**Abstract**

Over the past decade, our research team has conducted eight qualitative studies on women leaders in South Korea from the lens of work-life balance, career and leadership development, and cultural context, using semi-structured interviews. Given the emergence of computational analysis in HRD research, we felt a strong need to reanalyze those narratives to supplement the subjectivity issue of qualitative research, using topic modeling that is a computational technique to analyze a large volume of texts. The purpose of this study was to reanalyze women leaders’ narratives to examine the research topics identified through the topic modeling method. In the process, we identified key research topics on Korean women leaders’ narratives, compared those with the research themes of the previous qualitative studies, and provided implications for HRD research and practice for the future.

*Keywords*: women leaders, computational analysis, topic modeling, South Korea

**A Computational Analysis of Korean Women Leaders’ Narratives Using Topic Modeling**

Despite the remarkable economic growth of South Korea (Korea, hereafter), becoming one of the developed countries as in Organisation for Economic Co-operation and Development (OECD) in 1996, women’s status continues to lag behind. According to the World Economic Forum’s Global Gender Gap Report (2023), which shows a country’s state of global gender diversity and equality, Korea ranks 105th out of 146 countries surveyed. Korea’s gender gap index (0.68) falls below the global gender gap index (0.684), suggesting that Korea’s low rank largely comes from the lack of women leaders in government, politics, and business over the past decade.

Why then has Korea’s remarkable economic success failed to translate into women’s corresponding status? To answer this driving question, we have explored lived experiences of Korean women leaders in the workplace in their own voices and have collected a total of 200 narratives in diverse contexts: business, academic medicine, multinational corporations, entrepreneurship, higher education, and others over the past decade. In the eight qualitative studies conducted (Cho et al., 2015; Cho et al., 2016; Cho et al., 2017; Cho et al., 2019; Cho et al., 2020; Cho et al., 2021; Cho et al., 2022; Han et al., 2018), we found that women leaders experience difficulties in work-life balance, leadership development, and career development and success due to a gendered workplace where organizational constraints and traditional values coexist, such as informal networking (e.g., after-work drinking) is encouraged for team building purposes. We also found differences in women leaders’ narratives by position (team leaders and to management) and marital status (married and singles) concerning support, challenges, and leadership and by gender (women vs. men) in defining career success. The study findings also revealed that the gender divide is manifested in women’s roles in the home as opposed to men’s roles at work is unconsciously translated to the gendered workplace where men take up leadership roles and women occupy supporting roles.

However, due to the qualitative nature of the eight studies, we did not examine any associations and relationships between research constructs reviewed and thus we are not sufficiently confident if the study findings can be reproducible and generalized to other contexts beyond Korea. Although generalizability is not a goal of qualitative research, examining research topics that will be identified through topic modeling, a computational analysis method which compares research themes identified from the eight qualitative studies would give us the opportunity to revisit women leaders’ narratives from the perspective of quantitative analysis, which might be used as supplementing the subjective nature in qualitative research (Galdas, 2017, Peshkin, 1988).

Computational analysis ensures a high degree of reliability, trustworthiness, reproducibility, and efficiency in the data analysis process (Nelson, 2020). In this study, we used topic modeling as the method for computational analysis to reanalyze Korean women leaders’ narratives and compare key topics identified from this computational analysis with qualitative themes identified from eight qualitative studies. In analysis of 200 narratives using topic modeling to compare the themes and topics identified from the previous qualitative studies and computational analysis, this study will contribute to HRD research: first, this study will provide the possibility of exploring an uncharted area for a new methodological approach to qualitative research using computational analysis, supplementing the subjectivity of qualitative research. Second, this study will capture certain aspects of their experiences that might have been bypassed in the qualitative studies. Third, the challenges faced by women leaders across various organizational contexts, as identified in this study, will contribute to the development of policies and programs that support women in the leadership pipeline and gender equality in organizations, which has to do with the UN’s Sustainable Development Goal 5 (Gender Equality) for all women and girls (United Nations, 2015).

**Literature Review**

We reviewed the literature on computational analysis in qualitative research and topic modeling that are used in this study.

**Computational Analysis in Qualitative Research**

In qualitative research, a data analysis process considerably relies on the subjectivity of research, potentially undermining the reliability and consistency of the research findings (Nelson, 2020; O’Kane et al., 2021). Specifically, researchers’ subjective judgments play a substantial role in categorizing qualitative data and assigning labels to the emerging themes derived from these categories. The subjectivity could lead to reduced consistency and transferability of the analysis results, ultimately compromising the overall reliability of qualitative research (Miner et al., 2023; Nelson, 2020). To address this concern, qualitative researchers make efforts to establish rigor and ensure the trustworthiness of the data analysis process by employing various strategies such as intercoder reliability assessments and peer reviews. Nevertheless, these strategies might be inefficient as they demand a considerable amount of time and effort (Nelson, 2020).

Computational analysis of qualitative data offers reproducible and efficient analytical tools by integrating a quantitative approach (Baumer et al., 2016; Nelson, 2020). Within qualitative research, computational analysis can encompass various methods, though the terminology is often confusing and lacks consensus on precise definitions. The widely-used term, *computer-assisted* or *computer-aided analysis* refers to the use of qualitative data software like NVivo, QDA Miner, ATLAS. ti, and MAXQDA. This technique is known as CAQDAS (computer-aided qualitative data analysis software) (Maher et al., 2018; O’Kane et al., 2021). Through CAQDAS, researchers can perform analysis activities including text and code retrieval, coding co-occurrence, and visualization of connectedness between codes more efficiently, especially when analyzing large text data (O’Kane et al., 2021). CAQDAS enhances transparency and trustworthiness of analysis by allowing the recording of the analysis process specifically and checking inter-coder or intra-coder consistency. However, while utilizing a CAQDAS technique, researchers’ subjective interpretations are intertwined throughout the entire coding process - from selecting meaningful quotes to identifying the themes (O’Kane et al., 2021).

Recent developments in computational analysis in qualitative research suggests a shift towards *computer-generated* *or computer-led coding process* using machine learning, with topic modeling emerging as a commonly employed method (Baumer et al., 2016; Miyaoka et al., 2023). Nelson (2020) refers to a qualitative data analysis process where a computer detects and recognizes words and patterns in texts as "computational grounded theory." This computational analysis relies on the inherent text data itself rather than human perceptions, as the identification of words and patterns hinges on word frequency and co-occurrence within the text (Baumer et al., 2016). Consequently, computational analysis ensures a high degree of reliability, trustworthiness, reproducibility, and efficiency in the analysis process (Nelson, 2020). The computational analysis process usually follows three steps. For example, Hannigan et al. (2019) outline three steps of computational analysis: rendering corpora, rendering topics, and rendering theoretical artifacts. Similarly, Nelson (2020) proposes three steps of computational text analysis – pattern detection, hypothesis refinement, and pattern confirmation.

Despite the computational nature of the analysis, researchers’ interpretation still remains essential for identifying patterns and delving into meaningful topics from them (Miyaoka et al., 2023). Among Nelson's (2020) three steps of the computational grounded theory, pattern detection is conducted computationally, while the other two steps heavily rely on human perception and interpretation. Hence, the computational analysis, like other methods, cannot guarantee complete transparency and transferability (Miyaoka et al., 2023). Furthermore, the computational analysis, which relies on frequency and co-occurrence of words, has limitations such as disregarding the contextuality of texts, hampering the assessment of the significance and associations of detected words within a whole document (Miyaoka et al., 2023; Tonidandel et al., 2021). However, the recent emergence of Artificial Intelligence (AI) technologies, including advanced natural language process (NLP) models like ChatGPT and the GPT-4 model, is reshaping computational analysis techniques. AI automates the process of pattern identification, theme extraction, and result interpretation - typically led by human researchers’ subjective perception - through machine learning. This automation contributes to enhancing the transparency of the analysis process (Siiman et al., 2023).

Along with the current interest in computational analysis, there has been a noticeable increase in the use of computational qualitative data analysis in social sciences including sociology, psychology, management, and education. In particular, computation analysis has been often used in analyzing extensive textual data such as social media posts and documents. For example, Melville et al. (2019) examined the content of 78,783 posts including 3,221,794 words on a website to understand women’s encounters with sexism in their day-to-day lives using topic modeling. Tonidandel et al. (2021) also employed structural topic models to analyze 7288 managers’ leadership assessments and identified patterns of challenges they encounter. Some studies compare the findings of computational analysis with those of other computer-assisted techniques or human-generated techniques, and they have unveiled consistent results across different methods (e.g., Miner et al., 2023; Miyaoka et al., 2023). In the human resource development field, empirical studies employing computational analysis are surprisingly scarce despite its potential benefits and applicability in HRD research. For instance, Yoon and Chae (2022) conducted an analysis of words extracted from titles, abstracts, and keywords of articles published in the Human Resource Development Review between 2012 and 2021. They employed structural topic modeling (STM) to identify research topics and trends in HRD research. Lee and Chae (2019) also investigated employees’ perceptions of career success by analyzing responses from open-ended surveys using topic modeling. Therefore, there is a need for more empirical studies in the field of HRD to embrace the integration of computational analysis. This will not only bolster the reliability and trustworthiness of the qualitative research process but also offer new insights through diverse analytical approaches.

**Topic Modeling**

**Method**

The purpose of this study was to reanalyze Korean women leaders’ narratives using the topic modeling method, a computational analysis to compare with research themes identified in the previous eight qualitative studies. The research question that guided our inquiry was: how do we compare research themes identified in the previous qualitative studies with the key topics identified from the computational analysis using topic modeling?

To answer this question, we use structural topic modeling (STM) that charts a systematic pathway from the initial data intake to its eventual visualization, ensuring that researchers can seamlessly navigate through the algorithm-identified themes (Roberts et al., 2019). For analytical purposes, 200 women leaders’ narratives (transcripts) were converted to a structured dataset that is compatible with the R software. We chose sentences as the basic unit of analysis, as our research focus is to unearth overarching themes and discern patterns. From the 200 narratives (transcripts), we extracted 70,529 sentences, and for further analysis, we ended up with a corpus that has 69,102 sentences and 39,454 terms.

**Original Studies**

Table 1 (below) shows research purposes, theoretical backgrounds, themes identified, and the number of participating women leaders involved in the previous eight qualitative studies.

**Table 1**

*Original Studies on Women Leaders in Korea 2015-2022*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Research Purpose | Theoretical Background | Theme | Interview Participants |
| Cho et al. (2015) | To explore Korean women leaders’ challenges and opportunities that they face in the process of balancing work and family and developing leadership | Work-family balance, women’s leadership development, and cultural context | Commitment to work, struggles for a balance, defining success, gender identity, being a role model, and need for support | 18 women leaders in diverse sectors |
| Cho et al. (2016) | To investigate the lived experience of women leaders in Korean companies who manage their work and life and develop their leadership | Cultural context, work- life balance, and leadership development | Convergence (gendered workplace, work-life imbalance, leadership development, and definitions of success) and divergence (position and marital status) | 50 women leaders in business |
| Cho et al. (2017) | To compare Korean female executives’ definitions of career success with those of male executives | Career success, tokenism theory, and cultural context | Definitions of career success, success factors, and career development strategies | 15 male and 15 female executives in business |
| Han et al. (2018) | To investigate the lived experiences of Korean women physicians in academic medicine and social and cultural influences on the gendered process of their career journeys | Women’s career challenges, social and cultural contexts, and the Korean context | Gender discrimination in their early career, no work-life balance and identity crises, being ‘othered’ by social and cultural values, a lack of leadership aspirations and limited networking and mentoring, and pursuing individual excellence without organizational support | 21 female physicians in academic medicine |
| Cho et al. (2019) | To explore how multinational  corporations’ (MNCs) women leaders in Korea have become CEOs | Career success, Korean culture, and organizational culture | Becoming a CEO, Key success factors, MNC culture, career challenges, and career development strategies | 15 women CEOs in MNCs |
| Cho et al. (2020) | To conduct a case study of three Korean women entrepreneurs in the IT industry | Motivations, factors for business development, and Korean context | Case 1 (Business partnership with husband), Case 2 (Business principles and dreaming big), and Case 3 (Strong commitment with religious faith) | 3 women entrepreneurs |
| Cho et al. (2021) | To investigate women entrepreneurs’ motivations, challenges, and career  success | Motivational factors, challenges and opportunities, career success, and cultural context | Quantitative analysis and qualitative analysis (motivations, challenges and opportunities, and career success) | 23 women entrepreneurs |
| Cho et al. (2022) | To explore how highly educated women respond to career chance events in a Korean context | Career chance events, career interruptions, and Korean context | Before obtaining a doctoral degree, during and after the doctoral study, and responses to chance events | 50 highly educated women with doctoral degrees |

We conducted eight qualitative studies independently with different participating women leaders, using a basic qualitative study design (Merriam & Tisdell, 2016) to explore lived experiences of women leaders from the lens of work-life balance, career and leadership development, tokenism theory, career success and chance, and culture in diverse contexts including business, high education, academic medicine, entrepreneurship, multinational corporations (MNCs), and others.

**Data**

Table 1 shows a total of 195 interview participants involved in the eight qualitative studies. We added five narratives that were collected but not added, including: a women leader from the business sector and four from the non-business sectors (non-profit, public, high school, and local government) to ensure an equal number (15) of men and women in the business sector in the study on gender difference in the definitions of career success (Cho et al., 2017). In the original studies, all narratives were collected through semi-structured interviews using purposeful and snowball sampling techniques. All interviews were audio-recorded by participants’ permission and were transcribed, and the transcripts were sent to participants for member-checking to ensure validity.

**Analysis**

We will go through the following steps to analyze 200 narratives collected from the eight qualitative studies as follows:

* The dataset, comprising 200 interview transcripts, will be subjected to anonymization, followed by the creation of ancillary metadata encompassing interviewee identifiers (ranging from 2015-1 to 2022-50), industrial affiliations, designations, tenure, familial composition, and support structures. This anonymized metadata will be instrumental in discerning demographic-based patterns pertaining to female leadership.
* The topic modeling methodology will be deployed on a representative subset of the interview transcripts to ascertain its congruence with the data. In instances of informational lacunae within the metadata, deliberations will ensue on strategies to address such discrepancies before proceeding.
* Upon attaining consensus, the topic modeling methodology will be systematically applied to the entire dataset, utilizing computational methodologies delineated in the ensuing data analysis schema.
* As the computational outcomes materialize, an interdisciplinary consortium of researchers will embark on an evaluative discourse concerning the validity and reliability of the topic modeling outcomes. This analysis will entail juxtaposition against findings from the eight antecedent studies. The validation process will engage three researchers involved in previous qualitative investigations and a doctoral student who has joined the research team. The objective is to unearth latent research themes concerning women in leadership in Korea, harnessing the computational acuity of topic modeling to supplement the qualitative analysis. This amalgamation of computational and interpretative approaches is anticipated to yield less biased, robust, and reproducible research themes, thereby catalyzing future scholarly exploration and pragmatic applications.

The data analysis protocol is bifurcated into four cardinal stages: data assimilation and reformatting, data purification, preprocessing, and modeling. Initially, the 200 transcripts will undergo conversion to textual files and subsequently be structured into data frames within an Excel environment. This structure will demarcate interview questions, responses, and anonymized metadata. The second phase encompasses data cleansing, during which data validation is performed to ensure type fidelity and to ascertain the presence of any voids in the dataset. The preprocessing phase, which distinguishes text mining from traditional qualitative and quantitative analyses, involves transforming the text into an analyzable corpus. Techniques such as n-gram generation, stop-word filtering, and part-of-speech tagging are employed to construct tokenized word lists. Finally, structured topic modeling (STM) based on Latent Dirichlet Allocation (LDA) is applied to the processed data, yielding a matrix of topics and their associated probabilities. The model also affords criteria for the optimization of the number of topics. The research consortium will then undertake the responsibility of ascribing semantic labels to the topics, predicated upon keyword co-occurrence patterns.

**Findings**

**Discussion**

**Significance of the Study**

**Implications for HRD Research**

**Study Limitations**

**Conclusion**

**References**

Baumer, E. P., Mimno, D., Guha, S., Quan, E., & Gay, G. K. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology*, *68*(6), 1397-1410. <https://doi.org/10.1002/asi.23786>

Cho, Y., Park, J., Ju, B., Han, S., Moon, H., Park, S., Ju, A., & Park, E. (2016). Women leaders’ work-life imbalance in South Korean companies: A collaborative qualitative study. *Human Resource Development Quarterly, 27*(4), 461-487. <https://doi.org/10.1002/hrdq.21262>

Cho, Y., Park, J., Han, S., Ju, B., You, J., Ju, A., Park, C., & Park, H. Y. (2017). How do South Korean female executives’ definitions of career success differ from those of male executives? *European Journal of Training and Development*, *41*(6), 490-507. <https://doi.org/10.1108/EJTD-12-2016-0093>

Cho, Y., Park, J., Han, S. J., & Ho, Y. (2019). “A woman CEO? You’d better think twice!”: Exploring career challenges of women CEOs at multinational corporations in South Korea. *Career Development International*, *24*(1), 91-108. <https://doi.org/10.1108/CDI-03-2018-0078>

Cho, Y., Han, S. J., Park, J., & Kang, H. (2020). Business startups and development of South Korean women entrepreneurs in the IT industry [Special issue]. *Advances in Developing Human Resources*, *22*(2), 176-188. <https://doi.org/10.1177/1523422320907047>

Cho**,** Y., Kim, N., Lee, M., Lim, J. H., Han, H., & Park, H. Y. (2015). South Korean women leaders’ struggles for a work and family balance [Special issue]. *Human Resource Development International*, *18*(5), 521-537. <https://doi.org/10.1080/13678868.2015.1076562>

Cho, Y., Park, J., Han, S. J., Sung, M., & Park, C. (2021). Women entrepreneurs in South Korea: Motivations, challenges, and career success. *European Journal of Training and Development*, *45*(2/3), 97-119. <https://doi.org/10.1108/EJTD-03-2020-0039>

Cho, Y., You, J., Choi, Y., Ha, J., Kim, J., Kim, Y. H., Kang, S. H., Lee, S., Lee, R., & Kim, T. (2022). Career chance events of highly educated women with doctoral degrees in South Korea. *European Journal of Training and Development*. <https://doi.org/10.1108/EJTD-04-2022-0049>

Doldor, E., Wyatt, M., & Silvester, J. (2019). Statesmen or cheerleaders? Using topic modeling to examine gendered messages in narrative developmental feedback for leaders. *The Leadership Quarterly*, *30*(5), 101308. <https://doi.org/10.1016/j.leaqua.2019.101308>

Galdas, P. (2017). Revising bias in qualitative research: Reflections on its relationship with funding and impact. *International Journal of Qualitative Methods*, *16*, 1-2. <https://doi.org/10.1177/1609406917748992>

Han, H., Kim, Y., Kim, S., Cho, Y., & Chae, C. (2018). Looking into the labyrinth of gender inequality: Women physicians in academic medicine. *Medical Education*, *52*(10), 1083-1095. <https://doi.org/10.1111/medu.13682>

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., Kaplan, S., & Jennings, P. D. (2019). Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, *13*(2), 586-632. <https://doi.org/10.5465/annals.2017.0099>

Lee, J., & Chae, C. (2019). Differences and multi-dimensionality of the perception of career success among Korean employees: A topic modeling approach. *The Journal of the Korea Contents Association*, *19*(6), 58-71. (in Korean)

Maher, C., Hadfield, M., Hutchings, M., & de Eyto, A. (2018). Ensuring Rigor in Qualitative Data Analysis: A Design Research Approach to Coding Combining NVivo With Traditional Material Methods. *International Journal of Qualitative Methods, 17*(1). <https://doi.org/10.1177/1609406918786362>

Melville, S., Eccles, K., & Yasseri, T. (2019). Topic modeling of everyday sexism project entries. *Frontiers in Digital Humanities*, *5*, Article 28. <https://doi.org/10.3389/fdigh.2018.00028>

Merriam, S. B., & Tisdell, E. J. (2016). Six common qualitative research designs. *Qualitative research: A guide to design and implementation* (4th ed.) (pp. 22-42). Jossey-Bass.

Miner, A. S., Stewart, S. A., Halley, M. C., Nelson, L. K., & Linos, E. (2023). Formally comparing topic models and human-generated qualitative coding of physician mothers’ experiences of workplace discrimination. *Big Data & Society, 10*(1). <https://doi.org/10.1177/20539517221149106>

Miyaoka, A., Decker-Woodrow, L., Hartman, N., Booker, B., & Ottmar, E. (2023). Emergent Coding and Topic Modeling: A Comparison of Two Qualitative Analysis Methods on Teacher Focus Group Data. *International Journal of Qualitative Methods, 22*, 1-9. <https://doi.org/10.1177/16094069231165950>

Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, *49*(1), 3-42. <https://doi.org/10.1177/0049124117729703>

O’Kane, P., Smith, A., & Lerman, M. P. (2021). Building transparency and trustworthiness in inductive research through computer-aided qualitative data analysis software. *Organizational Research Methods*, *24*(1), 104-139. <https://doi.org/10.1177/1094428119865016>

Peshkin, A. (1988). In search of subjectivity–One’s own. *Educational Researcher*, *17*(7), 17-21. https://doi.org/10.2307/1174381

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*(2), 1-40. <https://doi.org/10.18637/jss.v091.i02>

Tonidandel, S., Summerville, K. M., Gentry, W. A., & Young, S. F. (2022). Using structural topic modeling to gain insight into challenges faced by leaders. *The Leadership Quarterly*, *33*(5), 101576. <https://doi.org/10.1016/j.leaqua.2021.101576>

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>

World Economic Forum. (2023). *The global gender gap report 2023*. <https://www3.weforum.org/docs/WEF_GGGR_2023.pdf>