J. Korean Soc. Math. Educ., Ser. D, Res. Math. Educ.
Vol. 24, No. 4, 313-335, December, 2021
http://doi.org/10.7468/jksmed.2021.24.4.313

What Topics Have Been Studied in Korean Mathematics Education for 15 Years: Latent Topic Modeling Analysis

Jihyun Hwang (Professor)
Korean National University of Education, jihyun-hwang@knue.ac.kr

(Received December 3, 2021; Revised December 16, 2021; Accepted December 27, 2021)

The purpose of this research is to identify topics discussed by Korean mathematics education studies and examine research trends for 15 years. I applied latent Dirichlet allocation (LDA) to the original text datasets including English abstracts of 3,157 articles published in eight journals indexed by the Korean Citation Index (KCI) from 1997 to 2019. I identified an LDA model with 60 topics, then research trends in 2,884 articles between 2002 and 2018 were as follows; mathematics educators have paid most attention to teacher education through 2010 to 2015 and curriculum analysis after 2016. The findings in this research can contribute to understand what have been discussed in Korean mathematics education society as well as what will and need to be emphasized more in the future compared to the global research trends. In addition, LDA has potentials to identify topics and keywords of manuscripts newly written and submitted to any journals in addition to information provided by authors.

Keywords: text mining, latent Dirichlet allocation, research trend.
MESC Classification: A30
MSC2010 Classification: 97A30

I. INTRODUCTION

Mathematics education in Korea has a long history, but the modern field of mathematics education has been started since 1953 when the Korean government established a modern educational system after the Korean war (J. S. Kim, 2012). As Leung (2006) argued, Koreans have laid “strong emphasis on the importance of education” (p. 43), which results in a gradual development of teacher education programs and research in mathematics education. Particularly, the 1990s is the decade where multiple mathematics education journals began to be published (Pang et al., 2019). In parallel with development of the mathematics education research field, the Korean Citation Index (KCI) has been also developed to encourage interactions in overall Korean research communities and to have valid evaluation tools to journals. KCI has applied to mathematics education and seven journals are found with the keyword like *mathematics education* and *school mathematics* in 2019.

Investigating articles indexed by KCI could be beneficial for international researcher in two folds: First, KCI can provide researchers opportunities to find high-quality studies which reflect unique contexts of Korea, as well as which are generalizable and applicable to other countries. Researchers all around the world can access to information about KCI studies through the databased provided by the Web of Science platform, sometimes including full texts. Second, the KCI database itself could be used for text mining research to reveal information from text, particularly to answer how has been the mathematics education research field in Korea has been changed. There are research databases like JSTOR[1] that provide datasets about articles for text mining. In addition, openness of the KCI database could be a great source to see research trends using a text mining approach. In this study, I utilize the KCI database to examine topics of Korean mathematics education studies from 2002 to 2018.

I applied one of the machine learning approaches, the latent Dirichlet allocation (LDA; Blei et al., 2003). This study can be a replication of the work done by Inglis and Foster (2018) and Shin (2020) recently, but this study used all terms except for stop words (Shin used only nouns and adjectives), different pre-processing, multiple indices to select the best model. I also should note that the main purpose of this study is not comparison between local and global trends although I discussed similarities and differences between them briefly. As Shin (2020) and Pang et al. (2019) mentioned, investigation on research trends could suggest pathways for future studies by understanding what topics have been emphasized or not. For this reason, there are many studies including this study that have investigated trends in Korean mathematics education research. However, most studies applied existing lists of codes (e.g., content areas and research methods) and some studies have examined article titles and keywords given by authors. However, with this machine learning approach called LDA, the latent topics are emerged from the dataset, not from pre-determined codes by researchers. Also, LDA allows me to utilize a larger-size database by saving time and efforts. And I can have rich information examining patterns in main texts or abstracts rather than titles and keywords.

[1] Available at https://www.jstor.org/dfr/

II. TRENDS IN KOREAN MATHEMATICS EDUCATION

A large body of literature has sought to find research trends in Korean mathematics education with various approaches. Y. Kim and Pang (2017) analyzed research articles about elementary mathematics education published since 2010 by coding the articles with multiple dimensions including research methods and content areas. According to Y. Kim and Pang, a significant number of studies are about affective constructs and textbook analysis. There are increasing attentions to teacher education, but fewer interests in technology and assessments. J.-E. Kwon and Choi (2008) also discussed research trends in elementary mathematics education with a list of codes (e.g., in-service or pre-service, students or teachers) ten years before the study of Y. Kim and Pang (2017). Similarly, curriculum or textbook analysis was one of the main topics, which is aligned to the findings of J.-E. Kwon and Choi (2008). A unique finding of J.-E. Kwon and Choi (2008) is that there are also many studies about instructional design and methods.

Some studies narrowed their foci to specific research areas like textbook analysis. According to the previous two studies (Y. Kim & Pang, 2017; J.-E. Kwon & Choi, 2008) about research trends in Korean elementary mathematics education, textbook analysis is one of the main interests. Furthermore, Hong and Cha (2005) provided more details about trends in textbook analysis studies. Hong and Cha argued that many studies conducted international comparison. Their finding showed that many studies have examined how to reflect the history of mathematics into designing textbooks, but more studies are required on how to include using technology in mathematics textbooks. The importance of textbook analysis is also supported by R. Y. Kim et al. (2012). Their study scrutinized 124 research articles about curriculum between 2000 and 2010 with a focus on trends in research design and methods. They found that 76% of the studies were about curriculum or textbook analysis, and those studies applied less rigorous analysis methods.

Previous studies about research trend in Korean have shown that curriculum and teacher education are the main research areas in Korean mathematics education. As discussed before, most studies utilized existing codes about topics, research methods, and content areas, which are possibly adapted from ZDM classification, MSC 2000 classification, PME classification. However, such topics are not emerged from a text dataset. Emerging topics might be very new, which means that it cannot be found in existing classifications. Furthermore, they might be a combination of multiple classifications because different dimensions of topics like research methods and content areas are not considered in topic identification process. To identify topics from text data, some alternative methods rather than coding are available. S. H. Kim and Kim (2018) applied one technique called the language network analysis. They collected and analyzed titles of 1201 papers presented in the PME 41, the PME-NA 39, the NCTM Annual

Meeting 2017 with focuses on frequencies of and connectedness between words in those titles.

Recently, there are two remarkable works on research trends in Korean mathematics education (Pang et al., 2019; Shin, 2020). The work of Shin (2020) was to compare global and local trends by applying LDA to two different sets of articles separately, indexed by KCI and the Social Science Citation Index. He found 16 similar topics from the text data including nouns and adjectives filtered by lemmatization and pre-processing. Because of possibility of over-simplification, my study as a replication study used a relaxed pre-processing. Furthermore, Pang et al. (2019) securitized 4,559 articles from 1963 to 2019 with a qualitative approach. They showed that teacher education, cognition and attitudes are a main research topic. They also argued necessity of research on learning with technology and assessments.

Grounding on the above literature, I applied LDA to identify topics discussed by Korean mathematics education studies from 1997 to 2019. This approach helps me find research trends appeared more than 2,800 articles published in eight KCI journals between 2002 and 2018. Based on the literature review, it would be interesting to see which words are characterizing articles about textbook and curriculum analysis as well as how many articles about textbook and curriculum analysis. In addition, it is expected to recognize some topics related to teacher education and their trends for 15 years because of increasing attentions to this research area (Y. Kim & Pang, 2017). Thus, the questions guiding this research are:

1. Through KCI-indexed research articles published from 2002 to 2018, what words characterize their latent topics uncovered in LDA?
2. What are research trends represented by the number of articles dealing with each latent topic since 2002?

III. LATENT DIRICHLET ALLOCATION

There have been increasing interests to the machine learning approach and some researchers have applied this approach to educational studies (e.g., Inglis & Foster, 2018; Lamb et al., 2019). One way to apply the machine learning approach is topic modeling with LDA (Blei et al., 2003), which is a three-level hierarchical Bayesian model. As an unsupervised machine learning technique, “[t]he basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words” (p. 996). In other words, topics are not directly observable cluster characterized by a distribution of words. This model is hierarchical because LDA model includes parameters at corpus level sampled in the process of a collection of documents (called corpus) as well as document-level and word-level variables. Because LDA involves corpus, document, and word levels, topics can be tested repeatedly within the document. Thus, multiple topics can be related to each document (Blei et al., 2003).

I highlight that both machine learning and traditional statistics approaches are concerned with the question how we learn from data while sharing some concepts with different terms - regression in statistics is corresponding to supervised learning in machine learning (Kuhn & Johnson, 2013). There are also differences between those approaches: Machine learning is mainly about predictions while statistics is about population, assumptions, and inferences. Thus, applying LDA enables researchers to see the patterns in topics of previously published articles as well as to predict topics previously unseen documents with well-defined inference procedures.

Although Inglis and Foster (2018) already applied LDA to examine how topics published in *the Journal for Research in Mathematics Education* and *the Educational Studies in Mathematics* have been changed for five decades, applying this method to articles written in Korean has some issues. For example, Korean has a different grammar structure from English, LDA with only nouns are applicable (Kang et al., 2013). An alternative way is using English abstract of each article because most Korean articles have their English abstracts. In this case, it is questionable if the lengths of article abstracts are long enough for LDA. However, there have been some studies with short documents like the content of Twitter (W. X. Zhao et al., 2011) and abstracts (Griffiths & Steyvers, 2004). Thus, I argue that LDA with English abstracts may discover topics discussed in Korean mathematics education journals. In the entire analysis procedure, the computer program R was used mainly with the R package *topicmodels* (Grün & Hornik, 2021).

I address some benefits from applying LDA above: latent topics emerged from patterns in a text dataset and a large-scale dataset could be analyzed with less time and efforts compared to qualitative studies on research trends. However, there are also some limitations of using LDA. First, naming latent topics should be done by researchers anyway although it is recommended that latent variables should not labeled because it is *latent* (Henson & Roberts, 2006). This indicates a risk that two latent topics are labeled with the same name, but it is possible that they are truly different. For example, Shin (2020) provided the label “algebra and algebraic thinking” to two latent topics – one is from the KCI model, and the other is from the SSCI model (see Table 3 in Shin, 2020). However, they have different sets of top 10 words, which means that they might be slightly different. Second, LDA considers frequencies of terms, not meaning of terms. A single term could have various meanings, but LDA do not reflect contexts where each term is used.

IV. REARCH METHOD

1. DATA DESCRIPTION

I collected the data of 3,157 articles published in the following eight journals[2] in mathematics education indexed by KCI from 1997 to 2019:

- Journal of Educational Research in Mathematics (JERM)
- Journal of Elementary Mathematics Education in Korea (JEMEK)
- Journal of Korea Society Educational Studies in Mathematics: School Mathematics (SM).
- Journal of the Korean School Mathematics Society (JKSM)
- Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education (ME)
- Journal of the Korean Society of Mathematical Education Series C: Education of Primary School Mathematics (EPSM)
- Journal of the Korean Society of Mathematical Education Series D: Research in Mathematical Education (RME(K)[3])
- Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education (CME)

The number of articles per year and journal is reported in Table 1. I accessed all articles through the KCI-Korean Journal database on the Web of Science platform[4]. I collected the following variables on April 5th, 2019: article ID, authors, published journal, title, volume, issue, and starting page number, and English abstract. The dataset included articles with English abstracts and exclude articles written by editorial board and articles without English abstracts. Thus, the total number of articles analyzed in this research is 2,884. Furthermore, I will report the results about articles published since 2002. This is because there were only at most 10 articles available in each year before 2002.

Although the data have texts of English abstracts, it is required to remove unnecessary parts when I tokenize texts into individual words. I removed all stop words provided in the R package tm with the option of "SMART." In addition, I added more terms to this set of stop words like education/educational, research/study, mathematics/math, and article/paper, which are possibly irrelevant to specific topics in mathematics education or too general to specify topics.

---

[2] In alphabetical order.

[3] I add "K" to distinguish this journal from Research in Mathematics Education, which is the official journal of the British Society for Research into Learning Mathematics.

[4] Retrieved from https://apps.webofknowledge.com/KJD_GeneralSearch_input.do?product=KJD&SID=5FlE7s38dgoyZQVu1Dv&search_mode=GeneralSearch

2. MODEL SPECIFICATION

Because LDA is an unsupervised machine learning approach, the number of topics cannot be pre-determined. It is necessary to evaluate models to avoid overfitting or underfitting the data. To identify the number of topics, I examined the four indices together: the maximum value of two indices suggested by Griffiths and Steyvers (2004), and Cao et al. (2009); and the minimum value of the other two indices provided by Arun et al. (2010), and Deveaud et al. (2014). Based on Figure 1, each index has a different recommendation: from 50 to 150 topics by the index of Cao et al. (2009), and between 50 and 200 by the index of Arun et al. (2010). The index of Deveaud et al. (2014) suggests that the number of topics is less than 25, but according to the index of Griffiths and Steyvers (2004), it ranges from 50 to 100. Except for the index of Deveaud et al. (2014), the indices show that the number of topics should be above 50. There is no rule of thumb in selecting the number of topics with all indices. However, comprehensively considering all indices, I argued that somewhere between 50 and 100 topics is optical for my dataset.

Table 1. The Number of KCI Articles published in 8 Mathematics Education Journals from in the Web of Science Database

Year	CME	EPSM	JEMEK	JKSM	RME(K)	SM	JERM	ME	Total
1997			9						9
1998			5						5
1999			5						5
2000			3						3
2001			7				3		10
2002			5	15		39	25	24	108
2003			6	19		30	26	44	125
2004			10	14		22	24	28	98
2005			9	29		24	25	37	124
2006			11	32		22	19	31	115
2007	36		10	32		27	24	31	160
2008	33		10	31		31	25	32	162
2009	52		15	35		37	31	27	197
2010	46	8	44	33	26	37	30	31	255
2011	36	20	30	29	25	36	24	30	230
2012	20	16	23	36	18	30	32	27	202
2013	32	19	27	42	18	52	33	32	255
2014	31	16	28	36	18	40	34	29	232
2015	36	11	32	24	14	28	42	21	208
2016	27	22	35	23		44	41	26	218
2017	28	21	29	26		40	42	23	209
2018	29	25	24	22		38	29	24	191
2019	3	7	9				9	8	36
Total	409	165	386	478	119	577	518	505	3157

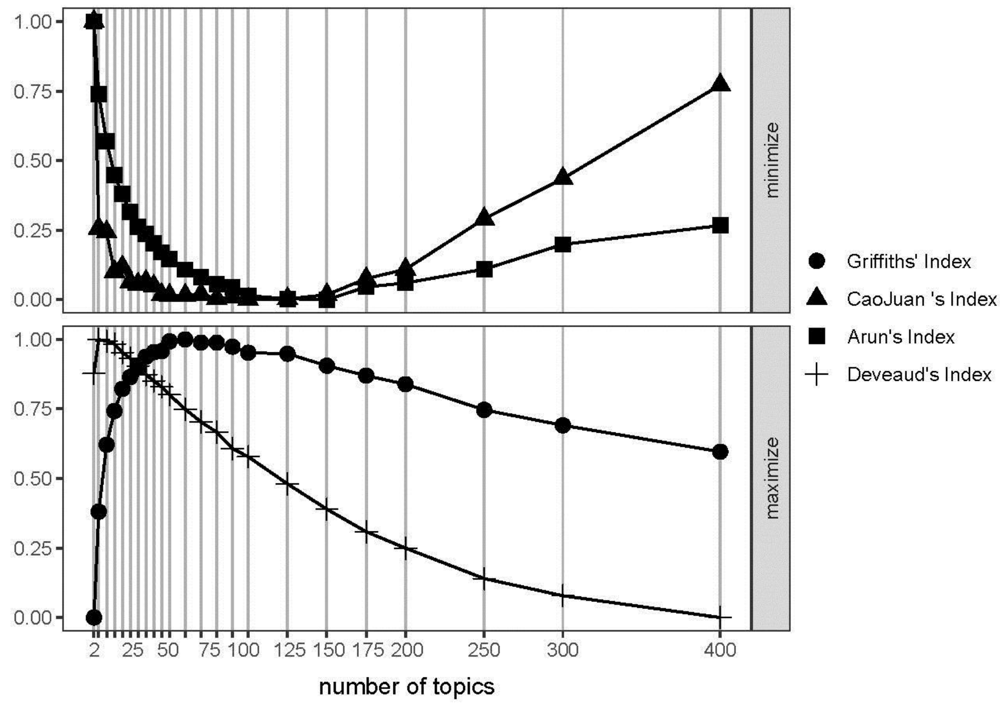

Figure 1. Results of Four Indices

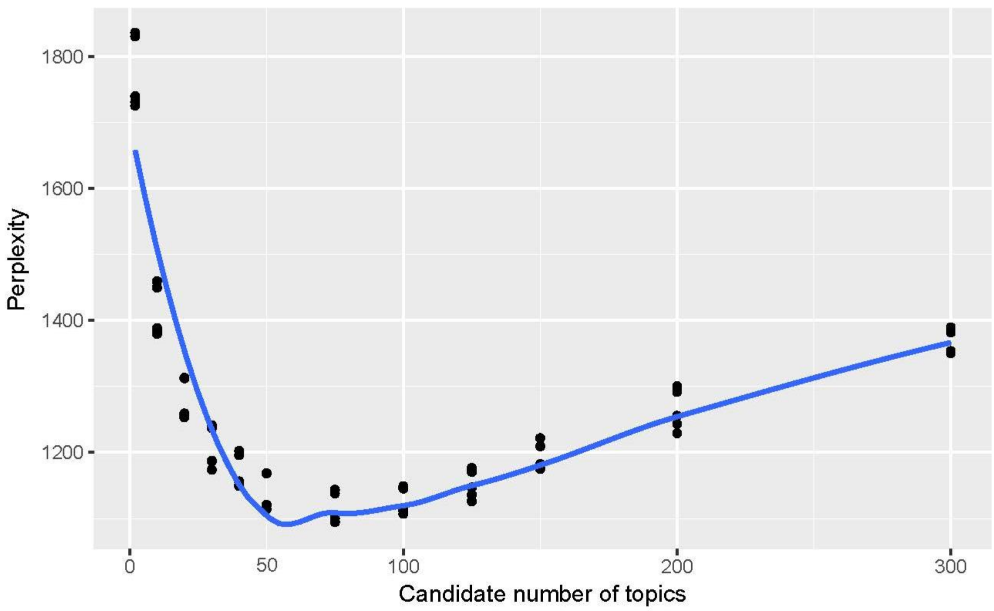

Figure 2. Five-Fold Cross-Validation of Topic Modelling with the KCI Database

I also utilized the perplex-based method as Inglis and Foster (2018) did. W. Zhao et al. (2015) said “Perplexity [italic added] is a commonly used measurement in information theory to evaluate how well a statistical model describes a dataset, with lower perplexity denoting a better probabilistic model” (p. 2). The difference from the study of Inglis and Foster (2018) is that a five-fold cross-validation with various numbers of topics was used. The key idea of five-fold cross-validation is that the data is divided into five equal-size subsets randomly, and one of the five subsets is used as a validation set while the others are used as a training set. This process is repeated five times using each subset in turn as the validation set. The cross-validation is useful for evaluating the overall model on the dataset that wasn’t involved in its training.

As seen in Figure 2, the results of perplexity were aligned with the interpretation of the four indices before. It is recommended between 50 and 75 topics. But perplexity of the model with 60 topics is likely to have the lowest value. Lastly, I conducted LDA with 60 topics on all 2,884 articles.

V. RESULTS

1. TOPICS IN KOREAN MATHEMATICS EDUCATION RESEARCH

I interpreted some of the 60 identified topics using the words with high probabilities to appear in research articles of each topic. For example, the following terms are characterizing the fifth topic in Table 2: curriculum, textbooks, revised, national, textbook, contents, elementary, Korean, and analysis. In addition, I examined articles that have the high probabilities to discuss each topic (in most cases, the probability is above 0.75). For example, the article with very high probability to be about the fifth topic is the work of Chang et al. (2016), which investigated the connections between curriculum and textbooks at fifth and sixth grade levels. Studying these studies also contributes to selection of labels of topics.

On one hand, I should be cautious in naming identified topics with specific terms because of the nature of an unsupervised machine learning approach (Henson & Roberts, 2006). Because LDA is an exploratory analysis to find latent topics, naming topics possibly results in highlighting only one aspect of each identified topic ignoring subtle nuances. On the other hand, naming topics are inevitable to clarify research trends. When referring to the topics suggested by Inglis and Foster (2018), I was able to compare research trends in Korea to their results. Because of space limitation, I selected 30 topics reported in Table 2, which many articles have discussed. Some of these topics are also comparable to the topics in the study of Inglis and Foster (2018).

Table 2 shows the 30 topics with at most 20 words charactering each and one exemplar article. To avoid any confusion in the results and discussion sections, numbers are assigned to the topics. For better understanding of all 60 topics and further discussion to characterize the topics, 20 words for each topic with high probability are available at [webpage address will be share later. Because all topics emerge from patterns in the data, these topics are not fitted into dimensions like content areas and research methods that prior studies have used to examine research trends. Some topics can be characterized by participants (e.g., gifted education and college students) or content areas (e.g., fractions). There are multiple topics about teacher education (e.g., teacher knowledge, teacher training/practicum). It should be noted that a term can be found to characterize different topics.

Table 2. The 30 topics with the words characterizing the topics and the article with the high probability to dealing with each topic

Topic Name (Alphabetical order)	Characteristic Words (Top 20)	Paper with the High Probability
1 Algebraic Expressions Formula	area students figures figure geometric activities geometrical algebra plane formula geometry understanding triangle inquiry variables letters solution methods areas volume	A study on students' understanding of letters and algebraic expressions in solving algebraic word problems with Excel (Ryu & Kim, 2004)
2 Analysis of Error	students errors error representations representation process language types understanding concepts results derivative learning discourse showed van type low high	Analysis of error on the process of solving the liner inequality- focusing on curriculum of the middle school - (Y. H. Kim & Oh, 2002)
3 Anxiety & Gender Difference	achievement students anxiety differences gender grade results school factors female characteristics high male variables naea analysis elementary academic difference student	Relationship between mathematics anxiety and mathematical achievement of middle school students according to gender and grade1 (S. Hwang & Lew, 2018)
4 Calculus (Real world or everyday life)	life real students series integral limit definite digital learning sequence concept map activities calculus infinite experience process mind textbook important	An alternatives of the definition of definite integral in <MathematicsⅡ> textbook under 2015-revised curriculum (Shin & Cho, 2018)
5 Cognitive Demand of Tasks	tasks task cognitive level students analysis noticing convergence connections percentage demand types answers high correct suggested purpose findings analyzed	The analysis of mathematical tasks in the high school mathematics (M. Kim & Kim, 2013)

6 College Entrance Examination and Freshmen	college university test basic ability engineering scholastic calculus science examination freshmen year students academic entrance achievement major level system survey	An analysis of the change in mathematical inclination of middle level engineering college freshmen (G. H. Lee & Lee, 2015)
7 Cooperative Learning & Attitude	group learning achievement experimental attitude groups students effects academic positive control effect small motivation cooperative significant method experiment showed difference	The effects of learning mathematics according to feedback method (Seo, 2007)
8 Curriculum implementation /Revised Curriculum	curriculum textbooks revised national textbook contents elementary Korean analysis analyzed competencies school implications revision terms results methods result related current	Analysis on connection of curriculum and textbooks in elementary school mathematics: Focused on 5~6 grades (Chang, et al., 2016)
9 Elementary Textbook Analysis	textbook textbooks first second elementary third angle measurement presented point angles curriculum length school fourth Korean definition related definitions	Research on the definitions of angle in the past Korean elementary mathematics textbooks (Soomi Kim, 2018)
10 Fractions	fraction division fractions knowledge students number algorithm teaching unit understanding mkt arithmetical quotient concepts elementary concept operation context analysis divisor	Division of fractions in the contexts of the inverse of a cartesian product (Yim, 2007)
11 Function in Calculus	function storytelling concept students functions teaching change graph calculus understanding concepts point curve tangent distance continuity definition method continuous constant	Analysis for the concept of smooth curve by velocity (M. S. Choi, Jeong, & Kim, 2012)
12 Gifted Education	gifted students mathematically elementary generalization characteristics ability regular second first children creative purpose results attitude career affective analysis third test	Comparative study between gifted math elementary students and non-gifted students in emotional intelligence and creative nature (E. H. Lee & Rim, 2014)
13 International Comparison	Korea Korean countries Japan South North school China curriculum states differences Singapore united compare compared content teaching on	Comparison of early tertiary mathematics in USA and Korea (S.-G. Lee, Seol, & Ham, 2009)
14 Item Difficulty and Cognitive load	writing difficulty test students data difficulties process journal item results presentation variables eye cognitive performance ma attitudes change regression	Estimating the regression equations for predicting item difficulty of mathematics in the College Scholastic Ability Test (S. H. Lee et al., 2007)

15 Mathematics Teaching Evaluation standards	evaluation standards teaching knowledge elements curriculum instruction instructional focused content essay understanding students methods teachers including subject skills development developed	The study on the investigation of the evaluation standards for mathematics teaching according to the teacher's opinion research (H. J. Hwang, 2013)
16 Novel Assessments	assessment teachers activity descriptive process students elementary performance method school assessments play student learning results observation focused interview portfolio developed	A study on the classroom application of observation assessment of mathematics assessments (Lee et al., 2013)
17 Operations with Whole numbers or Fractions	multiplication children operations addition fractions numbers subtraction arithmetic students division understanding teaching calculation understand elementary algorithm natural counting operation order	An analysis of students' understanding of operations with whole numbers and fractions (K. M. Kim & Whang, 2012)
18 Problem Solving	problem solving students process ability solve method knowledge student activity analysis result purpose information solution elementary understanding abilities case methods	Development and analysis of effect for problem solving model of student-based (Jung & Roh, 2014)
19 Proportional Reasoning	reasoning proportional proportion strategy students formal teaching ratio strategies multiplicative graders quantitative informal inductive results situations additive tasks elementary comparison	Teaching proportional reasoning in elementary school mathematics (Chong, 2015)
20 Proof in Geometry	lessons proof students proofs theorem lesson activities Geogebra proving symmetry results related understanding figures tangram fundamental career congruence gsp rules	Analysis of various proofs of Pythagorean theorem (Y. R. Kim, Noh, & Son, 2009)
21 Sociocultural theory	classroom students factors teacher practices intuitive social learning teaching intuition instruction culture student role preference decision important process system classrooms	A study on interaction between social practices and identities in elementary mathematics classroom (J. R. Kwon, 2007)
22 Spatial reasoning	posing spatial sense activities ability students visualization mental transformation geometric solid students factors space grade school activity figures orientation elementary	The application of embodied turtle schemes for the task of the spatial visualization (J. Y. Lee, Cho, & Song, 2013)
23 Statistics Education	data sample statistical sampling variability distribution students understanding population confidence results samples concept distributions control concepts normal disliking standard representativeness	The design and implementation to teach sampling distributions with the statistical inferences (Y. Lee & Lee, 2010)

24 Teachers' knowledge (PCK)	teacher teachers knowledge teaching content development professional pedagogical PCK prospective training elementary instruction programs community practice subject classroom program reflection	The research on pedagogical content knowledge(PCK) focused on instructional consulting for secondary beginning teachers (Choe & Hwang, 2009)
25 Teacher Motivation	teaching teachers data analysis reading learning classroom results characteristics understand understanding interviews time collected theory aspects practice effective	A qualitative case study about mathematics pre-service teachers' motivation, ways of dealing with university mathematics knowledge and tutoring experiences (Jun, 2013)
26 Teacher Training/ Practicum	teachers service pre teaching elementary survey teacher secondary results perception purpose perceptions experience investigate showed practicum participated training first	Elementary pre-service teachers' uses of mathematics teaching expertise according to the number of their mathematics instructions in the teaching practice (Gyu, 2018)
27 Teaching or Learning Programs	learning teaching program students method effective school develop developed activities instruction methods results process applied effects directed attitude purpose application	Mathematics instruction that classified by level, it applied the portfolio (J. I. Lee & Choi, 2007)
28 Technology	students computer mathematization gsp learning target software base student conceptual discrete project web process regular contents system method designed experiments	Development of a CAS-based virtual learning system for personalized discrete mathematics learning (Y. Jun, Kang, Kim, & Jung, 2010)
29 Triangles & Congruence	linear algebraic triangle students properties algebra calculators equations functions equation cas conditions geometric school method graphing related concepts calculator lines	Didactical analysis on triangle-determining conditions and triangle-congruence conditions (Yim, 2005)
30 Word Problems	problems type word problem solving students solve types questions situations solutions first second solved understanding situation answer showed graders made	Linguistic and cognitive factors that affect word problem solving (Sunhee Kim, 2004)

2. RESEARCH TRENDS OVER 15 YEARS

After identifying the topics in Korean mathematics education, I examined how many articles have been published from 2002 to 2018 for each topic. In this research, I will report the number of articles dealing with each topic rather than the proportion of the articles. This means that equal weights are assigned to the articles to examine research trends although there is variation in the total number of articles published in each year. In general, Table 3 shows that the topics that have most frequently discussed are teacher knowledge (99 articles), curriculum implementation (90 articles), teacher training (78 articles), and triangle and congruence (77 articles). Also, it is interesting that most topics in Table 3 have been covered continuously through 20 years.

Table 3. The number of research articles by topic (top 20)

	Topic	Year																	
		02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Total
24	Teacher Knowledge	0	0	3	2	6	5	7	8	6	8	10	7	8	9	7	6	7	99
8	Curriculum Implementation	1	0	3	3	1	0	4	2	11	3	9	8	5	7	8	13	10	90
26	Teacher Training	1	0	0	1	1	0	1	1	5	8	6	12	11	9	9	5	8	78
29	Triangles & Congruence	0	0	3	7	1	2	9	15	9	9	1	3	2	5	2	4	2	77
10	Fractions	0	1	0	3	3	2	2	5	7	3	8	6	9	4	6	4	3	68
16	Novel Assessments	2	0	3	1	4	4	6	2	3	5	6	9	4	2	2	6	4	65
27	Teaching or Learning Programs	0	0	6	4	2	1	3	3	4	4	4	7	5	6	3	2	4	64
11	Function in Calculus	0	0	1	2	3	1	7	2	5	3	8	6	4	2	6	9	4	63
2	Analysis of Error	1	0	2	2	0	5	3	4	7	2	6	6	5	3	8	6	3	63
6	College Entrance	0	0	2	2	0	6	5	8	3	11	1	6	2	7	4	4	1	63
3	Anxiety & Gender Difference	0	0	3	2	5	3	3	4	3	5	8	4	5	1	4	3	5	60
17	Operation with Numbers	0	0	2	1	1	2	1	7	5	3	5	1	4	2	5	8	8	58
25	Teacher Motivation	0	0	1	1	2	1	0	2	6	6	5	7	8	5	3	5	2	55
12	Gifted Education	0	0	2	2	5	3	3	4	2	4	4	8	7	0	4	1	2	53
7	Cooperative Learning	2	0	2	2	2	2	4	4	7	3	2	2	6	3	3	3	6	53
1	Algebraic Expressions	0	0	2	1	1	3	8	4	8	8	4	3	4	2	3	1	0	53
5	Cognitive Demand of Tasks	0	0	1	2	0	2	0	0	2	2	4	7	3	9	6	10	2	52
18	Problem Solving	1	1	1	4	3	4	2	3	4	5	3	2	2	4	2	3	5	51
13	International Comparison	0	0	3	2	3	1	2	2	6	1	3	2	8	6	5	1	4	51
28	Technology	1	1	4	3	1	1	5	2	9	7	3	2	0	4	1	0	2	48
	Total	108	125	98	124	115	160	162	197	255	230	202	255	232	208	218	209	191	3157

With a combination of teacher knowledge, teacher training, and teacher motivation, 232 articles (about 7%) of the KCI articles are about teacher education. Since there are 60 topics, this number of the articles about teacher education is remarkably high. As seen in Table 3 and Figure 3, teacher training/practicum has received researchers' increasing attentions around 2010 although researchers have continuously published multiple articles about teacher knowledge even before. The number of published articles about teacher education was at maximum approximately in 2013 and 2014, but decreases recently.

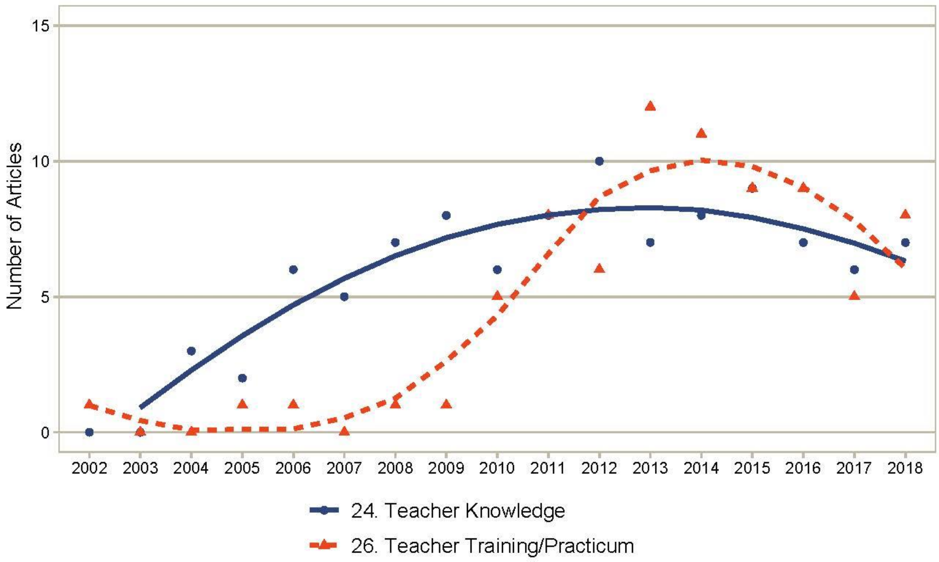

Figure 3. Trends in Teacher Education Research

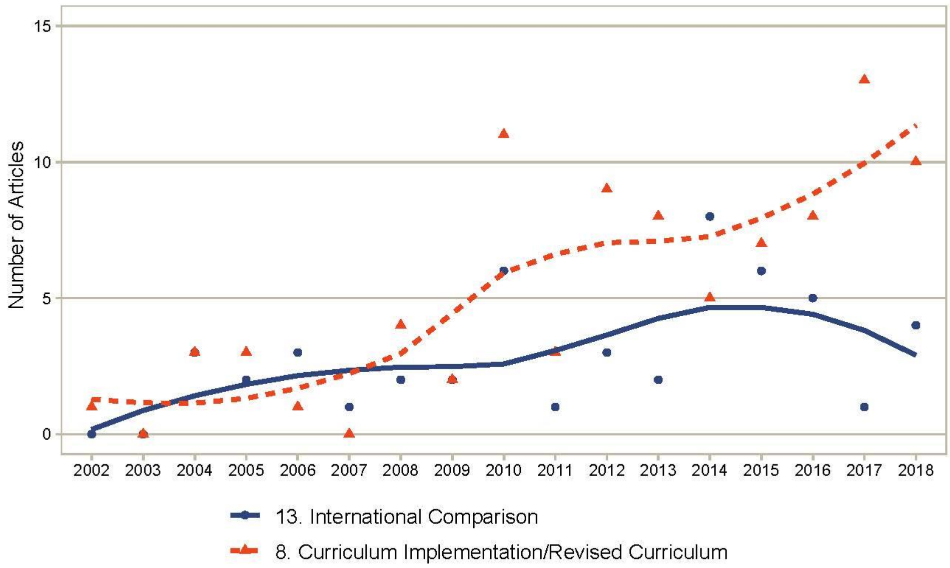

Figure 4. Trends in International Comparison or Curriculum Research

Shifting a focus from teacher education to curriculum implementation, I found that there are the increasing number of articles about textbooks, tasks, curriculum implementation and revision of curriculum. Another topic receiving much attention recently is cognitive demand of tasks (see Table 3). In addition, Figure 4 shows that international comparison studies have continuously published in the KCI journals. Some of these studies compare Korean mathematics curriculum to others' (e.g., Japan, North Korea, United States, and Singapore) because the term curriculum is one of the terms characterizing this topic (see the topic number 13 in Table 2). Then, it is reasonable to think that articles about curriculum analysis is the most notable research topic lately.

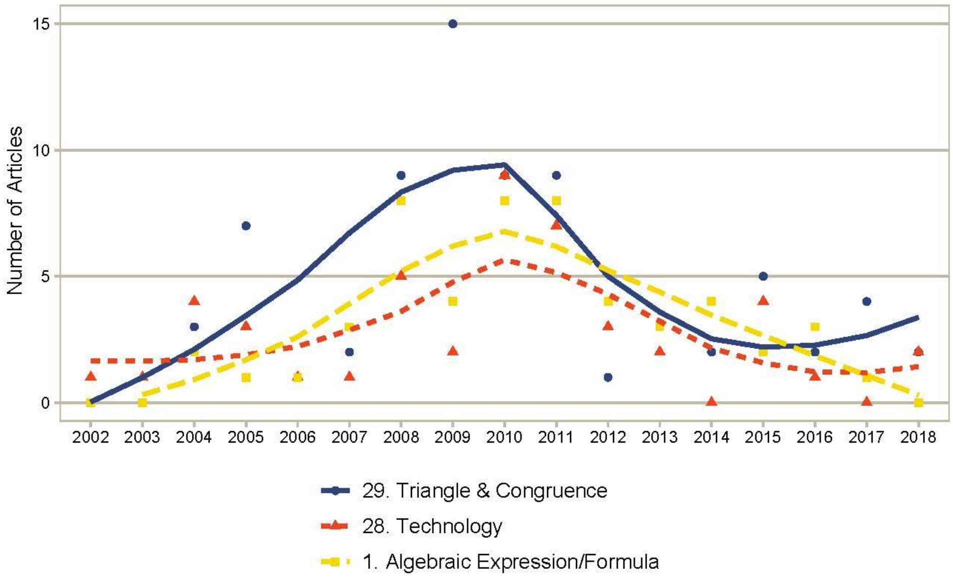

Figure 5. Trends in research about triangle, technology, and algebraic expressions/formula

Figures 5 and 6 show that mathematics educators have paid most attention to teacher education from 2010 to 2015 and curriculum analysis after 2010. Then, which topics were the main interests of Korean mathematics educators before 2010? Answering this question, Figure 5 shows three topics having bell-shape patterns with the maximum around 2010; triangle and congruence, technology, and algebraic expression/formula. In detail, Table 3 shows that topics related to content areas except for operation with numbers could show this pattern. It is also interesting that there were many studies about use of technology in mathematics classrooms in 2010 and 2011, but few studies since 2012.

I also examined topics related to assessments because Inglis and Foster (2018) showed that there are the increasing number of articles about novel assessment. Figure 6 shows that research about assessments with observation, interview, or portfolio has received increasing attentions until 2013. However, the findings about Korean journals are not corresponding to the finding Inglis and Foster (2018). Rather, I find no remarkable variance in the number of studies about assessments across years when I consider the number of articles about novel assessments in 2013 as an outlier.

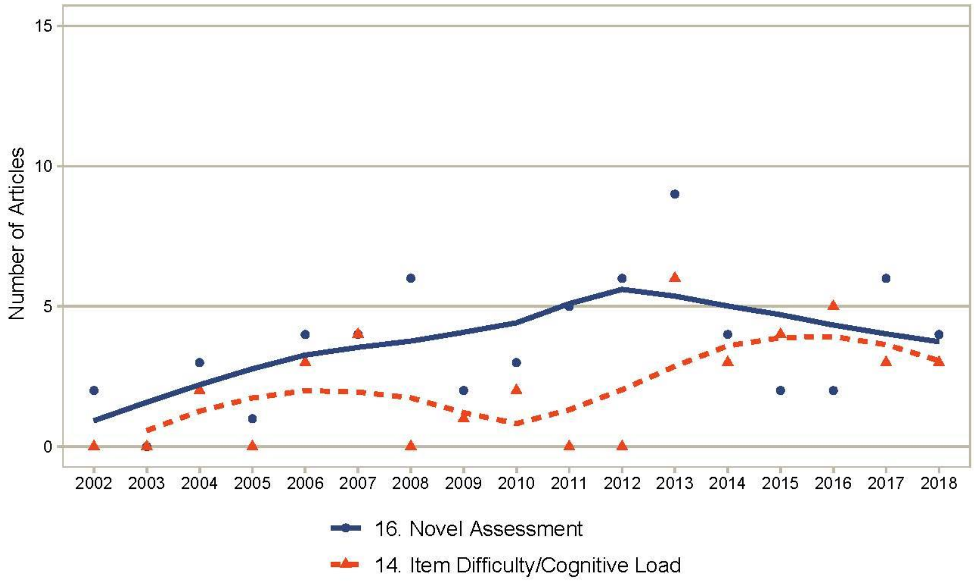

Figure 5. Trend in research about assessments

VI. DISCUSSION AND CONCLUSION

The aim of this research was to identify topics studied in Korea mathematics education by applying Latent Dirichlet Allocation (LDA). Then, I investigated research trends from 2002 to 2018. The model with 60 topics was selected based on the multiple indices and perplexity. It is interesting that more topics were identified using English abstracts of Korean mathematics education articles considering the 28 topics suggested by Inglis and Foster (2018). This might be because of two reasons. First, there are more specific topics situated in Korean education contexts, for example, college entrance exam and the Mathematics Teaching Evaluation Standards. Second, I collected the data from the seven KCI journals, not one or two leading journals. This resulted that more topics were discussed, and they are recognized in this research.

Korean mathematics educators have published a significant number of articles about teacher education for 15 years – teacher knowledge, teacher training/practicum, and teachers' motivation. However, I found an increasing number of articles about curriculum analysis with slightly less emphasis on teacher education. Furthermore, I found some similarities and differences between the research trend in Korean mathematics education journals, and that in JREM and ESM. I highlight that JRME also have published an increasing number of studies about curriculum analysis in more recent days. Moreover, there are similar patterns from 2002 to 2018 between dynamic geometry in Inglis and Foster (2018) and technology in this research. However, the two highest-quality journals, JRME and ESM (Nivens & Otten, 2017), have more attentions to the topic, *novel assessments*, which is about new approaches to evaluate students' achievement like using portfolio. Korean mathematics education journals have stably published articles about this topic. Another difference is that only few topics in Korean mathematics education are about theoretical perspectives and methodology. This is partially because I analyzed abstracts only, which might not discuss methods and theoretical framework in detail.

In this research, I tried to find the topics discussed in Korean mathematics education and their research trends over 15 years. In addition to identifying topics and research trends, latent Dirichlet allocation has potentials to identify topics and keywords of manuscripts newly submitted to journals when all latent topics are clarified and understood well. This technique could help editors have more information to classify new manuscripts in addition to self-reported classification and keywords. For example, I additionally collected English abstracts of two new articles not examined in this research: H. Lee (2019) and H. R. Choi and Shin (2019). H. Lee (2019) analyzed a teacher training program designed to help elementary teachers improve their content knowledge. The established topic model indicates that this article is about teacher knowledge with probability of 0.35. In addition, the study of H. R. Choi and Shin (2019) analyzed elementary mathematics textbooks following the 2015 revision curriculum. The topic model in this research suggests that this research is about curriculum implementation (the probability is 0.26). If providing only one topic is not enough, it is possible to examine the second and third candidates of relevant topics. In this way, researchers and journal editors can receive additional information about topics of manuscripts with the machine learning approach.

A limitation of this research is that characterizing and labelling the topics in Table 3 seems less systematic. It could be difficult to find an appropriate label that covers all 20 terms characterizing one latent topic. For example, Topic 1 is labelled with algebraic expression/formula although it is found that geometric, geometrical, and geometry are the terms characterizing Topic 1. To overcome this issue and provide more appropriate topic labels, I carefully read the articles dealing with each topic. The topics listed in Table 3 should be continuously improved. However, it should be noted that labelling topics is different from coding articles qualitatively. The topics are given by the LDA, not by a coding procedure.

In conclusion, Korean mathematics educators have increased their understanding about learning and teaching mathematics while publishing numerous articles about a wide range of topics. I hope that this research contributes to understand what have been discussed in Korean mathematics education society as well as what will and need to be emphasized more in the future.

REFERENCES

Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010). *On finding the natural number of topics with latent Dirichlet allocation: Some observations.* Paper presented at the 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2010, Hyderabad, India.

Blei, D. M., Ng, A. Y., & Jorda, M. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. Neurocomputing, 72, 1775–1781. https://doi.org/:10.1016/j.neucom.2008.06.011

Chang, H., Kang, T., & Lim, M. (2016). Analysis on connection of curriculum and textbooks in elementary school mathematics: focused on 5~6 grades. *Journal of Educational Research in Mathematics, 26*(1), 121–141.

Choi, H. R., & Shin, H. G. (2019). Analysis of mathematics textbooks in elementary school 1~2 grade from the perspective of project 2061. *Journal of Korea Elementary Education, 30*(1), 101–118. https://doi.org/: 10.20972/kjee.30.1.201903.101

Choi, M. S., Jeong, D. R., & Kim, J. S. (2012). Analysis for the concept of smooth curve by velocity. *The Journal of Educational Research in Mathematics*, 22(1), 23–38.

Choe, S. H., & Hwang, H. J. (2009). the research on pedagogical content knowledge (PCK) focused on instructional consulting for secondary beginning teachers. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics*, 11(3), 369–387.

Chong, Y.-O. (2015). Teaching proportional reasoning in elementary school mathematics. *Journal of Educational Research in Mathematics, 25*(1), 21-58.

Deveaud, R., SanJuan, É., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique, 17*(1), 61–84. https://doi.org/:10.3166/dn.17.1.61-84

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Science (PNAS), 101*(Suppl. 1), 5228–5235. https://doi.org/:10.1073/pnas.0307752101

Grün, B, & Hornik, K. (2021). *topicmodels*: Topic models. R package version 0.2-12.

Gyu, K. H. (2018). Elementary pre-service teachers' uses of mathematics teaching expertise according to the number of their mathematics instructions in the teaching practice. *Journal of Elementary Mathematics Education in Korea, 22*(1), 1–24.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66*(3), 393–416. https://doi.org/:10.1177/0013164405282485

Hong, M., & Cha, I. (2005). A meta analysis on comparative studies of mathematics textbooks. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 44*(2), 201–214.

Hwang, H. J. (2013). The study on the investigation of the evaluation standards for mathematics teaching according to the teacher's opinion research. *Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education, 27*(1), 39–62.

Hwang, S., & Lew, K. (2018). Relationship between mathematics anxiety and mathematical achievement of middle school students according to gender and grade. *Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education, 32*(2), 175–189.

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

Jun, P. G. (1999). Korean mathematics education: Past, present, and future (in Korean). *Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education, 9*, 327– 333.

Jun, Y. (2013). A qualitative case study about mathematics pre-service teachers' motivation, ways of dealing with university mathematics knowledge and tutoring experiences. *Journal of Educational Research in Mathematics, 23*(2), 269–284.

Jun, Y., Kang, Y., Kim S. H., & Chul, J. (2010). Development of a CAS-based virtual learning system for personalized discrete mathematics learning. *Journal of the Korean School Mathematics, 13*(1), 125–141.

Jung C. S. & Roh, E. H. (2014). Development and analysis of effect for problem solving model of student-based. *Journal of the Korean Society of Mathematical Education Series C: Education of Primary School Mathematics*, *17*(1), 57–75.

Kang, B., Song, M., & Jho, W. (2013). A study on opinion mining of newspaper texts based on topic modeling. *Journal of The Korean Society for Library and Information Science, 47*(4), 315–334.

Kim, J. S. (Ed.). (2012). *Education in Korea*. Korea Institute for Curriculum and Evaluation.

Kim, K. M, & Whang, W.-H. (2012). An analysis of students' understanding of operations with whole numbers and fractions. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 51*(1), 21–45.

Kim, M. & Kim, G. Y. (2013). The analysis of mathematical tasks in the high school mathematics. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 15*(1), 37–

59.

Kim, R. Y., Kim, G., & Kwon, N. (2012). Trends in research design and methods: Research on elementary and secondary mathematics curriculum. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 14*(3), 395–408.

Kim, S. [Sunhee] (2004). Linguistic and cognitive factors that affect word problem solving. *Journal of Educational Research in Mathematics*, *14*(3), 267–281.

Kim, S. [Soomi] (2018). Research on the definitions of angle in the past Korean elementary mathematics textbooks. *Journal of Educational Research in Mathematics, 28*(3), 265–282.

Kim. S. H., & Kim, S. M. (2018). The international research trends of mathematical education by language network analysis method - Based on the titles of 2017 international conference. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 20*(4), 591–608.

Kim, Y., & Pang, J. (2017). Research trends in elementary mathematics education: focused on the papers published in domestic journals during the resent seven years. *Journal of the Korean Society of Mathematical Education Series C: Education of Primary School Mathematics, 20*(1), 19–36. https://doi.org/: 10.7468/jksmec.2017.20.1.19

Kim, Y. R., Noh, H. S., & Son, E. H. (2009). Analysis of various proofs of Pythagorean theorem. *Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education, 23*(3), 887–921.

Kim, Y. H. & Oh, H. J. (2002). Analysis of Error on the process of solving the liner inequality-Focusing on curriculum of the middle school. *Journal of the Korean School Mathematics, 5*(1), 69–86.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer.

Kwon, J.-E., Choi, J.-H. (2008). An analysis of trends in elementary mathematics education research - Focusing on mathematics education journals in Korea. *Journal of Elementary Mathematics Education in Korea, 12*(2), 149–163.

Kwon, J. R. (2007). A study on interaction between social practices and identities in elementary mathematics classroom. *Journal of the Korean Society of Mathematical Education Series A: Mathematical Education, 46*(4), 389–406.

Lamb, R., Firestone, J., Schmitter-Edgecombe, Maureen, & Hand, B. (2019). A computational model of student cognitive processes while solving a critical thinking problem in science. *Journal of Educational Research, 112*(2), 243–254. https://doi.org/10.1080/00220671.2018.1514357

Lee, E. H. & Rim R. (2014). Comparative study between gifted math elementary students and non-gifted students in emotional intelligence and creative nature. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 16*(1), 181–199.

Lee, G. H., & Lee, J.-R. (2015). An analysis of the change in mathematical inclination of middle level engineering college freshmen. *Journal of the Korean Society of Mathematical Education*

Series E: Communications of Mathematical Education, 29(4), 745–762.

Lee, H. (2019). The effect of mathematics contents professional development programme for elementary school teacher. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 21*(1), 233–250.

Lee, J. I., & Choi, K. M. (2007). Mathematics instruction that classified by level, it applied the portfolio. *Journal of the Korean School Mathematics, 10*(4), 437–453.

Lee, J. Y., Cho, H. H., & Song, M. (2013). The application of embodied turtle schemes for the task of the spatial visualization. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 52*(2), 191–201.

Lee, K. S., Nan, H., Yong, S. H., Son, J. H., Jo, H. G., Lee, J. J., … Kang, O. K. (2013). A study on the classroom application of observation assessment of mathematics assessments. *Journal of the Korean School Mathematics, 16*(2), 289–318.

Lee, S.-G., Seol, H.-G., & Ham, Y. (2009). Comparison of early tertiary mathematics in USA and Korea. *Journal of the Korean Society of Mathematical Education Series E: Communications of Mathematical Education, 23*(4), 977–998.

Lee, S. H., Lee. B., J., & Son, H. C. (2007). Estimating the regression equations for predicting item difficulty of mathematics in the College Scholastic Ability Test. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 46*(4), 407–421.

Lee, Y. & Lee, E. H. (2010). The design and implementation to teach sampling distributions with the statistical inferences. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 12*(3), 273–299.

Leung, F. K. S. (2006). Mathematics education in East Asia and the West: Does culture matter? In F. K. S. Leung, K.-D. Graf, & F. J. Lopez-Real (Eds.), *Mathematics education in different cultural traditions – A comparative study of East Asia and the West: The 13th ICMI study* (pp. 21–46). Springer.

Nivens, R. A., & Otten, S. (2017). Assessing journal quality in mathematics education. *Journal for Research in Mathematics Education, 48*(4), 348–368.

Pang, J., Jin, S., Cho, S., Lee, Y., Kim, E., Kim, Y., Park, Y., Kim, K., Hwang, J., & Lee, H. (2019). Domestic research trends of mathematics education: An analysis of journals published from 1963 to 2019. *Journal of Educational Research in Mathematics, 29*(4), 709–739. https://doi.org/10.29275/jerm.2019.11.29.4.709

Ryu, H.-C., & Kim, H. J. (2004). A study on students' understanding of letters and algebraic expressions in solving algebraic word problems with Excel. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 6*(4), 411–433.

Seo, J. J. (2007). The effects of learning mathematics according to feedback method. *Journal of the Korean School Mathematics, 10*(1), 71–89.

Shin, S. J. & Cho, W., Y. (2018). An alternatives of the definition of definite integral in <MathematicsⅡ> textbook under 2015-revised curriculum. *Journal of Korea Society*

Educational Studies in Mathematics: School Mathematics, 20(4), 723-741.

Shin, D. (2020). A comparative study of domestic and international research trends of mathematics education through topic modeling. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 59*(1), 63–80.

Yim, J. H. (2005). Didactical analysis on triangle-determining conditions and triangle-congruence conditions. *Journal of Educational Research in Mathematics, 15*(2), 131–145.

Yim, J. H. (2007). Division of fractions in the contexts of the inverse of a Cartesian product. *Journal of Korea Society Educational Studies in Mathematics: School Mathematics, 9*(1), 13–28.

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics, 16*(S8), 1–10.

Zhao, W. X., Jiang, J., Weng, J., He, J., Ee-Peng, L., Yan, H., & Li, X. (2011). *Comparing Twitter and traditional media using topic models*. Paper presented at the 33rd European Conference on IR Research, ECIR 2011, Dublin, Ireland.