Research trends over 10 years (2010-2021) in infant and toddler rearing behavior by family caregivers in South Korea: text network and topic modeling

In-Hye Song¹, Kyung-Ah Kang²

¹Graduate Student, College of Nursing, Sahmyook University, Seoul; ²Professor, College of Nursing, Sahmyook University, Seoul, Korea

Purpose: This study analyzed research trends in infant and toddler rearing behavior among family caregivers over a 10-year period (2010-2021). Methods: Text network analysis and topic modeling were employed on data collected from relevant papers, following the extraction and refinement of semantic morphemes. A semantic-centered network was constructed by extracting words from 2,613 English-language abstracts. Data analysis was performed using NetMiner 4.5.0. Results: Frequency analysis, degree centrality, and eigenvector centrality all revealed the terms "scale," "program," and "education" among the top 10 keywords associated with infant and toddler rearing behaviors among family caregivers. The keywords extracted from the analysis were divided into two clusters through cohesion analysis. Additionally, they were classified into two topic groups using topic modeling: "program and evaluation" (64.37%) and "caregivers' role and competency in child development" (35.63%). Conclusion: The roles and competencies of family caregivers are essential for the development of infants and toddlers. Intervention programs and evaluations are necessary to improve rearing behaviors. Future research should determine the role of nurses in supporting family caregivers. Additionally, it should facilitate the development of nursing strategies and intervention programs to promote positive rearing practices.

Key words: Caregivers; Infant; Child rearing; Data mining

INTRODUCTION

Rearing, which begins with pregnancy and continues in earnest after childbirth, is the process of nurturing and bringing up a human being; it substantially influences the holistic growth and development of a child physically, emotionally, cognitively, and socially [1]. Specifically, early childhood, referring to the first 3 years after birth, is a critical period during which physical, emotional, cognitive, and social functional development progresses more rapidly than at any other life stage [2]. A child who does not receive appropriate care and love from family caregivers (FCs) during early childhood and does not form attachments through interactions with FCs may develop early life stress that can harm cognitive and emotional development, impeding brain development and physical growth [3]. Rearing behavior refers to the external and internal actions that parents universally exhibit while raising children and enabling their growth and development [4]. In this study, rearing behavior refers to the overall activities of FCs that contribute to the growth and development of infants and toddlers. Furthermore, we have included all relevant papers on children from birth to 6 years.

Although early childhood is a critical stage for children, during which the groundwork for growth and development is established, FCs also experience major transformations when they become directly involved in child-rearing after birth. FCs often face stress, physical fatigue, and lifestyle alterations due to their caregiving responsibilities. Consequently, during this vital period, FCs require specialized support to offer nurturing care and education to their children [2,5,6].

Nurses play a vital role in enhancing the rearing competency and confidence of caregivers, alleviating rearing stress, and promoting child development. This is possible due to their comprehensive practical knowledge and skills in human de-
child-rearing is seen as a shared responsibility between both parents, resulting from the economic participation of married women, further enhanced by the societal acceptance and obligations of married mothers [2,7]. In order to support the healthy growth and development of infants and toddlers by offering effective health education to caregivers, nurses must remain informed about recent research trends in rearing behaviors.

In the past, mothers were regarded as the primary caregivers. However, due to the rise in dual-income households resulting from the economic participation of married women, a sociocultural atmosphere has recently emerged in which child-rearing is seen as a shared responsibility between both parents. Various family members, including fathers and grandparents, also take part in childcare [8,9]. Consequently, it is essential to examine the research trends of prior studies that take into account the altered social and cultural characteristics of child-rearing practices among FCs.

Text network analysis entails gathering and examining papers associated with a specific research topic within a field of interest, then utilizing the identified knowledge structure to pinpoint distinct patterns and relationships among the dispersed data [10]. This approach aids in the analysis of extensive text data or granular topics within big data. It offers considerable benefits, as it supplies a quantitative foundation for extracting valuable information and has recently been employed in various disciplines, including nursing [10-12].

In this study, topic modeling was employed in conjunction with text network analysis. Topic modeling is a technique used to identify highly relevant keywords in documents, estimate the distribution of words and documents based on patterns composed of these keywords, model the main keywords and their classifications, and analyze the relationships and distributions of each topic [13-15]. By applying topic modeling alongside text network analysis, a clear understanding of research trends can be achieved based on the core topics of interest. Furthermore, since the meaning of a specific topic can be interpreted, the extracted topics can provide valuable information for representing documents [11,16].

In Korea, text network analysis has been employed to study various nursing research issues, including coronavirus disease 2019, clinical nursing practice, and nursing education [8,10,12,15,17]. However, research on the infant- and toddler rearing behavior of FCs published in the last decade has primarily been focused on survey studies examining specific variables, such as rearing stress and attitudes toward FCs. Although a systematic literature review [1] and meta-analysis studies [3,18,19] related to parenting have been published, they are limited in their understanding of the overall research trends in infant and toddler rearing behavior in Korea. Utilizing text network analysis and topic modeling methods to comprehensively understand research trends related to the infant and toddler rearing behaviors of FCs can provide valuable insights based on the most recent information, ultimately aiding in the care of infants and toddlers.

In this study, we explored the research trends in infant and toddler rearing behaviors among family childcare providers in South Korea over a 10-year period (2012-2021). To investigate these trends, we employed text network analysis and topic modeling techniques. The specific objectives were as follows: First, to identify the main keywords based on network centrality. Second, to explore subgroups within keyword networks using cohesion analysis. Third, to name the clusters identified by cohesion analysis based on topic modeling analysis of the major themes associated with early childhood rearing behavior of FCs.

### METHODS

#### Ethics statement

The Institutional Review Board (IRB) of the Sahmyook University (No. SYU 2022-09-002) reviewed this study. The committee decided that this study was an exempt research study.

#### 1. Study Design

In this study, we employed text network analysis and topic modeling to investigate prominent research trends associated with infant and toddler rearing behaviors. To achieve this, we constructed a co-occurrence keyword network derived from the abstracts of articles published in leading Korean databases between January 2012 and December 2021 (Figure 1-A).

#### 2. Data Selection and Study Procedure

1) **Keyword selection**

The last search was conducted on October 5, 2022, covering articles published between 2012 and 2021. Five major Korean databases were utilized for this purpose: KISS, RISS, ScienceON, KMBase, and KoreaMed.

In this study, the keywords used for the search were related to the "rearing target," "rearing providers," and "rearing behavior." For each category, we first included MeSH terms, then the synonyms of the MeSH terms followed by commonly used dictionary synonyms. In particular, we aimed to report the overall research trends related to child-rearing by including synonyms and keywords representing important concepts in child-rearing behavior. The final search keywords used for each category were as follows. Keywords related to the rearing target included "infant," "toddler," and "young children"; keywords related to the rearing providers included...

2) Abstract collection

Information on authors, publication year, title, and abstract was extracted using the reference management program Endnote 20.3 (Clarivate) from papers published between 2012 and 2021 in five representative databases. The aim of this study was to review research trends in South Korea across various disciplines. Consequently, we included all pertinent papers focusing on infants and young children aged from birth to 6 years.

In the initial search, 30,155 papers were retrieved, and duplicates were removed (18,080 articles; 59.96%) using the automatic duplicate removal function in Endnote 20.3 (Figure 1-B). After eliminating duplicate articles, 12,075 papers remained (40.04%). Based on the topic of interest, the titles of these 12,075 papers were reviewed, and 11,573 papers were selected, excluding 502 duplicate publications. Subsequently, the researchers examined the titles and abstracts, excluding documents unrelated or less related to the subject of this study. The exclusion criteria applied in this study were as follows: lack of relevance to the main research topic, foreign or non-resident Korean research participants, special situations involving disabilities and illnesses, topics related to cognition and learning, and papers for which the original text could not be verified. A total of 8,960 papers were excluded. Analysis of the 2,613 valid abstracts was conducted using NetMiner version 4.5.0 (Cyram Inc.) [20] and network analysis software.

3) Pre-processing

(1) Node filtering

Node filtering is the process of extracting semantic morphemes, which are the smallest units of meaning in language, from unstructured text and identifying words containing these morphemes. In network analysis, nouns are commonly employed as morphemes [21]. Consequently, in this study, only nouns were utilized as keywords for the morpheme analysis.

- Word refinement and selection

Using NetMiner's unstructured data function, 8,507 words were initially identified in the main node set. These words were subsequently sorted in descending order, and their frequencies were examined. We established a frequency cutoff of less than five as an appropriate exclusion criterion to eliminate words significantly unrelated to the research topic. Although no specific basis was used to determine the cutoff, we referred to the cutoff criteria employed in previous studies [22,23].

- Dictionary development

After selecting the words that aligned with the research objective, three dictionaries (thesaurus, exception, and defined words) were created for analysis [21].

In the thesaurus dictionary, representative words (e.g., "affection") were selected from the thesaurus (e.g., "affection," "love," and "warmth") and placed in the same row as the representative words in the first column of a Microsoft Excel (Microsoft Corp.) spreadsheet row. Representative words were selected based on the frequency of results from the node filtering process in NetMiner, prioritizing the highest frequency. However, if the word with the highest frequency did not adequately represent the thesaurus, words were chosen in order of frequency. Words not closely related to the research topic were added to the exception dictionary (e.g., "Amos," "bootstrapping"), and compound nouns consisting of two or more words were added to the defined dictionary (e.g., "Child rearing," "Emotion coaching"). The excluded and defined words were entered into text files, with each word in a separate row. The dictionary refinement process was repeated six times for the final analysis. Finally, keywords included in study titles ("Child," "Rearing," and "Family caregiver") were excluded.

- Keyword extraction using a dictionary

After completing the iterative dictionary refinement process, we applied the finalized dictionary and identified 338 keywords with a frequency of five or more in the main node set.

(2) Development of a keyword co-occurrence matrix (from a two-mode to a one-mode network)

The fundamental structure of a text network consists of nodes (keywords) and links (connections between nodes) [21]. In this study, we converted a word-document network (a two-mode network) into a word-word network (a one-mode network) for word co-occurrence analysis. Term frequency-inverse document frequency (TF-IDF) is a statistical metric that signifies the importance of a word within a specific document when a collection of documents is utilized for information retrieval and text mining [24]. Since the co-occurrence of words was employed as the unit of analysis, we generated a co-occurrence word matrix using the standardized TF-IDF value with a one-mode network, applying the relevant function of the NetMiner Program.
Figure 1. The process of this study on recent trends in research on infant and toddler rearing behavior by family caregivers. (A) Study flow. (B) Data selection process.
(3) Link filtering

The generated one-mode network comprised 25,744 links, which was too many to effectively visualize the network results. The researchers collectively determined the cutoff criteria for link reduction through a simulation. In the link reduction process, 62 links were chosen in the data analysis.

3. Data Analysis

The data were analyzed using NetMiner version 4.5.0.

1) Network visualization

Network visualization involves presenting nodes and links in a standardized format, allowing for an intuitive and visual display of the connection structure [21]. A one-mode network map was created, simulations were conducted, and the researchers ultimately decided to use a link frequency of $\geq 62$, as it best represented the research trends in infant and toddler rearing behaviors of FCs.

2) Network analysis

(1) Centrality

Centrality analysis identifies the central nodes (keywords) within a text network, which helps determine the most important keywords and the level of concentration in the network structure [21]. A keyword with a high centrality value is regarded as a core keyword. In this study, we employed degree and eigenvector centralities to examine word co-occurrence. Degree centrality quantifies the number of links in a node, with high values signifying core keywords. A high eigenvector centrality value suggests the existence of numerous nodes with elevated centrality in a particular area [21].

(2) Cohesion

Cohesion refers to the identification of subtopic groups within a network of co-occurring words and the subgroups that form within that network [21]. We extracted a giant component and conducted community analysis. Visualizing the results of network analysis was challenging due to the complexity of interpreting the connections between numerous nodes and links. To create a graph for the centrality and coherence analysis, we employed the Pathfinder network filtering algorithm, which filters only the essential links for each node.

(3) Topic modeling

Topic modeling is a text mining technique that interprets the distribution of words in a document based on the identified words [25]. In this study, we used a latent Dirichlet allocation (LDA) model to perform topic modeling and verify the cohesion analysis results. LDA is primarily employed to identify related topics within documents [21,25]. We established a criterion for extracting appropriate topics based on LDA parameters of $\alpha = .01$, $\beta = .01$, and 1,000 iterations [21]. Subsequently, we selected indicators for determining topics by referring to Griffiths, Steyvers, and Deveaud [21]. We then conducted six simulations using the aforementioned criteria to identify the optimal number of topics. Two researchers maximized the two indicators to determine the validity of each topic, and the final number of topics was established through consensus.

The optimal number of topics was determined to be two, as it corresponded to the most distinct distances and boundaries between topics. We named the topic groups by examining each topic’s top words and referring to the original texts of the documents with a high topic probability.

RESULTS

1. Characteristics of Included Studies

This study analyzed 2,613 papers published between 2012 and 2021. The average number of papers on infant- and toddler rearing behaviors by FCs was 261.3 (standard deviation=27.48). The number of published papers was 244, 226, 264, 312, 282, 275, 287, 251, 234, and 238 in 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013, and 2012, respectively (Supplement 1). The number of papers published on the rearing behavior of FCs was highest in 2018 and lowest in 2020. Of the total number of papers, 1,483 (56.8%) were published in academic journals, and 1,130 (43.2%) were theses. Of the 1,483 papers published in academic journals, only 46 (3.1%) were related to the nursing discipline.

2. Network Analysis

1) Word frequency, degree centrality, and eigenvector centrality of co-occurring keywords in abstracts

Table 1 presents the top 20 keywords based on frequency of appearance, degree centrality, and eigenvector centrality analysis in the abstracts of research papers on rearing behaviors toward young children by FCs published in South Korea over the past 10 years. The top 10 core keywords in all three criteria included "program," "scale," and "education." Further, the top keywords by frequency included "attitude" (3,411), "stress" (3,382), "competence" (2,989), "scale" (2,910), and "program" (2,118). The top five keywords by degree centrality and eigenvector centrality were "program," "scale," "kindergarten," "education," and "age." The degree of centrality index was 17.73%. The trends of each keyword analyzed based on de-
gree and eigenvector centrality displayed a similarity with 100% overlap.

2) Cohesion

For the cohesion analysis, a giant component was extracted from a one-mode network that represented the relationship between the set of keywords (main node to main node) and the document (abstract). This giant component consisted of a 67-keyword network with a link frequency of 62 or greater, based on one-mode TF-IDF code values. Two communities were identified with the optimal modularity value of 0.161. The keywords categorized by community can be found in Table 2, while the clusters are illustrated in Figure 2.

Cluster 1 comprised 36 keywords, including "scale," "program," and "education." Cluster 2 contained 31 keywords, such as "attitude," "stress," and "competence." Figure 2 displays the top 10 keywords for each cluster and the community map.

3) Topic modeling

In topic modeling, keyword order refers to the keywords that most accurately represent a given topic [21]. For Topic 1, the highest-probability keywords were "attitude" (8.1%) and "competence" (3.6%). For Topic 2, they were "stress" (7.8%) and "program" (7.0%) (Table 2). The documents associated with Topic 2 (n = 1,682) exhibited a classification resembling that of Cluster 1 in the cohesion analysis, which contained the largest number of documents. Topic 1, which included 931 documents, yielded results similar to those of Cluster 2. The topic modeling results are depicted in Figure 3-C.

4) Comparison of cohesion results with topic modeling and naming

The research team examined the similarity between the keywords found in the cohesion results and the topics, which included representative keywords from the topic modeling.
results, based on the keyword trends of each community analyzed through cohesion. Tables 1 and 2 demonstrate the agreement to name Topic 2 "program and evaluation," based on Cluster 1. Furthermore, Topic 1 was named as "caregivers' role and competency in child development," based on Cluster 2. Considering these results, most studies on Korean FCs' rearing behaviors toward infants and toddlers are related to Topic 2 (program and evaluation). Topic 1 (caregivers' role and competency in child development) has been less frequently researched.

### DISCUSSION

This study examined 2,613 papers published between 2012 and 2021, focusing on infant and toddler rearing behaviors by FCs in South Korea. The researchers identified core topics and knowledge structures of research trends over the past decade using text network analysis and topic modeling. The study revealed two topic groups based on the keywords within each cluster. These findings offer valuable quantitative and essential scientific evidence for future research on major trends and topic areas related to rearing behaviors for infants and toddlers by FCs in South Korea.

Analysis of the 2,613 papers included in this study indicated that the largest number of papers was published in 2018, with a total of 312. However, the number gradually declined thereafter, with 244 publications in 2021. This decline can be attributed to decreased interest due to the ongoing reduction in the number of infants and toddlers in Korea, which results from the dramatic decline in the birth rate [26]. Furthermore, of the 1,483 papers published in academic journals that were analyzed, only 46 (3.1%) were related to nursing disciplines, while the majority were focused on childcare, education, child development, counseling psychology, and sociology. Although nurses are considered the foremost experts and advocates for FCs raising infants and toddlers [1], insufficient research still exists in this area. This highlights the need for further investigation into various aspects of infant care from a nursing perspective.

Through text network analysis, we identified keyword frequency, degree centrality, and eigenvector centrality, with the top 10 core keywords in all three criteria including "program,"
"scale," and "education." This suggests that these three core keywords have stronger connections with other keywords and play particularly important roles within a cluster. Subsequently, a cohesion analysis was used to categorize keywords related to a research topic into two clusters. This analysis verified the similarities between keywords in each cluster [27]. However, analytical limitations were present because cohesion analysis cannot provide clear guidelines or mathematical evidence to designate each cluster as a core topic. In contrast, topic modeling can identify topics that represent each community based on the probabilistic distribution of words within documents [11,16]. Therefore, to overcome these limitations [12,28], we determined the two main topics of this study by comparing the results of cohesion analysis and topic modeling analysis.

The topic modeling results of this study identified two topical groups: "program and evaluation" and "caregivers' role and competency in child development." Comparing the two extracted topics with the two clusters categorized through cohesion analysis revealed that the extracted topics effectively represented the core keywords of each cluster. Consequently, the outcomes of both topic modeling and text network analysis were consistent, demonstrating that the topics representing recent research trends in infant and toddler rearing behaviors by FCs were scientifically and appropriately extracted.

The first topic, "program and evaluation," accounted for 64.37% (1,682 papers) of all articles within the subject of research. The papers reflecting this topic were primarily focused on the development and efficacy validation of various intervention programs, such as education and play for infant- and toddler rearing. Various programs for infant and toddler FCs have been developed, including a parent education program utilizing a social network system [29] and a grandparent education program for FCs [30]. These programs and interventions have demonstrated positive effects on various caregiver-related variables, such as parenting efficacy and parenting stress. This aligns with previous research [1] involving a systematic review of 11 studies on group-oriented parenting.
support programs targeting parents (FCs) of infants and toddlers; findings revealed that parenting support programs had the strongest impact on improving parenting competence. However, in the systematic literature review, the socio-psychological aspects of FCs and child development were found to have moderate effect sizes compared to the recent research trend focusing on interventions to enhance parenting competence in the rearing of infants and toddlers. Few studies have investigated the socio-psychological factors of caregivers that significantly influence parenting or the specific effects on the growth and development of young children, who are the target of parenting in this topic group. Given that the ultimate goal of parenting is the healthy growth and development of young children [1], more comprehensive research is needed in the future. This research should explore other factors that may affect rearing behaviors among FCs, the quantitative and qualitative effects on the growth and development of infants and toddlers as the target of rearing, and standardized measurement tools that can serve as criteria for measuring such effects.

The second topic, "caregivers' role and competency in child development," accounted for 35.63% (931 papers) of the articles examined. The studies within this topic primarily focused on the role and competence of FCs in the normal development of children, using outcome variables such as parenting attitudes, parenting stress, and parenting efficacy. These studies demonstrated that socio-psychological factors among caregivers, like parenting attitudes, positively influenced the development of infants and toddlers. A systematic review and

Figure 3. Visualization of the analysis of recent trends in research on infant and toddler rearing behavior by family caregivers. (A) Community map. (B) Word cloud based on topic modeling results. (C) Network map of topic groups.
In this study, the researchers examined the relationship between parenting attitudes and the resilience of infants and children during early childhood [31]. They discovered that resilience was effectively fostered among children when they experienced maternal affection and autonomous parenting attitudes. Additionally, the study was conducted in conjunction with two other investigations: a qualitative meta-synthesis of 11 papers exploring fathers' parenting involvement experiences during early childhood [19], and another qualitative meta-synthesis study of 10 papers investigating the experiences of grandparents in supporting their grandchildren's parenting [18]. These analyses aimed to identify the recent research trend of recognizing fathers or grandparents as crucial caregivers. This shift is attributed to contemporary societal changes that diverge from traditional, maternal-centered parenting. Consequently, future research should prioritize establishing and reinforcing the role and competency of FCs, pinpointing various factors that enhance caregiving competency, and examining the relationships among these factors.

In this study, we examined previous research trends concerning infant and toddler rearing behaviors by FCs, identifying two primary topic groups: "program and evaluation" and "caregivers' role and competency in child development." The aim of this study was to determine the knowledge structures and essential research topics that reflect current trends, offering practical research directions for support programs designed to assist FCs in the holistic development of infants and toddlers. To achieve this, the study utilized text network analysis and validated the findings through cohesion analysis with topic modeling, providing quantitative evidence for the validity of the resulting keyword clusters and topic groups. This study holds methodological significance and value.

The study involved the analysis of papers published over the past 10 years using academic databases in South Korea to identify research trends related to the rearing behavior of FCs. However, this research had limitations, as it did not consider research conducted outside of Korea. Therefore, collecting and comparing research papers published in international academic journals with domestic research trends will allow us to better understand the strengths and weaknesses of domestic research. This will enable us to gain richer insight into the state of nursing knowledge. Furthermore, a limitation existed in analyzing research trends over time, as the papers were not analyzed chronologically. Since parenting practices can vary by era, it is necessary to analyze them separately by time period in future studies. Moreover, studies on cognitive abilities or learning in early childhood were excluded from the analysis due to significant differences among individuals and environments. To minimize the impact of these limitations and enhance reliability and validity, this study compared and discussed the results via a systematic literature review and meta-analysis of infant- and toddler rearing behaviors of FCs. Future studies should address these issues and provide a more precise and helpful direction for researching infant and toddler rearing behaviors by FCs in the nursing discipline.

**CONCLUSION**

This study aimed to identify trends in research on the rearing behavior of infants and toddlers by FCs. The findings revealed that research has predominantly focused on "program and evaluation" and "caregivers' role and competency in child development" over the past decade. This study is significant as it provides information on research trends, which will assist in the development of future research topics, intervention programs, and support resources for the rearing behavior of infants and toddlers by FCs. Further research is required on the role and competence of FCs in enhancing their skills. Additionally, it is needed to analyze the various factors that affect caregiver competence and develop and evaluate nursing interventions and educational programs suitable for the current generation.

**ORCID and ResearcherID**

In-Hye Song  
https://orcid.org/0000-0002-0166-7371  
https://researcherid.com/rid/IUP-7885-2023

Kyung-Ah Kang  
https://orcid.org/0000-0002-3799-9554  
https://researcherid.com/rid/AGI-8092-2022

**Authors’ contribution**

Conceptualization: all authors; Data collection, Formal analysis: all authors; Writing-original draft: all authors; Writing-review and editing: all authors; Final approval of published version: all authors.

**Conflict of interest**

No existing or potential conflict of interest relevant to this article was reported.

**Funding**

None.

**Data availability**

Please contact the corresponding author for data availability.
Acknowledgements
None.

REFERENCES


23. Kang KA, Han SJ, Chun J, Kim HY. Research trends related to

https://doi.org/10.4094/chnr.2021.27.3.201

https://doi.org/10.16952/pns.2019.16.1.12


https://doi.org/10.22251/jlcci.2020.20.11.185

https://doi.org/10.4040/jkan.2019.49.5.538

https://doi.org/10.22590/ecce.2021.25.3.407

https://doi.org/10.22251/jlcci.2018.18.10.279

https://doi.org/10.15854/jes.2019.9.50.3.153
Supplement 1. Annual Numbers of Published Papers on Infant and Toddler Rearing Behavior by Family Caregivers.