International Journal of Progressive Education, Volume 20 Number 6, 2024

Emerging Research Themes in Mathematics Education: A Topic Modeling Analysis of Most Influential Journals (2019-2023)

Kenan Gökdağ[i]
Gaziantep University

Mehmet Fatih Özmantar[ii]
Gaziantep University

Abstract

As in various scientific fields, the volume of publications in mathematics education is rapidly increasing, complicating the detailed examination of academic outputs. Latent Dirichlet Allocation (LDA)-based topic modeling algorithms have gained popularity for their ability to synthesize vast amounts of data and provide an overview of emerging research themes within specific fields. This article conducts a topic modeling analysis on 1,000 scholarly articles published between 2019 and 2023 in the five most influential journals in mathematics education. The study identifies 18 research themes, comparing these with a comprehensive topic modeling study conducted in 2018 (Inglis & Foster, 2018). Newly emerging themes include Mathematical Performance and Assessment, Lesson Study, Mathematical Modeling, Social Justice, Teacher Practice, Statistical Literacy, Prospective Teachers' Noticing of Student Thinking, and Framework Design and Development. The findings indicated teacher professional development and education-related studies have been the most prolific research areas over the past five years. Additionally, the research themes and keywords highlighted the ongoing social transformation and shifts in research focuses within mathematics education. This study is expected to be a resource for researchers who conduct research and determine the research theme.

Keywords: Topic Modeling, Mathematics Education, Research themes, Scientific Publications

DOI: 10.29329/ijpe.2024.1078.2

Submitted: 01/07/2024 **Accepted:** 30/10/2024 **Published:** 01/12/2024

---

[i] **Kenan Gökdağ,** Mathematics and Science Education, Gaziantep University, ORCID: 0000-0001-8087-563X

Correspondence: kenangokdag0133@gmail.com

[ii] **Mehmet Fatih Özmantar,** Prof. Dr., Mathematics Education, Faculty of Educatıon, Gaziantep University

INTRODUCTION

Research in mathematics education stands as a crucial area within the educational research landscape. Numerous studies in this field have been conducted, yielding results that are extensively discussed within academic circles. Mathematics education encompasses a variety of sub-topics, including different learning processes and the challenges students face during these processes. As knowledge within the field increases daily, traditional methods of analysis are becoming increasingly inadequate for researchers, posing significant challenges. These challenges prevent a comprehensive understanding of the literature and a broad perspective of the field. The dramatic increase in available data compels researchers to seek alternative methods of investigation. Furthermore, determining the trends and preferred topics of research within the field has become more complex (Inglis & Foster, 2018).

Researchers utilize various methods such as bibliometric analysis and systematic reviews to identify these topics and trends (Nie & Sun, 2017). Systematic reviews offer a general perspective on selected studies in terms of subject context and proven techniques. However, maintaining systematic reviews in a dynamically evolving field can be challenging (Chen, 2016). According to Chen (2016), the primary reason for this challenge is the inability of systematic reviews to provide a comprehensive and holistic view in rapidly developing areas. On the other hand, bibliometric analysis creates a network model based on data such as citations, authors, or keywords from studies conducted within a designated research area. Network modeling and the visualization of these models can reveal conceptual maps of the research field (Chen, 2016). Despite these advantages, bibliometric analysis also has its drawbacks. It primarily generates conceptual maps based on keywords alone, not incorporating the main texts or abstracts of the studies (Chen & Luo, 2019). Choi and Kwak (2019) recommend analyzing the abstract sections of studies in this field, as they typically summarize all contents and results, highlighting their significance. These abstracts contain valuable information about research trends in mathematics education. However, working with the vast textual data available in this long-established field presents various challenges in terms of time and effort, hence limiting its use.

In their study, Griffiths and Steyvers (2004) eliminated this limitation of text mining, which is an extension of data mining and defined as a method of extracting meaningful patterns from a collection of unstructured texts, by using the topic modelling method offered by text analysis on the abstracts of journals (Olson & Delen, 2008). Text data mining progresses through five stages: data gathering, data cleaning, data processing, data mining, and evaluating results (Zengul et al., 2021). Significant text mining applications include clustering, information extraction, and topic mapping. These techniques provide tools that save time and energy in processing and analyzing information (Delen and Olson, 2008).

One of the applications of text mining is topic modeling, which analyzes texts to identify themes. LDA has become increasingly popular in recent years as a method for topic modeling (Blei, 2012). LDA is a probabilistic generative model that aims to identify the main themes within a collection of textual documents (Blei et al., 2003). This model significantly reduces the dimensions (words) in a text while preserving the crucial connections among all dimensions and the main topics in subsequent documents (Blei and Lafferty, 2007). LDA assumes that a document contains multiple topics, with each word in the document presumed to stem from a particular topic (Hung, 2012). Modeling is based on estimating the probability that a particular word belongs to a specific topic. Algorithms use the most suitable probability to allocate words to a certain topic. Topic modeling not only uncovers the discussed topics within a corpus of studies but also reveals overlooked research areas. This method provides recommendations for future research, serving as a predictive tool for emerging trends within a field. Topic modeling employs unsupervised machine learning techniques using various statistical models to identify themes from large sets of textual data, making it a popular method in recent literature reviews (Asmussen & Møller, 2019; Griffiths & Steyvers, 2004; Paul & Girju, 2009). Researchers like Griffiths and Steyvers (2004) and Blei and Lafferty (2007) have applied topic modeling to understand the themes and trends emerging from academic journal literatures.

Similarly, Bittermann and Fischer (2018) employed topic modeling to identify current trends in psychology.

In the field of mathematics education, topic modeling has been effectively applied to analyze research trends. Inglis and Foster (2018) utilized this approach on articles from highly influential journals such as Educational Studies in Mathematics and the Journal for Research in Mathematics Education. Their analysis revealed 28 distinct topics, which they compared to understand evolving trends within the field. Similarly, Chen and Luo (2019) applied topic modeling to 3,963 articles from the Computers and Education Journal, identifying 24 topics. Herfort, Tamborg, Meier, Allsopp and Misfeldt (2023) conducted topic modeling on 336 studies presented at the European Society for Research in Mathematics Education (CERME) conference, focusing on technology in mathematics education. Choi and Kwak (2019) analyzed 2,556 publications from five leading international journals in mathematics education, revealing current trends in five-year intervals. The studies utilized Latent Dirichlet Allocation (LDA), a method proposed by Blei, Jordan and Ng (2003), which employs a Bayesian framework for text mining to uncover thematic structures within a field. LDA considers the textual structure and words of studies as data within a system that models topics integral to the field. Additionally, LDA not only provides an overview of the topic structure but also details the distribution of topics across individual documents. For such analyses, Knime is frequently used as a data visualisation and analysis program (KNIME, 2020). These implementations of topic modeling offer significant insights into the themes and directions evolving within mathematics education research.

Despite the presence of studies using topic modeling in mathematics education, there has been a notable absence of such applications over the last five years specifically aimed at reviewing the emerging topics and trends through research. Therefore, this study applies LDA topic modeling to abstracts from influential journals in mathematics education to identify the themes and trends from publications over the past five years (2019-2023). This approach is crucial for understanding how the field has evolved and for drawing implications for future research. By analyzing the scientific research published in influential academic journals within the field using LDA topic modeling, this study evaluates the themes derived to comprehend the development of the field and to provide insights that will guide subsequent inquiries.

This study employs LDA-based topic modeling analysis to identify the research themes focused on in the top five influential academic journals in mathematics education from 2019-2023. In this context, answers to the following two research questions will be searched:

1. What are the emerging themes in mathematics education over the last five years?

2. How have these themes evolved or differed from the existing research themes identified in previous studies?

Answering these questions will provide a detailed understanding of the dynamic shifts and advancements in the field since the last major review by Inglis and Foster in 2018. Specifically, the contributions of this research to the field can be articulated as follows. Firstly, by pinpointing the current themes, this study will outline the most actively researched areas within mathematics education over the past five years. This is crucial for researchers, curriculum developers, and policy makers to understand where the field is currently focusing and what gaps might exist. Secondly, comparing the newly identified themes with those from previous studies, particularly the comprehensive review by Inglis and Foster (2018), will highlight shifts and continuities in research focus. This comparison is vital for tracking the evolution of the field, understanding how certain areas have expanded or receded in focus, and identifying new research frontiers. Finally, the analysis will provide a clear depiction of how the field is evolving, offering insights into potential areas that require further investigation or could benefit from renewed focus. This foresight can help steer future research efforts, funding priorities, and educational policy decisions.

METHOD

The research method of the study consists of two main parts. These are data collection and LDA topic modelling. We aim to determine the research themes and trends of the field of study through the dataset obtained with LDA topic modelling.

Creation of Dataset

For the dataset creation, this study draws upon the work of Williams and Leatham (2017), who presented twenty journals that significantly influence the field of mathematics education and assessed the quality of scientific publications. Their study incorporated two types of analysis: citation-based and opinion-based. Given that the Williams and Leatham (2017) study is the most recent comprehensive assessment of journal quality in mathematics education, it has been used to guide journal selection for this research. Additionally, data were collected from the Web of Science (WoS) Social Science Citation Index (SSCI), a globally recognized database containing over 21,000 peer-reviewed high-quality scientific journals ("Web of Science Core Collection - Web of Science Group", 2020). Consequently, the citation-based evaluation from the Williams and Leatham (2017) study was utilized to rank the academic journals in mathematics education, the results of which are shared in Table 1.

This approach ensures a rigorous and systematic creation of a relevant dataset to provide a comprehensive overview of the current research landscape in mathematics education. By leveraging a citation-based selection of journals, the study aligns with established scholarly metrics, enhancing the validity of the research findings. Moreover, the use of a reliable database like WoS guarantees the inclusion of high-quality and impactful studies, providing a robust dataset for subsequent LDA topic modeling.

Table 1 Citation Ranking of 10 Journals That Publish Research in Mathematics Education (Williams & Leatham (2017)

Ranking	Journal	No. of citations (including self-citations)	No. of citations (excluding self-citations)
1	Educational Studies in Mathematics (ESM)	2,729	1,872
2	Journal for Research in Mathematics Education (JRME)	2,188	1,854
3	Journal of Mathematical Behavior (JMB)	848	554
4	For the Learning of Mathematics (FLM)	625	507
5	Mathematical Thinking and Learning (MTL)	490	429
6	Journal of Mathematics Teacher Education (JMTE)	630	427
7	Zentralblatt für Didaktik der Mathematik (ZDM)	740	376
8	Mathematics Education Research Journal	263	175
9	International Journal of Math Education in Science and Technology (IJMEST)	526	166
10	School Science and Mathematics (SSM)	307	122

As seen in Table 1, the journals were ranked as follows based on their influence in the field of mathematics education: Educational Studies in Mathematics (ESM) at first position, followed by the Journal for Research in Mathematics Education (JRME), Journal of Mathematical Behavior (JMB), For the Learning of Mathematics (FLM), and Mathematical Thinking and Learning (MTL). However, as FLM is not indexed in the Web of Science Citation Index (www.webofknowledge.com), it was excluded from the analysis, and the Journal of Mathematics Teacher Education (JMTE) was included instead. Consequently, the journals ESM, JRME, JMB, MTL, and JMTE were selected for this study, and their published research articles were incorporated into our analysis. From 2019 to 2023, a total of 1,000 articles published in these journals were collected. Only research articles were included in the analysis, excluding any other types of research publications. The dataset used in this study comprised abstracts extracted from these collected articles.

Latent Dirichlet Allocation (LDA) Topic Detection

In this research, LDA topic detection was employed to uncover themes and determine trends. For this analysis and model creation, KNIME data mining software was used (KNIME, 2020). KNIME facilitates the analysis process with its easy-to-use visual interface that offers pre-built structures instead of complex base codes. The code process for LDA topic detection was developed by researchers using the KNIME data mining software, which includes ready-to-use code blocks for essential analysis structures like text preprocessing, LDA, and lemmatization. Figure 1 represents our text mining process.

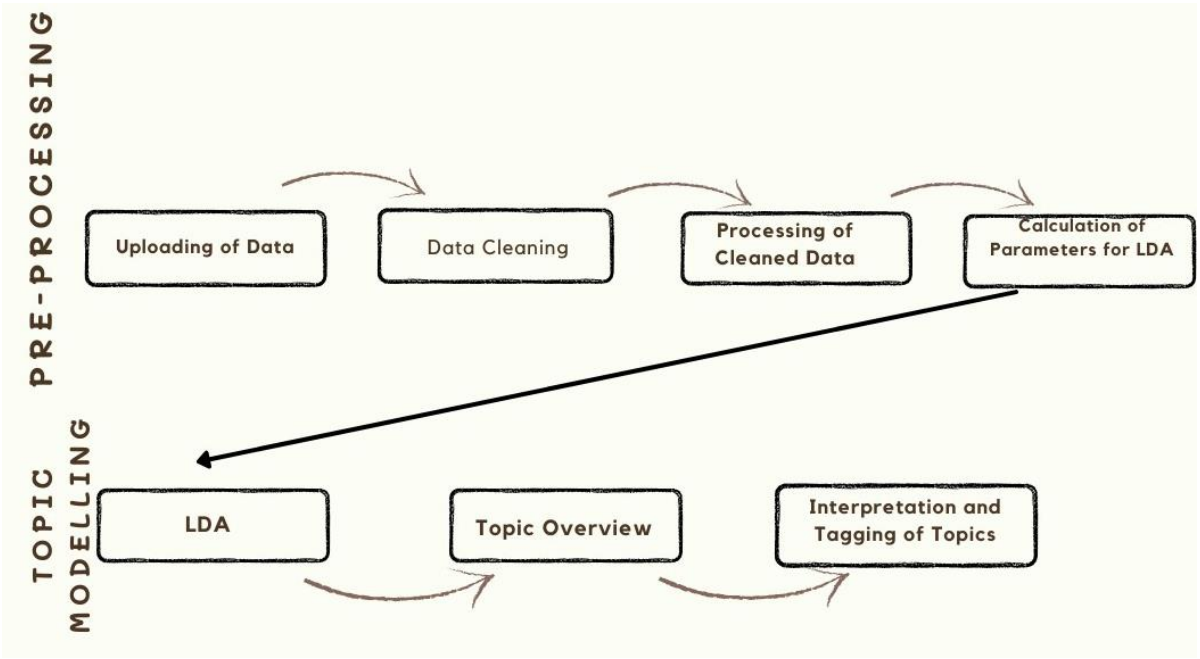

Figure 1. LDA topic detection process

In our LDA topic modeling process using the KNIME data mining software, the initial step involves loading the dataset into the system as an Excel file. The next step is to apply text preprocessing techniques to the abstracts within the dataset. These techniques include tagging, filtering, and lemmatization, transforming words into their most meaningful forms based on context. The objective of this step is to remove various punctuation marks, word suffixes, repetitions, and common terms such as "the" that may affect the analysis.

After cleaning the data, the processed words in their purest and cleanest forms are transferred to the LDA topic extraction node. It is essential that the parameters for the LDA topic extractor node are set in advance. Based on recommendations from prior studies, the parameters were set with alpha (α) at 0.1 and beta (β) at 0.01 (Gurcan et al., 2022). Another critical parameter for the LDA topic extraction node is the number of topics to analyze. Various methods, such as the elbow method and silhouette method, are available to determine the best number of topics. In our study, we utilized perplexity, as used by Inglis and Foster (2018), to ascertain the most suitable number of topics.

Once the optimal number of topics was determined using perplexity, the model generated was reviewed by experts who examined the most representative studies and the words with the highest representation weight to assign the most appropriate names to the topics. This process ensures that the topics identified are both significant and relevant to the current trends and themes in mathematics education research.

FINDINGS

The findings from our LDA-based topic modeling analysis are presented under three main headings. Firstly, the results of the perplexity analysis are presented, which focuses on the number of research themes around which the studies published in the top five mathematics journals indexed by WoS from 2019 to 2023 are centered. Secondly, we will present the themes identified through the LDA-based topic modeling analysis in response to our first research question. Lastly, in alignment with our second research question, we discussed how the identified research themes have differed from those identified by Inglis and Foster (2018) in the last five years.

Perplexity Plot: Number of Topics

Determining the appropriate topic-word distribution in LDA topic modeling analysis is of significant importance. This measure helps to ascertain the suitable number of research themes from the 1000 scientific publications collected from the five most influential academic journals in the field of mathematics education. The results of the perplexity analysis used to determine the number of topics are shared in Figure 2.

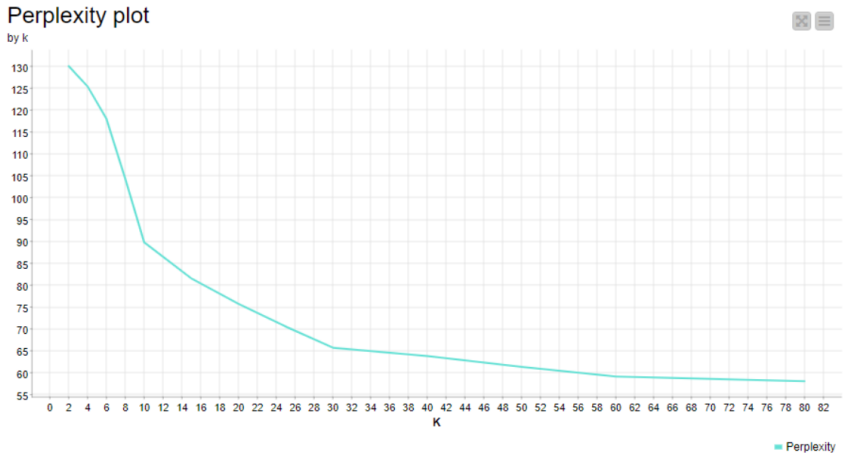

Figure 2. Perplexity K topics

As presented in Figure 2, the number of topics for the 1000 studies reviewed ranges from 2 to 80. According to Maier et al. (2018), the most suitable number of topics is determined by examining the points where the perplexity K topics graph shows sudden and significant breaks. This approach, referred to as the elbow method, was proposed by Dewi and Thiel (2017). Considering these methodologies, a clear break in the graph at 10 and 30 topics can be observed in Figure 2. Upon detailed examination of these values and intervals, our research team has identified a maximum of 18 distinct topics that meaningfully correlate with the dataset. Therefore, the optimal number of topics has been determined to be 18.

Research Themes Extracted Through Topic Modeling

The main findings of our study are based on the 18 topics derived from the perplexity analysis. Using the KNIME software, the analysis results in a map created from the studies clustered under each topic, using keywords and abstracts. The naming of these topics has been conducted in light of the relevant literature, considering the data from keywords and abstracts. During this naming process, the weights of the keywords presented under each topic were also taken into account. Moreover, data files generated through the KNIME program were analyzed to examine the weights of the studies presented under each topic.

This part of the study details how the topics were named, with the keywords used in the topic naming and the percentage of studies under each topic within the total body of work presented in Table 2. This section will explain how the topics were named based on these data.

Table 2 Topic name, topic keywords and proportion of the clusters within the dataset

Topic	Topic name	Top 10 most relevant terms(sorted by relevance)	% of tokens
Topic-0	Mathematical performance and assesment	student, mathematics, assessment, study, belief, self-efficacy, performance, test, item, attitude	6.4
Topic-1	Function concept	function, student, concept, calculus, study, activity, theory, participant, result, understand	3.1
Topic-2	Proof and argumentation	proof, student, argumentation, example, argument, mathematical, mathematician, definition, collective, conjecture	6.1
Topic-3	Algebraic thinking	student, reasoning, meaning, task, relationship, algebraic, equation, function, quantity, mathematical	7.5
Topic-4	Digital environment	student, environment, datum, generalization, pattern, geometry, dynamic, representation, statistical, process	5.1
Topic-5	Addition and subtraction	child, word, study, strategy, skill, task, intervention, ability, knowledge, numerical	4.3
Topic-6	Lesson study	study, lesson, mathematics, factor, problem-solving, adaptation, activity, analysis, sfon, japanese	2.4
Topic-7	Prospective teachers noticing of student thinking	prospective-teacher, pst, course, student, feedback, prospective, study, pt, student-thinking, productive	4.7
Topic-8	Language	language, student, concept, mathematics, classroom, programming, space, mathematical, learner, multilingual	5.0
Topic-9	Mathematical discourse	discourse, student, mathematical, classroom, discussion, lesson, interaction, framework, learning, authority	7.5
Topic-10	Mathematical modelling	task, mathematical, strategy, student, model, modeling, solution, process, study, modelling	6.6
Topic-11	Curriculum	textbook, curriculum, context, approach, question, difference, school, explanation, educational, material	3.0
Topic-12	Fraction	fraction, unit, reasoning, multiplicative, student, proportional, multiplication, concept, equivalence, understanding	4.5
Topic-13	Socail justice	mathematics, Mathematics-education, social, research, student, community, school, black, experience, narrative	5.8
Topic-14	Teacher Practice	teacher, practice, teaching, mathematics, instructional, instruction, pedagogical, program, knowledge, identity	8.0
Topic-15	Statistical literacy	Datum, mathematical, mathematics, reading, mathematics-education, critical, literacy, pandemic, statistical, Covid-19	3.1
Topic-16	Framework desing and development	research, knowledge, desing, learning, activity, framework, approach, paper, perspective, development	7.3
Topic-17	Professional development	teacher, practice, Mathematics-teacher, teaching, classroom, lesson, experience, study, context, professional-development	8.3

Topic 0: Mathematical Performance and Assessment

The first research theme identified through the LDA topic modeling analysis conducted with the KNIME software is named "Mathematical Performance and Assessment." This theme encompasses 64 scientific studies, accounting for 6.4% of the total dataset. The keywords "mathematics," "assessment," and "performance" were directly employed in naming this topic. Upon reviewing the publications clustered under this theme, it was found that they primarily discuss the assessment of students' mathematical knowledge and skills, how such assessments impact students' performance, and how mathematics teaching can enhance this performance. These findings align with research like that of Dowker et al. (2019), which examines similar aspects.

Topic 1: Function Concept

The theme "Function Concept" comprises 31 studies, representing 3.1% of all the analyzed works. The naming of this topic directly utilized the keywords "function" and "concept." Studies within this theme focus on the learning and understanding of basic ideas about functions, as illustrated

by research such as that conducted by Martínez-Planell and Trigueros (2019), which explores conceptual understanding in this area.

Topic 2: Proof and Argumentation

For Topic 2, the keywords "proof" and "argumentation" were directly used in the theme's naming. This theme includes 61 studies, corresponding to 6.1% of the total dataset. The publications under this theme are primarily concerned with the evaluation of arguments produced in mathematics education and teaching and discussions based on proofs. Research by Ko and Rose (2021) exemplifies this focus, exploring how students and teachers engage with mathematical argumentation and proofs. The use of these keywords provides a general descriptor that resonates with the existing literature on the topic.

Topic 3: Algebraic Thinking

Under Topic 3, the analysis revealed that 75 scientific studies are clustered, constituting an examination of students' transformations in algebraic thinking skills through various algebraic concepts and applications. The keyword "algebraic" emerged with the highest weight and was directly incorporated into the theme's naming. The reviewed literature under this theme predominantly discusses how students engage with and understand algebraic concepts and their applications, as seen in the work of Ellis et al. (2020). Additionally, studies like those by Pittalis et al. (2020) emphasize discussions around students' thought processes concerning algebraic applications, leading to the inclusion of "thinking" in the theme's title to align with the existing literature, thus coining the term "Algebraic Thinking."

Topic 4: Digital Environment

Topic 4 encompasses 51 scientific studies (5.1% of the dataset), focusing on how students and teachers utilize digital tools and the impacts of these tools on teaching. The keyword "environment" had the highest weight and was directly used in the theme's naming. Research under this theme, like that by Yao and Manouchehri (2019), examines the use of digital tools in educational settings and their effects on teaching methodologies. Further literature review reveals a consistent focus on digital tools in teaching, prompting the use of "digital" to refine the theme's title to "Digital Environment," reflecting the integration of digital tools in educational contexts.

Topic 5: Addition and Subtraction

For Topic 5, 43 studies (4.3% of the dataset) are clustered under the theme "Addition and Subtraction." This theme was named based on the prevalent focus in the clustered studies on the effects of teaching experiments and practices concerning addition and subtraction, as demonstrated in the research by Kullberg, Björklund, Brkovic and Runesson Kempe (2020). The analysis of the existing literature indicates that the concepts used in teaching addition and subtraction practices align well with the keywords identified under this topic, confirming the appropriateness of the theme's naming.

Topic 6: Lesson Study

For Topic 6, clustered around 24 scientific studies (2.4% of the dataset), the theme has been named "Lesson Study." The keywords "lesson" and "study" were directly utilized in the theme's naming. This theme covers research on how teachers and pre-service teachers collaboratively develop, implement, observe, and evaluate lesson plans, enhancing their professional knowledge and teaching skills. Studies like those by Lendínez Muñoz, García García, Lerma Fernández and Abril Gallego (2019) focus on these aspects, indicating the practical and collaborative nature of lesson studies in professional development.

Topic 7: Prospective Teachers Noticing of Student Thinking

Under Topic 7, which includes 47 studies (4.7% of the dataset), the theme "Prospective Teachers Noticing of Student Thinking" emerged. This theme was named highlighting the key phrase "prospective teacher," which carries significant weight among the keywords. The clustered studies discuss how pre-service teachers' noticing skills are applied and developed through various instructional outcomes. Research by Warshauer et al. (2023) exemplifies this focus, showing an emphasis on how these future educators perceive and respond to student thinking, aligning closely with the concept of "noticing" from the literature.

Topic 8: Language

For Topic 8, the keyword "Language," holding the highest weight, guided the naming of the theme. The 50 clustered studies explore how language is handled within the context of mathematics classrooms, particularly focusing on linguistically and racially marginalized students. The research, as illustrated by Barwell (2020), debates how linguistic frameworks affect students' learning outcomes and the concretization of mathematical ideas, emphasizing the critical role language plays in educational equity and understanding.

Topic 9: Mathematical Discourse

In Topic 9, which clusters 75 scientific articles, the theme "Mathematical Discourse" has been identified. The keywords "Mathematical" and "discourse" were directly employed to name the theme. The research within this theme examines the effects of classroom practices, such as discussions and interactions (notated by the same keywords) on the instructional activities used in mathematics teaching. Studies by Weingarden, Heyd-Metzuyanim and Nachlieli (2019) highlight how classroom discourses influence the educational experiences and learning of students, emphasizing the importance of communicative practices in enhancing mathematical understanding.

Topic 10: Mathematical Modelling

Topic 10, named "Mathematical Modelling," directly employs the keywords "mathematical" and "modelling." This theme, clustering 66 studies, focuses on discussions around concepts and applications related to mathematical modelling. Significant research, such as the study by Krawitz et al. (2022), explores how mathematical modelling is utilized to solve real-world problems and enhance mathematical understanding.

Topic 11: Curriculum

In Topic 11, identified as "Curriculum," 30 academic publications (3.0% of the dataset) are clustered. The keyword "curriculum" was used directly in the theme's naming. Research under this theme often discusses the impacts of educational reforms on mathematics curricula across various countries, with studies such as those by Drijvers, Kodde-Buitenhuis and Doorman (2019) examining the evolution of curricular standards and their implementation challenges.

Topic 12: Fraction

Topic 12 has been named "Fraction," centering around 45 studies (4.5% of the dataset), with "fraction" being the highest weighted keyword used directly in the theme's naming. This theme explores the conceptual structures students form around fractions, highlighted in research by González-Forte et al. (2022), which delves into students' understanding and misconceptions regarding fractions.

Topic 13: Social Justice

Topic 13, termed "Social Justice," includes 58 scientific studies. The theme employs the keyword "social" directly in its naming. The clustered publications examine issues of social injustice within mathematics classrooms and their implications for teaching. Research, such as that by Chen (2023), discusses how educators address social inequalities that emerge in the classroom environment, often focusing on dimensions such as race, gender, and cultural backgrounds. The inclusion of the term "justice" complements this focus, aligning with scholarly discussions on equity and inclusion in mathematics education.

Topic 14: Teacher Practice

Topic 14, "Teacher Practice," comprises the largest cluster of studies in the dataset, with 80 scientific studies accounting for 8.0% of all analyzed works. The theme was directly named using the two highest weighted keywords: "teacher" and "practice." Research within this theme primarily explores the changes and effects in mathematics teachers' methods and knowledge due to various practices and techniques. For instance, studies like those by Woods and Weber (2020) discuss how different teaching approaches impact teacher practice and student outcomes, highlighting the dynamic nature of educational methodologies.

Topic 15: Statistical Literacy

The theme for Topic 15 is "Statistical Literacy," which clusters 31 studies (3.1% of the dataset). This theme has gained particular relevance due to the interpretation and discussion of statistical data that became prominent during the COVID-19 pandemic, as noted in research by Borba (2021). The focus on statistical literacy reflects the growing importance of data interpretation skills in education, particularly in understanding real-world events through a mathematical lens.

Topic 16: Framework Design and Development

In Topic 16, 73 studies (7.3% of the dataset) are grouped under the theme "Framework Design and Development." The keywords "design" and "development" were employed directly in naming the theme, which examines the theoretical and practical approaches used in mathematics education. Research, such as that by Simon (2022), focuses on designing and developing frameworks that enhance the effectiveness of mathematical instruction and learning strategies, addressing the needs of both students and educational goals.

Topic 17: Professional Development

Topic 17, "Professional Development," contains the highest number of clustered studies, with 83 research articles. The theme name directly incorporates the phrase "Professional Development," reflecting the content of the clustered publications. Studies within this theme, such as those by Matranga and Silverman (2022), discuss the design and critique of various professional development programs for teachers, emphasizing their crucial role in teacher growth and improvement in teaching practices.

Differentiation in Research Themes

In this section, we compare the research themes identified in this study with those established by Inglis and Foster (2018) to highlight both newly emerged themes and those that have continued relevance in the field of mathematics education. Table 3 presents a comparative analysis of the topics identified by Inglis and Foster (2018) against the themes extracted from this study. In the table, themes that are similar are indicated by their names, while differing themes are marked with an asterisk (*). This comparison is based on both the theme names and the key terms found within each theme. Despite some themes having different names, they may aggregate under the same key terms,

indicating conceptual similarities. For example, as seen in Table 3, the "Fraction" theme identified in our model shows similarities with the "Rational Numbers" theme from Inglis and Foster (2018).

Table 3 Cross thematic comparisons

Research topics identified in this study	Research topics identified by Inglis and Foster (2018)
Mathematical performance and assesment	*
Function concept	Analysis
Mathematical modelling	*
Curriculum	Curriculum (especially Reform)
Fraction	Rational Numbers
Social Justice	*
Teacher Practice	*
Statistical literacy	*
Framework desing and development	*
Professional development	Teacher knowledge and beliefs
Proof and argumentation	Proof and argumentation
Algebraic thinking	School algebra
Digital environment	Dynamic geometry and visualization
Addition and subtraction	Addition and subtraction
Lesson study	*
Prospective teachers noticing of student thinking	*
Language	Multilingual learners
Mathematical discourse	Observations of classroom discussion

This comparative analysis enables the identification of new topics that have emerged in the last five years (2019-2023), illustrating how shifts within the field have influenced the evolution of research themes. Based on the information provided in Table 3, eight themes distinctly different from those in the Inglis and Foster (2018) model have been identified. These are:

- Mathematical Performance and Assessment
- Lesson Study
- Mathematical Modelling
- Social Justice
- Teacher Practice
- Statistical Literacy
- Prospective Teachers Noticing of Student Thinking
- Framework Design and Development

DISCUSSION AND CONCLUSION

In the field of mathematics education, an LDA analysis conducted on publications from 2019 to 2023 in the top five academic journals indexed by the Web of Science (WoS) has identified 18 research themes. Out of these themes, 10 show similar characteristics to those identified by Inglis and Foster (2018) through their topic modeling analysis. Three of these themes have been directly matched using the same nomenclature: Curriculum, Proof and Argumentation, and Addition and Subtraction. The other seven themes were aligned based on similar key terms and a review of the literature, indicating they cover comparable scopes. These themes are Digital Environment, Algebraic Thinking, Fraction, Function Concept, Professional Development, Language, and Mathematical Discourse.

Correspondingly, Inglis and Foster (2018) named these themes as Dynamic Geometry and Visualization, School Algebra, Rational Numbers, Analysis, Teacher Knowledge and Beliefs, Multilingual Learners, and Observations of Classroom Discussion.

During the comparison of these research themes, both the literature and the keywords clustered under each theme were examined. For example, the theme named Dynamic Geometry and Visualization by Inglis and Foster (2018) was renamed as Digital Environment in our model. This change reflects an expanded use of digital technologies not limited to the teaching of geometry but encompassing a wider range of teaching areas, such as statistics and probability, as illustrated by research like Herford et al. (2023). The findings indicate that while some themes have maintained their relevance, others have evolved or expanded in scope to adapt to new educational technologies and methodologies.

Our LDA analysis has redefined the research theme previously named "Multilingual Learners" by Inglis and Foster (2018) as "Language." This reclassification stems from a broader examination of the role of language diversity and cultural concepts within mathematics education as discussed in recent studies. Ryan (2022) focuses on language diversity in classroom mathematics instruction, while Morris (2021) discusses the importance of language and culture in teaching probability. These insights have guided us to adopt a more encompassing term, "Language," reflecting the broad impact of linguistic factors across various aspects of mathematics education.

Our findings indicate that out of the 18 identified themes, 10 have maintained their relevance as research themes over the past five years. However, it is evident that some themes have evolved in response to new disciplines, approaches, and global events. Among the notable shifts observed through our LDA-based topic modeling analysis is the emergence of eight research themes: Mathematical Performance and Assessment, Lesson Study, Mathematical Modelling, Social Justice, Teacher Practice, Statistical Literacy, Prospective Teachers Noticing of Student Thinking, and Framework Design and Development.

Particularly, "Statistical Literacy" has gained prominence in the context of the COVID-19 pandemic, highlighting how significant societal and global events can influence the focus of mathematics education and teaching. The relevance of statistical literacy has been underscored by research such as Kwan et al. (2021), who examined the interpretation of COVID-19 data in Korea, and da Silva et al. (2021), who discussed integrating statistical education through graphical representations of data in Brazil. These discussions illuminate the critical role of statistics in understanding and responding to global crises within educational settings.

Furthermore, the integration of diverse disciplines in educational research, supported by systematic reviews and topic modeling studies such as those by Vijayan (2021) and Rodríguez and Pulido-Montes (2022), validates the broadening scope and interdisciplinary nature of current educational research. These shifts signify an adaptive and responsive mathematics education field that not only reacts to immediate educational needs but also incorporates broader socio-economic and cultural dimensions, thereby enriching the academic discourse and practical applications within the discipline.

The theme of "Social Justice" reflects a transformative impact on educational practices influenced by societal shifts toward greater equity and inclusion, such as gender and racial equality (Gutiérrez, 2013). This focus aligns with findings that suggest an ongoing evolution in how social justice is integrated within mathematics education (Leyva, 2021). In contrast to Inglis and Foster (2018), who might not have emphasized this dimension as prominently, our findings indicate a significant shift towards addressing these critical issues within educational settings.

Similarly, the theme of "Mathematical Modelling" has shifted towards a more process-oriented focus in recent years. While Inglis and Foster (2018) might have touched on this theme, the current emphasis is on the establishment of real-life situational problems and their resolution through

mathematical processes (Shahbari & Tabach, 2020). This suggests a transition from solving predefined problems to engaging with the processes and design of modeling in educational practices, marking a significant development in the application of mathematical modeling in education.

"Mathematical Performance and Assessment" is another theme that has evolved over time. Recent literature, such as that by Evans and Jeong (2023), discusses how assessments of mathematical knowledge and skills can impact student performance and how teaching can enhance this performance. This theme could be seen as an evolution of the "Constructivism" theme discussed by Inglis and Foster (2018), reflecting a shift towards understanding and improving instructional outcomes based on constructive feedback and learning theories.

Furthermore, the theme "Framework Design and Development" emphasizes the creation of systematic and structured approaches to enhance teaching processes and student learning (Simon, 2022). This theme shares similarities with "Mathematical Performance and Assessment" in that both focus on developing student-centered teaching strategies, providing a solid foundation for improving educational practices. These related themes highlight the importance of structured educational frameworks that are crucial for implementing effective teaching strategies and enhancing learning outcomes.

The theme "Prospective Teachers Noticing of Student Thinking" emphasizes the significance of teacher candidates recognizing and understanding students' thought processes. Research by Lee and Lee (2023) and Tyminski et al. (2021) illustrates how this awareness is integrated into instructional practices, providing valuable insights into teacher training, pedagogical knowledge, and student learning processes. This area supports the transformation from 'Pupil' to 'Student' as discussed by Herford (2023), and underscores the evolving focus of teacher training towards leveraging technology not just as a tool, but as an integral part of pedagogical strategies.

Herford (2023) also highlighted how teachers are increasingly shaping their research interests based on the pedagogical needs they encounter, rather than solely on available technological solutions. This is in line with the "Professional Development" theme identified in our model, which has shown significant overlap with "Teacher Practice." These themes underscore a growing interest in teacher education and practice over the last five years, as reflected in studies by Matranga and Silverman (2022), and indicate a broader trend within mathematics education focusing on effective and student-centered teaching methodologies.

Moreover, the "Lesson Study" theme, identified in our model, resonates with findings from Lendínez Muñoz et al. (2023) and Miyakawa and Winsløw (2019), illustrating how collaborative approaches among teachers can develop more effective teaching methods. This theme highlights the professional development tool's role in enabling mathematics educators to collaboratively enhance their instructional strategies, reflecting a shift towards more collaborative and reflective teaching practices.

These emerging themes not only signify a transformation in the priorities and approaches within mathematics education but also suggest a departure from the themes identified by Inglis and Foster (2018). Our analysis reveals a heightened emphasis on teacher professional development, reflecting a shift towards more nuanced and targeted educational strategies that address both teacher and student needs in contemporary educational settings.

This LDA-based topic modeling study has explored how research themes in the top five influential journals in mathematics education have evolved since the 2018 study by Inglis and Foster. We propose that systematic reviews be conducted on these eight newly identified themes to further understand their development and implications in the field. Such studies would facilitate tracking these emerging topics, documenting research under each theme, and gaining new insights that could guide future educational practices and research in mathematics education.

Conflicts of Interest: The authors declare no conflicts of interest.

Funding Details: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRediT Author Statement: Kenan Gökdağ and Mehmet Fatih Özmantar contributed to the conceptualization of the study. Kenan Gökdağ was responsible for the methodology, formal analysis, original draft preparation, and visualization. Mehmet Fatih Özmantar contributed to the review and editing of the manuscript and provided supervision throughout the project.

Ethical Statement: This study does not require ethical approval as it does not involve human or animal participants.

REFERENCES

Asmussen, C. B., & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data, 6*(1), 1-18.

Barwell, R. (2020). Learning mathematics in a second language: Language positive and language neutral classrooms. *Journal for Research in Mathematics Education, 51*(2), 150-178.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

Blei ve Blei, D. M., & Lafferty, J. D. (2009). Topic models. In *Text mining* (pp. 101-124). Chapman and Hall/CRC.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84.

Bittermann, A., & Fischer, A. (2018). How to identify hot topics in psychology using topic modeling. *Zeitschrift für Psychologie*.

Borba, M. C. (2021). The future of mathematics education since COVID-19: Humans-with-media or humans-with-non-living-things. *Educational Studies in Mathematics, 108*(1), 385-400.

Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, *3*(3), 191-209.

Chen, C. (2016), CiteSpace: A Practical Guide for Mapping Scientific Literature, Novinka, available at: http://cluster.cis.drexel.edu/∼cchen/citespace/books/.

Chen, H., & Luo, X. (2019). An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Advanced Engineering Informatics*, *42*, 100959.

Chen, G. A. (2023). Teacherly response-ability: Ethical relationality as protest against mathematical violence. *Educational Studies in Mathematics, 114*(2), 277-296.

Choi, J. A., & Kwak, M. H. (2019). Topic changes in mathematics educational research based on LDA. *Journal of Education & Culture*, *25*(5), 1149-1176.

Da Silva, A. S., Barbosa, M. T. S., de Souza Velasque, L., & et al. (2021). The COVID-19 epidemic in Brazil: How statistics education may contribute to unravel the reality behind the charts. *Educational Studies in Mathematics, 108*, 269-289. https://doi.org/10.1007/s10649-021-10112-6

Drijvers, P., Kodde-Buitenhuis, H., & Doorman, M. (2019). Assessing mathematical thinking as part of curriculum reform in the Netherlands. *Educational Studies in Mathematics, 102*(3), 435-456.

Dowker, A., Cheriton, O., Horton, R., & Mark, W. (2019). Relationships between attitudes and performance in young children's mathematics. *Educational Studies in Mathematics, 100*, 211-230.

Evans, T., & Jeong, I. (2023). Concept maps as assessment for learning in university mathematics. *Educational Studies in Mathematics, 113*(3), 475-498.

Gutiérrez, R. (2013). The sociopolitical turn in mathematics education. *Journal for Research in Mathematics Education, 44*(1), 37-68. https://doi.org/10.5951/jresematheduc.44.1.0037

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences, 101*(suppl_1), 5228-5235.

González-Forte, J. M., Fernández, C., Van Hoof, J., & Van Dooren, W. (2022). Profiles in understanding operations with rational numbers. *Mathematical Thinking and Learning, 24*(3), 230-247.

Gurcan, F., Dalveren, G. G. M., Cagiltay, N. E., Roman, D., & Soylu, A. (2022). Evolution of software testing strategies and trends: Semantic content analysis of software research corpus of the last 40 years. *IEEE Access, 10*, 106093-106109.

Herfort, J. D., Tamborg, A. L., Meier, F., Allsopp, B. B., & Misfeldt, M. (2023). Twenty years of research on technology in mathematics education at CERME: A literature review based on a data science approach. *Educational Studies in Mathematics, 112*(2), 309-336.

Inglis, M., & Foster, C. (2018). Five decades of mathematics education research. *Journal for Research in Mathematics Education, 49*(4), 462-500.

KNIME (2020), "KNIME j open for innovation", available at: https://www.knime.com/ (accessed 4September 2020).

Krawitz, J., Chang, Y. P., Yang, K. L., & Schukajlow, S. (2022). The role of reading comprehension in mathematical modelling: Improving the construction of a real-world model and interest in Germany and Taiwan. *Educational Studies in Mathematics, 109*(2), 337-359.

Ko, Y. Y., & Rose, M. K. (2021). Are self-constructed and student-generated arguments acceptable proofs? Pre-service secondary mathematics teachers' evaluations. *The Journal of Mathematical Behavior, 64*, 100912.

Kullberg, A., Björklund, C., Brkovic, I., & Runesson Kempe, U. (2020). Effects of learning addition and subtraction in preschool by making the first ten numbers and their relations visible with finger patterns. *Educational Studies in Mathematics, 103*(2), 157-172.

Kwon, O. N., Han, C., Lee, C., & et al. (2021). Graphs in the COVID-19 news: A mathematics audit of newspapers in Korea. *Educational Studies in Mathematics, 108*, 183-200. https://doi.org/10.1007/s10649-021-10029-0

Lee, M. Y., & Lee, J. E. (2023). An analysis of elementary prospective teachers' noticing of student pattern generalization strategies in mathematics. *Journal of Mathematics Teacher Education, 26*, 155-177. https://doi.org/10.1007/s10857-021-09520-5

Lendínez Muñoz, E. M., García García, F. J., Lerma Fernández, A. M., & Abril Gallego, A. M. (2023). Increase in self-efficacy in prospective teachers through theory-based lesson study. *Journal of Mathematics Teacher Education*, 1-26.

Leyva, L. A. (2021). Black women's counter-stories of resilience and within-group tensions in the white, patriarchal space of mathematics education. *Journal for Research in Mathematics Education, 52*(2), 117-151.

Martínez-Planell, R., & Trigueros, M. (2019). Using cycles of research in APOS: The case of functions of two variables. *The Journal of Mathematical Behavior, 55*, 100687.

Matranga, A., & Silverman, J. (2022). An emerging community in online mathematics teacher professional development: An interactional perspective. *Journal of Mathematics Teacher Education, 25*(1), 63-89.

Miyakawa, T., & Winsløw, C. (2019). Paradidactic infrastructure for sharing and documenting mathematics teacher knowledge: A case study of "practice research" in Japan. *Journal of Mathematics Teacher Education, 22*(3), 281-303.

Morris, N. (2021). Learning probability in the Kingdom of Tonga: The influence of language and culture. *Educational Studies in Mathematics, 107*, 111-134. https://doi.org/10.1007/s10649-020-10022-z

Nie, B., & Sun, S. (2017). Using text mining techniques to identify research trends: A case study of design research. *Applied Sciences*, *7*(4), 401.

Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Springer Science & Business Media.

Paul, M., & Girju, R. (2009, September). Topic modeling of research fields: An interdisciplinary perspective. In *Proceedings of the International Conference RANLP-2009* (pp. 337-342).

Rodríguez, M. L., & Pulido-Montes, C. (2022). Use of digital resources in higher education during COVID-19: A literature review. *Education Sciences, 12*(9), 612.

Ryan, U. (2022). Solidarity acts in the context of language diversity: Students' encounters with epistemological difference. *The Journal of Mathematical Behavior, 66*, 100946. https://doi.org/10.1016/j.jmathb.2022.100946

Shahbari, J. A., & Tabach, M. (2020). Features of modeling processes that elicit mathematical models represented at different semiotic registers. *Educational Studies in Mathematics, 105*, 115-135. https://doi.org/10.1007/s10649-020-09971-2

Simon, M. A. (2022). Contributions of the learning through activity theoretical framework to understanding and using manipulatives in the learning and teaching of mathematical concepts. *The Journal of Mathematical Behavior, 66*, 100945.

Tyminski, A. M., Simpson, A. J., Land, T. J., & et al. (2021). Prospective elementary mathematics teachers' noticing of children's mathematics: A focus on extending moves. *Journal of Mathematics Teacher Education, 24*, 533-561. https://doi.org/10.1007/s10857-020-09472-2

Warshauer, H. K., Herrera, C., Smith, S., & Starkey, C. (2023). Examining preservice teachers' noticing of equity-based teaching practices to empower students engaging in productive struggle. *The Journal of Mathematical Behavior, 70*, 101045.

Weingarden, M., Heyd-Metzuyanim, E., & Nachlieli, T. (2019). The realization tree assessment tool–examining explorative participation in mathematics lessons. *The Journal of Mathematical Behavior, 56*, 100717.

Williams, S. R., & Leatham, K. R. (2017). Journal quality in mathematics education. *Journal for Research in Mathematics Education, 48*(4), 369-396.

Woods, C., & Weber, K. (2020). The relationship between mathematicians' pedagogical goals, orientations, and common teaching practices in advanced mathematics. *The Journal of Mathematical Behavior, 59*, 100792. https://doi.org/10.1016/j.jmathb.2020.100792

Vijayan, R. (2021). Teaching and learning during the COVID-19 pandemic: A topic modeling study. *Education Sciences, 11*(7), 347.

Zengul, F. D., Zengul, A. G., Mugavero, M. J., Oner, N., Ozaydin, B., Delen, D., ... & Cimino, J. (2021). A critical analysis of COVID-19 research literature: Text mining approach. *Intelligence-based medicine, 5*, 100036.