents.

Author Contribution

Authorship contributions: 1st author 60%, 2nd author 25%, 3rd author 15%.

Conflict of Interest Statement

The authors declare no conflict of interest.

Ethics Committee Statement

Ethics committee approval is not required for this research.

REFERENCES

- Anwar, L., Sa'dijah, C., Muzam Roudloh, Y. F., Yerizon, Y., & Maiyastri, M. (2025). Enhancing students' reasoning skills with GeoGebra-based digital worksheets. *Salud, Ciencia y Tecnología*, 5, Article 2137. <https://doi.org/10.56294/saludcyt20252137>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4), 959-975. <http://dx.doi.org/10.1016/j.joi.2017.08.007>
- Bayrak, A., & Aslanci, S. (2022). Realistic mathematics education: A bibliometric analysis. *Shanlax International Journal of Education*, 10(4), 52–62. DOI: <https://doi.org/10.34293/education.v10i4.5174>
- Black, P., & Wiliam, D. (2012). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 11–32). Sage.
- Blömeke, S., & Delaney, S. (2012). Assessment of teacher knowledge across countries: A review of the state of research. *ZDM*, 44(3), 223–247. <https://doi.org/10.1007/s11858-012-0429-7>
- Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning* (Rev. and expanded ed.). Lawrence Erlbaum Associates.
- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, 110(3), 608–645. <https://doi.org/10.1111/j.1467-9620.2008.00863.x>
- Bozkurt, K., & Izci, K. (2020). A desirable but difficult goal: Equitable assessment. *Education Research Highlights in Mathematics, Science and Technology*, 160, 177.
- Callingham, R., Pegg, J., & Wright, T. (2009). Changing teachers' classroom practice through developmental assessment: Constraints, concerns and unintended impacts. *Crossing Divides*, 81–88.
- Cerado, E. C. (n.d.). Assessment and evaluation in mathematics. Academia.edu. [https://www.academia.edu/45577568/Assessment and Evaluation in Mathematics](https://www.academia.edu/45577568/Assessment_and_Evaluation_in_Mathematics). Retrieved on March 25, 2026,
- Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E. and Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci.*, 62: 1382–1402. <https://doi.org/10.1002/asi.21525>
- Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers* (Vol. 11). American Mathematical Society.

- Craig, J., & Guzmán, L. (2018). Six propositions of a social theory of numeracy: Interpreting an influential theory of literacy. *Numeracy*, 11(2), Article 2. <https://doi.org/10.5038/1936-4660.11.2.2>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Demir, M. (2025). A systematic review of literature on mathematics curricula in use at primary school level. *Cogent Education*, 12(1), Article 2557617. <https://doi.org/10.1080/2331186X.2025.2557617>
- Drijvers, P. (2018). Digital assessment of mathematics: Opportunities, issues and criteria. *Mesure et Évaluation en Éducation*, 41(1), 41–66. <https://doi.org/10.7202/1056314ar>
- Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM Mathematics Education*, 56, 281–292. <https://doi.org/10.1007/s11858-023-01530-2>
- Eryani, R., Samitra, D., Arisandi, D., Amelia, W., & Bayumi, N. (2024). Mathematics curriculum evaluation: Challenges and opportunities in meeting global education standards (Case: State junior high school). *Journal of Education Research*, 5(4), 6294–6307. DOI:10.37985/jer.v5i4.1839
- Glanzel, W. (2003). *Bibliometrics as a research field: A course on Theory and Application of Bibliometric Indicators* (Course Handout). Retrieved from: https://www.researchgate.net/publication/378543284_Modelo_para_o_desenvolvimento_de_estrategia_de_busca_para_a_recuperacao_da_informacao_cientifica_e_tecnologica_o_caso_da_impressao_3D
- Gökçe, S., & Güner, P. (2021). Forty years of mathematics education: 1980-2019. *International Journal of Education in Mathematics, Science and Technology*, 9(3), 514–539. <https://doi.org/10.46328/ijemst.1531>
- Gradini, E. (2022). Development of rubric of higher order thinking skills assessment on mathematics learning. *Jurnal Teori dan Aplikasi Matematika*, 6(1), 61–74. <https://doi.org/10.31764/jtam.v6i1.5342>
- Guskey, T. R., & Brookhart, S. M. (2019). *What we know about grading: What works, what doesn't, and what's next*. ASCD.
- Hackenberger, B. K. (2020). “R Software: Unfriendly but Probably the Best”, *Croatian Medical Journal*, 61, 66–68. DOI:10.3325/cmj. 2020.61.66.
- Hasanah, N. R., Turmudi, T., Juandi, D., & Amiruddin, M. Z. B. (2024). Exploring research trends on PISA in mathematics: A systematic review using bibliometric analysis. *Avalokana Journal of Literature Review*, 1(1), 14–29.
- Hasumi, T., & Chiu, M. S. (2022). Online mathematics education as bio-eco-techno process: Bibliometric analysis using co-authorship and bibliographic coupling. *Scientometrics*, 127(8), 4631–4654. <https://doi.org/10.1007/s11192-022-04374-2>
-

- Hongli, T., Liu, J., Xu, J., & Zhang, M. (2024). Educational data mining for student performance prediction: Feature selection and model evaluation. *J. Electr. Syst.*, 1063–1074. DOI:[10.52783/jes.3434](https://doi.org/10.52783/jes.3434)
- Howard, A., & Donaghue, H. (2014). *Teacher evaluation in second language education*. Bloomsbury Academic.
- İnceoğlu, Ç. (2014). Türkiye’de sinemayı konu alan doktora tezleri üzerine bibliyometrik bir çözümleme. *Galatasaray Üniversitesi İletişim Dergisi*, (21), 31-50. doi: <https://doi.org/10.16878/gsuilet.96674>
- Johnson, P., & Fitzmaurice, O. (2025). Navigating mathematics curriculum reform: Teachers’ efforts and implementation challenges. *Journal of Curriculum Studies*, 1–25. <https://doi.org/10.1080/00220272.2025.2495973>
- Kellaghan, T., & Stufflebeam, D. L. (Eds.). (2003). *International handbook of educational evaluation*. Kluwer Academic Publishers.
- Kubiszyn, T., & Borich, G. D. (2024). *Educational testing and measurement*. John Wiley & Sons.
- Lamichhane, B. R. (2018). Assessment practices in mathematics: Local to global contexts. *Saptagandaki Journal*, 9, 1–16. <https://doi.org/10.3126/sj.v9i0.20876>
- Luo, Z., Mai, J., Feng, C., Kong, D., Liu, J., Ding, Y., ... Zhu, Z. (2024). A method for prediction and analysis of student performance that combines multi-dimensional features of time and space. *Mathematics*, 12(22), Article 3597. <https://doi.org/10.3390/math12223597>
- Manizade, A., & Orrill, C. H. (2020). International perspectives on frameworks for mathematics teachers’ knowing and quality of mathematics instruction. *Research in Mathematics Education*, 22(2), 93–97. <https://doi.org/10.1080/14794802.2020.1798809>
- McAlpine, M. (2002). *Principles of assessment*. University of Glasgow, Robert Clark Center for Technological Education.
- Meylani, R. (2025). AI-powered assessments in mathematics education: A systematic review of contemporary research literature. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, 66, 3642–3674. <https://doi.org/10.53444/deubefd.1527781>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Bmj*, 339: b2535 doi: 10.1136/bmj.b2535
- Moses, R. P., & Cobb, C. E. (2001). *Radical equations: Civil rights from Mississippi to the Algebra Project*. Boston, MA: Beacon Press.
- Nortvedt, G. A., & Buchholtz, N. (2018). Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity. *ZDM*, 50(4), 555–570. <https://doi.org/10.1007/s11858-018-0950-6>
- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment.

Eurasia journal of mathematics, science and technology education, 19(8), em2307.
<https://doi.org/10.29333/ejmste/13428>

- Radmehr, F., & Vos, P. (2020). Issues and challenges of 21st century assessment in mathematics education. In L. Leite, E. Oldham, A. S. Afonso, F. Viseu, L. Dourado, & H. Martinho (Eds.), *Science and mathematics education for 21st century citizens: Challenges and ways forward* (pp. 437–462). Nova Science Publishers.
- Ruthven, K. (1994). Better judgement: Rethinking assessment in mathematics education. *Educational Studies in Mathematics*, 27, 433–450. <https://doi.org/10.1007/BF01273382>
- Tjøstheim, D., Jullum, M., & Løland, A. (2023). Statistical embedding: Beyond principal components. *Statistical Science*, 38(3), 411–439. <https://doi.org/10.1214/22-STS881>
- Toker, Z., & Aksoy, E. (2024). A bibliometric review of studies on mathematics teacher professional development with an emphasis on mathematics coaching research. *Journal of Mathematics Teacher Education*, 1–75. DOI:[10.1007/s10857-024-09671-1](https://doi.org/10.1007/s10857-024-09671-1)
- Weir, C. J. (2005). Limitations of the Common European Framework for developing comparable examinations and tests. *Language testing*, 22(3), 281–300. DOI:[10.1191/0265532205lt309oa](https://doi.org/10.1191/0265532205lt309oa)
- Yambi, T. D. A. C., & Yambi, C. (2018). *Assessment and evaluation in education*. Universidade Federal do Rio de Janeiro.
- Zupic, I., & Cater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>
-