Educational Technology International
2023, Vol. 24, No 2, 169-202

https://doi.org/10.23095/ETI.2023.24.2.169

Transforming Pre-service Teachers into Data-Driven Educators: A Developmental Research*

Huijin SEOK
Ewha Womans University
Korea

Jiwon LEE
Ewha Womans University
Korea

Eunjeong SONG
Dongguk University
Korea

Jeongmin LEE**
Ewha Womans University
Korea

This study aims to develop instructional design strategies included in educational programs that can effectively improve the educational data literacy of pre-service teachers. We used the design and development model proposed by Richey and Klein and investigated its internal and external validity. Internal validity assessment involved the input of five experts who evaluated the initial instructional strategies. We conducted an educational data literacy education program with 29 pre-service teachers from Korean colleges and graduate schools for external validity. The effectiveness of the program was verified by the Wilcoxon Rank Sum Test, which revealed a meaningful statistical difference between Wilcoxon Rank Sum Test post-scores after the four weeks of online classes. Therefore, this study developed instructional strategies followed by the steps of data-based decision-making: the final instructional strategies encompass 21 strategies, categorized for implementation before, during, and after classes, accompanied by 38 detailed guidelines. This approach bears notable significance as it encapsulates actionable and effective instructional strategies thoughtfully tailored to the unique circumstances and educational setting of the field, as well as the specific characteristics and requirements of the learners.

Keywords : Educational data literacy, Data-based decision-making, Pre-service teacher, Instructional strategy, Educational data

---

* This paper is a revised and updated version of the first author's master's thesis.

** Corresponding author: Department of Educational Technology, Ewha Womans University, jeongmin@ewha.ac.kr

Introduction

Following the enactment and introduction of the "Act on Vitalization of Data-Based Administration" in Korea in 2020, there has been a noticeable shift toward embracing a data-based society. This entails using data to make informed decisions across social, economic, and political domains (Pentland, 2013).

In tandem with this, changes are being initiated across all levels of the education system. The Ministry of Education revamped the Education Data Statistics System (EDS) to enable personalized education using big data (Park, 2020). Additionally, a local education office in Korea has developed an integrated platform enabling the collection and utilization of educational data (Lee, 2023). On the one hand, the research is increasingly supporting the analysis and prediction of data and expanding the potential applications of data (Kim & Kim, 2021; Oh & Oh, 2018; Shin et al., 2021). It is evident, however, that promoting data literacy development and researching practical uses of data are necessary to drive change at a school level (Choi et al., 2023; Kwon, 2019). Teachers, specifically, are expected to support learning through facilitating learner-centered and personalized education (Luckin et al., 2016), which demands the development of data literacy skills, including the ability to evaluate various forms of data as well as the accuracy of data-based information (Hong et al., 2017).

However, improving educational data literacy requires sufficient time for continuous education and experiential learning. This, consequently, requires a systematic approach to developing human resources by providing educational opportunities from the beginning of teacher preparation (Mandinach & Gummer, 2016; Reeves, 2017).

Accordingly, a variety of approaches have been taken to improve the data literacy of pre-service teachers. Piro et al. (2014) developed "Data Chat" to increase data literacy behaviors, and Reeves and Honig (2015) implemented a program for pre-

service teachers to utilize assessment data to identify learners' strengths, weaknesses, and levels of understanding. The major observation was that the pre-service teachers who participated in the program demonstrated significant improvements in data self-efficacy. Furthermore, a study by Kim and Kim (2022) aimed to enhance data literacy competencies by creating a program on big data analytics.

However, it should be noted that previous studies have certain limitations regarding their applicability and scalability. Pre-service educators should receive training on a methodology for "data-driven decision-making". Teaching all of the necessary analytical methods, as demonstrated in previous studies, is not feasible due to the increased amount of data available. Pre-service teachers should be encouraged to engage in collaborative, communicative, and critically reflective activities with peers within the data-based decision-making process (Henderson & Corry, 2021; Mandinach & Gummer, 2016). Furthermore, data literacy is relevant across various domains as it extends beyond specific subjects and disciplines. It is important to consider pre-service teachers with diverse backgrounds, experiences, majors, and interests (Kim & Kim, 2022).

Consequently, this study aims to develop and validate instructional design strategies to improve the educational data literacy of pre-service teachers. The objective is to equip pre-service teachers with the ability to engage in collaborative data-based instructional decision-making in the real-world contexts, encouraging them to use data to make informed instructional decisions. To this end, the study poses the following research questions:

1) What instructional strategies can be employed to improve the educational data literacy of pre-service teachers?

2) Do these instructional strategies designed to improve the educational data literacy of pre-service teachers demonstrate validity?

Theoretical Background

Educational Data Literacy

Educational data literacy is a multifaceted concept encompassing education, data, and literacy. While data literacy broadly refers to the ability to read and write data, its definition varies across disciplines (Kim & Suh, 2022; Song et al., 2021). For instance, in social science, Shields (2005) included the understanding of various tools that can be employed in data acquisition and manipulation. In media studies, Deahl (2014) defined data literacy as the competency to collect, understand, and visualize all types of data to support opinions. On the other hand, in computer science, Long and Magerko (2020) explained data literacy as the understanding of machine learning processes rooted in data generated by AI algorithms and the critical interpretation of learning datasets.

The definitions discussed above highlight the variability of the concept of data, depending on the user and their intended application. This study examines diverse interpretations of data literacy as presented in previous studies, intending to characterize data literacy in education, as illustrated in Table 1.

When examining the differences between data literacy and other concepts, information literacy is the most common concept. Nonetheless, due to the semantic differences between 'information' and 'data', there is a clear distinction between the two types of literacy (Bae, 2019). According to Liew (2007), information refers to content extracted from data while data is presented in the form of raw, unprocessed data. Data literacy which collects, analyzes, purifies, and converts various raw data into information is more suitable for a modern data-rich society.

This study further posits that statistical literacy and assessment literacy are the subcomponents of data literacy within the process of data-based decision-making, building on Henderson and Corry's (2021) conceptualization of data literacy. We argue that statistical and evaluative literacy acts as foundational components of data

Table 1
Concepts of related data literacy presented in previous studies

Concept	Definition
Information literacy	▫ The ability to understand the significance of responsible and ethical usage of information, to have a critical awareness and appropriate information behavior in order to obtain information according to their information needs through channels or media in society (Johnston & Webber, 2003).
Statistical literacy	▫ The ability encompasses fundamental and critical skills needed to use statistical information and to interpret results, as well as an understanding of statistical concepts, symbols, vocabulary, and measures of uncertainty such as probability (Ben-Zvi & Garfield, 2004).
Assessment literacy	▫ The ability to understand educational assessment and the skills to use that knowledge to measure the achievement of diverse students (Stiggins, 1991).
Data literacy	▫ The ability to identify, acquire, evaluate, clean, analyze, and visualize data is necessary, as well as the ability to handle data using information technology (Association of College and Research Libraries, 2013).
	▫ The ability to use data critically and ethically (Calzada Prado & Marzal, 2013).
	▫ The ability to make data-based decisions incorporating statistical literacy, assessment literacy, and pedagogical knowledge (Henderson & Corry, 2021).
Teacher data literacy	▫ The educators' aptitude to consistently, ethically, and appropriately acquire, comprehend, apply, and convey various kinds of data relevant to their roles and duties is significant (Data Quality Campaign, 2014).
Data literacy for teaching	▫ The ability to use data to inform instructional decisions involves the merging of data analysis with knowledge and expertise in pedagogy (Mandinach & Gummer, 2016).

literacy, enabling higher-level behaviors like data-based decision-making.

Likewise, data literacy in education has distinct characteristics which differentiate it from other relevant concepts. At first, data literacy in education includes the ability to objectively comprehend the educational context. It is crucial that the use of data

in education is grounded in content knowledge, educational objectives, learner characteristics, curriculum knowledge, instructional strategies, content pedagogy, and the educational environment (Mandinach & Gummer, 2016; Mandinach et al., 2015). This perspective also requires educators to be aware of ethical and critical matters in using data in education (Calzada Prado & Marzal, 2013; Koltay, 2017).

Second, in education, data literacy strongly emphasizes the importance of making decisions and communicating based on data (Data Quality Campaign, 2014; Henderson & Corry, 2021). As supported by research on data-based problem-solving in educational contexts (Gill et al., 2014; Mandinach & Gummer, 2013), the significance of an inquisitive mindset and collaboration in deriving actionable instructional knowledge and practices is noteworthy (Gummer & Mandinach, 2015).

Consequently, in the context of this study, data literacy refers to more than just the ability to understand and interpret the process of collecting and analyzing data. We defined educational data literacy as the ability to collect, analyze, understand, and interpret data and transform it into actionable pedagogical knowledge and behaviors as well as the ability to communicate collaboratively based on it. (Henderson & Corry, 2021; Mandinach & Gummer, 2016).

Data-based Decision-making in Educational Context

To improve educational data literacy, adopting an instructional approach that differs from traditional data literacy programs is necessary. This prioritizes statistical principles and data analysis methods. In line with the findings of Gummer and Mandinach (2015), Kippers et al. (2018), and Means et al. (2011), it should be structured as a comprehensive instructional module that encompasses the full spectrum of data-based decision-making processes in educational contexts. The data-based decision-making framework for teachers to enhance their skills is shown in Figure 1 (Schildkamp & Ehren, 2013).

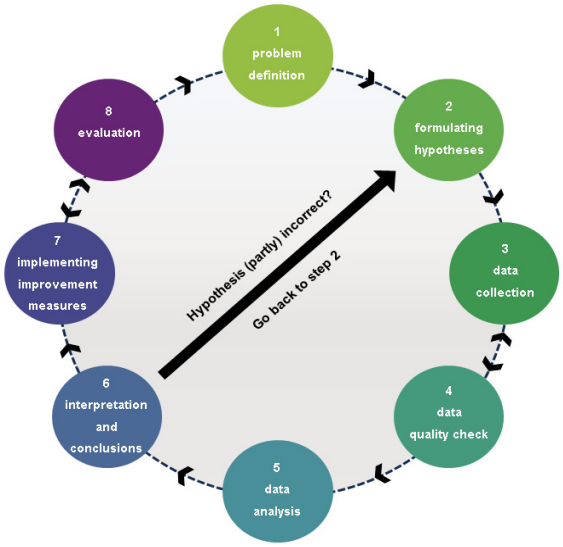


Figure 1. The Procedure of Data-Based Decision-Making (Schildkamp & Ehren, 2013, p. 56)

The process is in the formation of sequential steps, enabling teachers to collaborate effectively as a 'data team' when making data-based decisions. All stages of this training process are interconnected, following a logical progression.

In the first step 'problem definition', the data team selects an educational problem and establishes a goal. Following this, they hypothesize in the next step 'formulating hypotheses'.

Moving on to the third step 'data collection', information is gathered based on the identified problems and formulated hypotheses. This may comprise various types of data, such as evaluation data, test reports, and test outcomes, with data collection methods encompassing quantitative, qualitative, and blended approaches. Subsequently, in the 'data quality check' step, the data team is tasked with checking the reliability and validity of the collected data.

The data, ranging from descriptive to inferential statistics, are examined through different methods like regression analysis in the 'data analysis' step. This step also

entails data computation and summarization, followed by a comparative analysis. Then in the 'interpretation and conclusions' step, a determination is made regarding the validity of the hypothesis. If the hypothesis is true, the acquired information is utilized to determine the subsequent course of action. Otherwise, one sets up a new hypothesis or reverts to the second step to obtain additional information.

In the step of 'implementing improvement measures', the team deliberately considers the problem-solving goals and the necessary steps to achieve them. Each team member takes responsibility for determining the required resources and executing the actions. Additionally, they monitor these actions and make decisions on the data collection to assess the efficacy. In the final 'evaluation' step, the team assesses the effectiveness of the measures, the achievement of the objectives, the level of satisfaction with the outcome, and whether the problems were resolved. Depending on the result, they can proceed to a new problem, which means to initiate a new round from 'step 1'.

Drawing from the eight-step framework for data-based decision-making, Table 2 presents an overview of prior research that explored diverse approaches to enhancing data literacy.

The implications drawn from the analysis of instructional strategies in previous research are as follows. First, it is essential to implement strategies that enable pre-service teachers to gain practical, hands-on experience. The absence of a training environment results in limited practical experience in the phases of problem identification, data collection, and implementation of development measures. As established by previous research, participating in teacher internships can offer diverse opportunities for data exploration (Miller-Bains et al., 2022).

Second, the establishment of strategies for supporting self-directed learning is essential. The initial stage involves the introduction of data-based decision-making and its significance on data utilization so that students are conscious of the program's need and purpose (Miller-Bains et al., 2022; Song et al., 2021). In this process, instructors can support structured learning experiences by providing protocols and

Table 2
Analysis of data literacy instructional strategies for pre-service teachers

Steps	Piro et al. (2014)	Reeves (2017)	Kippers et al. (2018)	Song et al. (2021)	Miller-Bains et al. (2022)
0. Preparation and setting up a class environment	▫ Team considering the school levels and majors	▫ Step-by-step protocols and materials	▫ Cooperative team (4-6 teachers, 1-2 leaders, data experts)	▫ Election of analysis tools	
1. Introduction to a class		▫ Instructor as facilitator ▫ Guiding protocols and worksheets		▫ Guiding class progress and learning objectives ▫ Definition, importance, and data-based discussion rules	▫ Information on the usefulness of the evaluation data
2. Problem definition 3. Formulating hypotheses			▫ Encourage thinking about the purpose of using data ▫ Coaching	▫ Data exploration and problem discovery guide	▫ Providing prompts for problem definition
4. Data collection	▫ Provide local school evaluation data	▫ Provide real learner data	▫ Collect data to test hypotheses with the team	▫ Data definition and collection assistance	▫ Explore and collect in teaching internship
5. Data quality check			▫ Verify reliability and validity		
6. Data analysis	▫ A direct teaching of basic statistics ▫ Provide evaluation data analysis guidelines	▫ Support data analysis and visualization using appropriate tools	▫ Test hypothesis based on data	▫ Demonstrate data analysis process ▫ Provide scaffolding for error resolution	

Table 2

Analysis of data literacy instructional strategies for pre-service teachers *(continued)*

Steps	Piro et al. (2014)	Reeves (2017)	Kippers et al. (2018)	Song et al. (2021)	Miller-Bains et al. (2022)
7. Interpretation and conclusions	▫ Guide to interpret based on evidence and data	▫ Encourage discussion among students of the same major and grade	▫ Encourage inferences based on analysis and interpretation of data ▫ Test hypothesis		▫ Guide to evidence-based data interpretation and decision-making for instruction
8. Implementing improvement measures	▫ Evidence-based evaluation and instruction		▫ Develop improvement measures with team discussions	▫ Guide to implementing and monitoring	
9. Evaluation	▫ Presentation ▫ Provide guidelines for the presentation content		▫ Collect data for evaluation ▫ Assess satisfaction and Achievement	▫ Feedback from instructor and peer ▫ Evaluate achievement ▫ Self-evaluation and reflection	

worksheets in advance, as well as acting as facilitators to support learners (Kippers et al., 2018; Reeves, 2017).

Third, the decision to directly teach and demonstrate statistical and data analysis skills depends on the learner's prior knowledge. Previous studies have utilized strategies such as direct instruction, demonstration, and providing guidelines to help students better understand statistical theory and various analysis methods for data analysis (Piro et al., 2014; Song et al., 2021). Additionally, the selection of data analysis tools was influenced by the learner's proficiency level (Song et al., 2021). It is essential to note that the primary objective of data literacy education extends beyond acquiring theoretical knowledge and skills.

Fourth, encouraging collaboration among students is critical. Data-based decision-making in schools requires collaboration and communication between multiple stakeholders. Also, collaboration further fosters a deeper engagement in activities and allows pre-service teachers to seek assistance when facing challenges (Reeves, 2017). Previous studies have gathered a group of pre-service teachers based on their education levels and academic majors and engaged them in practice tasks or discussions as a part of a collaborative approach (Kippers et al., 2018; Piro et al., 2014).

Lastly, presenting and sharing the accomplished collaborative work is essential. Learners are able to engage in data communication practice through presentations. Song et al. (2021) implemented measures to facilitate the exchange of feedback, both from their peers and instructors. They further evaluated students' academic performance at the end of the class and provided opportunities for self-reflection. Therefore, it is necessary to support pre-service teachers in practicing data-based communication and reflecting on their data-based decision-making process.

Methods

Research Procedure

This study developed instructional strategies to improve the educational data literacy of pre-service teachers. The development process adhered to the design and development research method outlined by Rickey and Klein (2007), as followed by

Table 3
Procedure for developing instructional strategies

	Classification	Procedure	Details
A	The development of Instructional strategies	A Review of Previous Literature	▫ Review of previous literature related to data literacy, data literacy education for teachers and pre-service teachers, and data-based decision-making
		Initial Development of Instructional strategy	▫ Through a literature review, extract procedural steps for developing initial instructional strategies using a data-based decision-making model.
			▫ Analyze the instructional strategies of previous studies on educational data literacy instructional design based on data-based decision-making steps from 1 to 8.
			▫ Extract strategies for initial development based on previously proven effective strategies among the previous studies.
B	Validation of Instructional strategies	Internal Validation	▫ Expert review of the initial instructional strategies
			▫ Derivation of revised instructional strategies and supplements based on expert opinions
		External Validation	▫ Instructional design and application in class
			▫ Verification of the effectiveness of educational data literacy
			▫ Derivation of the final instructional strategies after collating validation results

previous studies Li et al. (2023) and Song et al. (2021) on instructional strategy development. The research method is categorized into two types: developing products and tools (Type 1) and model research (Type 2). This study follows type 2, with the detailed research procedure shown in Table 3.

Participants

Participants for internal validation. Verifying validity is critical in the instructional strategy development process (Richey & Klein, 2007). In this study, experts were consulted to obtain feedback on the development of instructional strategies on the basis of data-based decision-making stages and to improve the content. A panel of five educational professionals then evaluated the initial instructional strategies. Table 4 illustrates more details on the experts' profiles.

Table 4
Profile of experts participating in internal validation

Experts	Career	Position	Major	Highest degree
A	16 years	Professor	Educational Technology	Doctoral
B	10 years	Professor	Education Measurement and Evaluation	Doctoral
C	15 years	Teacher	Computer Education	Bachelor
D	5 years	Researcher	Educational Technology	Doctoral Student
E	6 years	Researcher	Educational Technology	Doctoral Student

Participants for external validation. External validation of instructional strategies is crucial to ensure the applicability of the strategies in real-world educational contexts (Rickey & Klein, 2007). To this end, a total of 29 pre-service teachers enrolled in either a college of education or a graduate school of education were recruited to participate in an educational data literacy class.

The invited class instructor had significant expertise in teaching educational data analysis to pre-service teachers. Her career is recognized in the field of education, so

she teaches at a university as a professor. We accordingly selected experts with an advanced understanding of educational data literacy and the capability to skillfully implement instructional strategies in the classroom.

Participants were selected using convenience sampling. We promoted our education program by posting on pre-service teachers' online communities to recruit participants. Twenty-nine participants were carefully selected among the received Google Forms responses as a result. The group consisted of two graduate students with majors in education methods and Korean language education and 27 college students with a diversity of majors from educational technology to early childhood education, social education, and mathematics education. It is worth noting that the majority of the participants had prior exposure to educational statistics or programming coursework.

Data Collection

In this study, we collected expert opinions during the internal validation process concerning instructional strategies through the expert validation checklist. Moreover, a pretest-posttest assessment took place in the external validation stage to verify significant differences in educational data literacy. The educational data literacy scale developed in the prior study was applied to both assessments' scales.

For the expert validation assessment, we established review criteria as a checklist referring to the standards outlined by Rha and Chung (2001), which have previously been adapted and refined by Kim (2019) and Song (2021). The evaluation criteria comprised 12 elements, mainly consisting of questions regarding levels of explanation, validity, appropriateness, universality, and comprehensibility of instructional strategies. We collected and analyzed the narrative responses. Sample evaluation questions include: "Is there an organized system and structure?" and "Does it employ methods that generate interest and motivation?". In expert selection, we adhered to the criteria outlined by Davis (1992); a minimum of two experts who have either authored academic papers in relevant professional fields or who can

credibly demonstrate their professional qualifications must be chosen. Five experts accordingly took a role as a panel, providing their assessments on program improvement and potential revisions to instructional strategies.

To assess the external validity of instructional strategies, we employed an educational data literacy measurement scale developed by Seok (2023). This assessment comprises a total of six components with 35 items, as depicted in Table 5. Responses were recorded on a 5-point Likert scale. The overall Cronbach's α for all items signified a high level of reliability of .98.

Table 5
Educational data literacy variables

Variables	Sub-factor	Reliability	The number of questions	Examples
Educational data literacy	Problem identification and definition	.68	3	I know what types of data can be used in an educational context.
	Data collection and management	.82	8	I can determine what data needs to be collected to solve educational problems.
	Data analysis and interpretation	.86	8	I can make a plan for data analysis to validate the hypotheses.
	Data visualization and representation	.89	5	I understand the characteristics of various data visualization way such as tables, charts, and graphs.
	Instructional decision making	.93	7	I can diagnose student needs using the data.
	Reflection	.90	4	I can use data ethically and legally.
All questions		.98	35	

Results

Initial Instructional Strategy Development

This study developed instructional strategies that adhere to the sequential stages of data-based decision-making to improve the educational data literacy of pre-service teachers. The following features characterize the strategies. First, it guides how lesson materials are prepared and outlines plans to support learners. Second, it guides instructors to support self-directed learning experiences as facilitators. Third, it supports the design of instruction that considers the characteristics and needs of learners by analyzing their prior knowledge and motivation as well as providing personalized support. Fourth, it promotes authentic experiences by providing actual educational data, practical tasks, and other methods. Finally, it facilitates collaboration and communication among the learners and encourages critical reflection through team-based tasks and feedback sharing. As a result, as shown in Table 6, this study presents initial instructional strategies encompassing 18 strategies and 31 detailed guidelines.

Internal Validity

This study carried out expert reviews to validate the initial instructional strategies. The reviews directed attention to three significant areas of modification: consideration of learner motivation, consideration of learner proficiency levels, and presentation of practical assignments.

One of the experts emphasized the value of helping pre-service teachers comprehend the significance of educational data literacy and data-based decision-making with regard to learner motivation. The strategies were modified to incorporate specific examples and illustrations of data-based decision-making in an educational context.

Table 6
Initial instructional strategies

Steps	Strategies	References
0. Prepare for class and set up environment	Select appropriate tools and environments for the class	Reeves (2017), Song et al. (2021)
	Form groups for practical assignments	Kippers et al. (2018), Piro et al. (2014)
	Create worksheets for practical assignments	Piro et al. (2014), Reeves (2017)
1. Introduce class	Introduce the basic concept of educational data literacy and data-based decision-making	Miller-Bains et al. (2022), Reeves (2017)
	Provide practical assignment	Miller-Bains et al. (2022), Piro et al. (2014), Reeves (2017), Song et al. (2021)
2. Define problem	Introduce educational data that can be collected to define a problem	Miller-Bains et al. (2022), Piro et al. (2014), Reeves (2017), Song et al. (2021)
	Explain the process of identifying and defining a problem	Kippers et al. (2018), Miller-Bains et al. (2022), Song et al. (2021)
3. Formulate hypotheses	Guide the process of formulating hypotheses	Kippers et al. (2018)
4. Collect data	Guide data collection	Miller-Bains et al. (2022)
5. Verify data quality	Provide instructions on how to evaluate the reliability of data	Kippers et al. (2018), Song et al. (2021)
6. Analyze data	Consider learners' prior knowledge and experience	Piro et al. (2014), Song et al. (2021)
	Introduce the data analysis tool	Reeves (2017)
	Introduce fundamental concepts of statistics and various data analysis methods	Kippers et al. (2018)

Table 6
Initial instructional strategies *(continued)*

Steps	Strategies	References
6. Analyze data	Provide practice opportunities	Reeves (2017), Kippers et al. (2018), Miller-Bains et al. (2022), Song et al. (2021)
7. Interpret and draw conclusions	Guide the way to interpret data logically based on evidence	Reeves (2017), Miller-Bains et al. (2022), Piro et al. (2014)
8. Implement improvement actions	Provide opportunities to implement modification	Kippers et al. (2018), Miller-Bains et al. (2022)
9. Evaluate	Share experiences about the process of data-based instructional decision-making	Piro et al. (2014), Song et al. (2021)
	Critical reflection on the process of data-based instructional decision-making	Song et al. (2021)

Second, the instruction is advised to be presented incrementally from simple to advanced knowledge, considering the learner's proficiency level, when teaching statistical concepts and data analysis techniques. Furthermore, recognizing different learning levels within individual learners prompted the inclusivity in providing additional support for supplementary and advanced learning. Learning materials for previewing and reviewing were integrated in response.

Lastly, in terms of practical assignments, concerns were raised about the possibility of free-rider tendencies in teams with four to five members. In response to these concerns, the size of the group contracted to two to three individuals. Furthermore, an expert recommended providing learners with instructional guidelines and specifying assessment criteria for peer feedback on each team's practical tasks in class.

External Validity

Instructional design and application in class. With the instructional strategies modified based on internal validation, external validation was implemented by a lecturer with expertise in educational data analysis. The process entailed a total of 8 classes held over four weeks, employing a block scheduling system via Zoom. The participants involved in this study consisted of 29 pre-service teachers.

Before delivering educational programs, we analyzed the learners and developed content and tasks such as motivation, prior knowledge, and experience based on their characteristics. However, a significant number of the pre-service teachers who participated in the study had no prior experience and were unfamiliar with statistical concepts. Adjustments were made to the structure of the session in general to lower the entry barrier and make the concept of 'data-based decision-making' more accessible. Data analysis was introduced as the first step. Then, we offered explicit teaching on evidence-based statistical principles and the utilization of data analysis software.

Next, practical tasks were developed to provide hands-on learning. Actual data from Kaggle was shared, accompanied by a protocol worksheet for making data-based decisions. There were teams of 2-3 pre-service teachers for these tasks based on their respective majors. They had the opportunity to work together on practical tasks (see Figure 2). In the last session, the teams presented their work and received feedback from peers and the instructor. Pre-service teachers recorded improvements to their work while listening to the presentations (see Figure 3). After each lesson, they were encouraged to reflect critically on data-based decision-making by completing reflection journals.

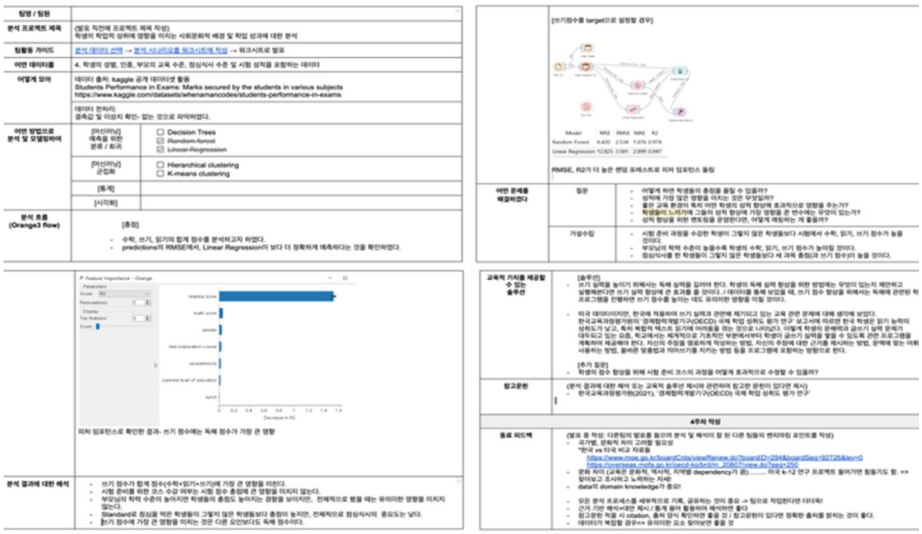


Figure 2. Examples of Worksheet

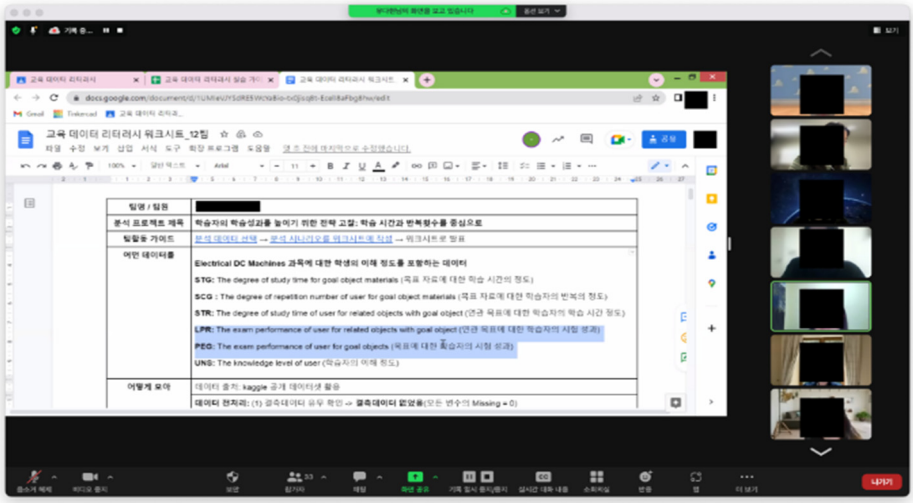


Figure 3. A sample screen of the educational data literacy program

Effectiveness of educational data literacy. In this study, we conducted pre-and post-tests to assess the educational data literacy of the students and the effectiveness of the developed instructional strategies. Given the relatively small sample size of fewer than 30 participants, we subjected each sub-variable within the educational data

literacy variable to the Shapiro-Wilk normality test. After excluding the two variables, namely 'Data collection and management' and 'Data visualization and representation', the significance probability for the remaining four variables was found to be less than .05. This result suggests that the data did not fit the normality, leading us to implement the Wilcoxon rank-sum test—a nonparametric test used for paired sample t-tests.

Wilcoxon rank sum test results showed meaningful statistical differences ($p < .001$) between pre and post-scores for all factors. Detailed results of the analysis are presented in Table 7.

Table 7
Results of the Wilcoxon rank-sum test for educational data Literacy

Variable	Factors	Pre-test	Post-test	Post-Pre				
				M	*SD*	*V*	*p*	
Educational data literacy	Problem identification and definition	3.43	4.45	1.02	0.65	0	***
	Data collection and management	3.14	4.10	0.96	0.52	0	***
	Data analysis and interpretation	3.39	4.25	0.85	0.61	14.5	***
	Data visualization and representation	3.73	4.29	0.56	0.6	26.5	***
	Instructional decision-making	3.16	4.21	1.06	0.84	18.5	***
	Reflection	3.74	4.40	0.66	0.69	4.5	***
	Total			0.87	0.51	2	***

***$p < .001$

Final Instructional Strategies

The final instructional strategies that underwent internal and external validation are outlined as follows. These strategies encompass ten stages, consisting of eight

steps aligned with the data-based decision-making process, along with two preparatory and introductory steps. Table 8 displays the final 21 strategies, including their timing for implementation (before, during, and after the class). In addition, the strategies are comprehensively elaborated upon with a total of 38 guidelines, offering specific insights into how to support learners.

The final instructional strategies can be summarized as follows. To begin with, the preparation for the class and setting up the learning environment entails analyzing students and selecting appropriate tools or platforms for class analysis. Moreover, acknowledging the diversity in proficiency levels among learners, it becomes imperative to provide pre-study and review materials in advance. Furthermore, arrangements for group organization and the creation of worksheets should be made to facilitate the practice tasks.

As a first formal step, '1. Introduce the class' comprises two general guidelines: introduce learners to the concepts of educational data literacy and data-based decision-making and their importance.

In the second step '2. Define the problem', it is imperative to introduce accessible data for exploration and to provide a detailed explanation on the steps to identify and define problems. Similarly, it is essential to guide students through the execution of steps '3. Formulating hypotheses' and '4. Collect data'.

Preceding the execution of '6. Analyze data' step, it is crucial to verify the quality of the data which brings the '5. Verify data quality' step between steps 4 and 6. In step 6, learners are tasked with selecting a topic for data analysis based on their prior knowledge and experience in their virtual educational context. During the class, the instructor introduces data analysis tools like 'Orange 3', which is used in the external validation of this research. Additionally, essential foundational statistical concepts and methodologies are presented in consideration of the learner's proficiency levels. Finally, it is of paramount importance to furnish students with practical experience in an educational context.

The instructor provides worksheets and protocols to support learners in logical

interpretation of data in step '7. Interpret and draw conclusions'. Then, in step '8. Implement improvement actions', it is crucial to develop actionable improvement measures even when the execution may not be immediately feasible. The final step '9. Evaluate' involves presentation and reflection activities. It is imperative to furnish learners with predefined criteria and guidelines for these activities. Before concluding the class, it is important to allocate time for learners to share their experiences with the data-based decision-making process and to receive feedback from the instructor and peers. This process enables learners to engage in critical self-reflection regarding their own performance.

Table 8
Final instructional strategies

Steps	Period	Strategies	Detailed Guidelines
0. Prepare for class and set up environment	***Before***	***0-1. Analyze the needs, skill level, and prior experience of learners***	▫ ***Perform preliminary interviews to gain insight into learners' motivations, prior knowledge, levels, and expectations.***
			▫ ***Determine the level of data-based decision-making practice tasks based on the learner's skill level and prior experience.***
	Before	0-2. Select appropriate tools and environments for the class	▫ Choose necessary online learning tools that consider both learners and instructors.
			▫ Choose interaction tools and strategies to facilitate instructor-student and student-student communication.
	Before and after	***0-3. Provide sufficient materials and time for preview and review***	▫ ***Provide materials in advance for preview and review.***
	Before	0-4. Form groups for practical assignment	▫ ***Assign 2 to 3 learners in one group based on their prior experiences, knowledge level, and major.***
	Before	0-5. Create worksheets for practical tasks	▫ Provide worksheets that help students follow data-based decision-making processes in practical tasks
			▫ Provide realistic and contextualized data for practical tasks

Table 8

Final instructional strategies *(continued 1)*

Steps	Period	Strategies	Detailed Guidelines
1. Introduce class	During	1-1. Introduce the basic concept of educational data literacy and data-based decision-making	▫ Introduce the concept and necessity of educational data literacy.
			▫ Introduce the process and cases of instructional s decision-making.
			▫ ***Provide examples of data-based decision-making processes in the educational context by in-service teachers.***
			▫ Notify teachers in training to understand defining specific problems or collecting actual data.
	During	1-2. Provide practical tasks	▫ Provide guidance on the structure and utilization of worksheets for the practical tasks.
			▫ Support learners to follow the recommended guidelines and complete the practical tasks.
2. Define problem	During	2-1. Introduce educational data that can be collected to define a problem	▫ Elucidate the concept and different types of educational data to facilitate learners' comprehension.
			▫ Conduct a discussion activity to let learners brainstorm alternative educational data utilization strategies beyond the provided examples.
	During	2-2. Explain the process of identifying and defining a problem	▫ Provide prompts to help learners understand the data and problem situations.
			▫ Guide learners through the principles of defining a problem and provide them with practice tasks
3. Formulate hypotheses	During	3-1. Guide the process of formulating hypotheses	▫ Practice identifying problems and formulating hypotheses from the given data.
4. Collect data	During	4-1. Provide guidance on data collection	▫ Guide on collecting various types of data (public, private, and self-collected) beyond the data utilized in the classroom.
5. Verify data quality	During	5-1. Provide instructions for how to evaluate the reliability of data	▫ Provide instructions for accurately confirming the format and source of the data.
			▫ Guide the process of ensuring validity and reliability before analyzing the data.
6. Analyze data	Before	6-1. Consider learners' prior knowledge and experiences	▫ Considering learners' prior knowledge and experiences, choose data analysis tools that learners can effectively utilize.

Table 8

Final instructional strategies *(continued 2)*

Steps	Period	Strategies	Detailed Guidelines
6. Analyze data	Before	6-1. Consider learners' prior knowledge and experiences	▫ Based on learners' prior knowledge and experiences, determine whether direct instruction on basic statistics and data analysis methods is necessary.
	During	6-2. Introduce the data analysis tool	▫ Explain the functions and features of the selected data analysis tool.
	During	6-3. Introduce fundamental statistics concepts and various data analysis methods	▫ Assess conceptual understanding and identify learner misconceptions by responding to questions and providing necessary corrections.
			▫ ***Present statistical concepts and analysis methods progressing from easy to more challenging.***
			▫ ***To motivate learners, introduce problems and examples of results that can be achieved using data analysis results.***
	After	6-4. Provide practice opportunities	▫ Provide actual educational data, related educational contexts, and problem situations.
			▫ During practice sessions, instructors provide advice and feedback as facilitators.
7. Interpret and draw conclusions	During	7-1. Guide the way to logically interpreting data based on evidence	▫ Provide worksheets or methods for logical interpretation and drawing conclusions.
			▫ Highlight the need for data modification based on objective empirical evidence.
8. Implement improvement actions	After	8-1. Provide opportunities to implement modification	▫ Provide an opportunity to experience data-based decision-making within a practical context and implement corrective actions.
9. Evaluate	Before	9-1. Share the evaluation criteria before the evaluation session	▫ Explain the evaluation criteria in advance and explain the main evaluation items.
	Before and during	9-2. Share experiences with the process of data-based instructional decision-making	▫ Set presentation times, Q&A procedures, and feedback processes for each group.
			▫ Provide guidelines for the presentation content, volume, and format of the materials
			▫ Instructors provide feedback on the presentation of practical assignments.
			▫ Share peer feedback
	After	9-3. Critical reflection on the process of data-based instructional decision-making	▫ Provide pre-formulated reflective questions and guide learners on submitting post-class reflection journals.

Discussion

The study aimed to develop instructional design strategies to improve pre-service teachers' educational data literacy skills. We accordingly completed internal validation through expert review and external validation through classroom implementation.

This study offers a novel approach to data literacy education in teacher education. Given the emphasis on educational data literacy in Korea, several studies have proposed ways to enhance the data literacy of pre-service teachers (Kim & Kim, 2022; Suh & Han, 2021). This study proposed specific examples of data-based instructional decisions that data-literate teachers are expected to make. Additionally, it offers step-by-step instructional strategies to facilitate those decisions. We anticipate improved statistical knowledge, data analysis, communication, and data-based decision-making skills through these strategies. The instructional strategies can also be applied in a wide range of contexts, including but not limited to instructional design, student guidance and counseling, and student teaching practice courses across various academic disciplines and education levels. This has significance as it allows the integration of data literacy skills into the objectives of traditional teacher education programs.

Secondly, instructional strategies are crucial as they contribute to improving teachers' effectiveness and expertise. In this study, 29 pre-service teachers were subjected to instructional strategies for four weeks to evaluate their effectiveness. Results showed a considerable improvement in educational data literacy competence from pre- to post-class, with a significance level of .001. In comparison to prior research on pre-service teachers, this study focused on improving cognitive competencies, namely 'data literacy', instead of affective competencies such as 'The Concepts of Assessment-III (COA-III)' and 'Teacher Self-Efficacy Scale (TSES)' (Miller-Bains et al., 2022). Among them, 'instructional decision-making' exhibited the greatest difference with a mean of 1.06. 'Problem identification and definition' displayed the second largest difference with 1.02. It can be inferred that implementing

these instructional strategies assisted pre-service teachers in comprehending the various data at their disposal and in identifying and hypothesizing the problems they aimed to address. In addition, their decision-making skills were enhanced to recognize students' needs, strengths, and weaknesses through data analysis. The participants are now capable of providing appropriate feedback on students' performance or development; they are able to provide explanations on the decision-making processes and persuade others through data. As a result, pre-service teachers are expected to identify and solve various problematic situations based on data they may encounter in an on-site school environment, ultimately contributing to students' positive academic, attitudinal, and social growth (Goe et al., 2008).

The limitations and recommendations of this study are as follows: First, more diverse methods should be tried to facilitate authentic learning. The study relied on data from the 'Kaggle' platform, provided in advance during the external validation phase. As a result, the scope of problem definition, hypothesis formulation, and data collection was constrained. Limited access to authentic situations and data is a challenge for teacher candidates. Future research is encouraged to explore opportunities to access data through collaborative efforts with nearby educational institutions, as demonstrated by Piro et al. (2014) and Reeves (2017). This collaborative approach can potentially increase the practicality of defining problems and implementing data-based instructional interventions.

Second, the external validation of this study included participants consisting of both undergraduate and graduate students from teacher colleges and graduate schools of education, primarily in metropolitan areas, with a remarkable 96.6% of the participants being female. In addition, 48.3% of the participants majored in educational technology, including English education, mathematics education, social studies education, and early childhood education. It should be noted that generalizing the findings of this study may be challenging due to the skewed demographic characteristics of the participants. Since data-informed decision-making has been identified as a necessary competency for teachers in all major subject areas and at all

levels of education, it is conceivable that this program could be applied to pre-service teachers from various disciplinary backgrounds.

Finally, this study has limitations in terms of research methods. It focused solely on internal validity in deriving strategies. Additional needs analysis and pilot testing would have increased the completeness of the derived instructional strategies. In addition, the study is limited by the lack of collection and analysis of various qualitative data, such as learner interviews and reflective journals, for external validity. Therefore, future studies should strive to include a broader range of research methods and greater participation of pre-service teachers in the process of identifying and evaluating instructional strategies.

References

Association of College and Research Libraries. Working Group on Intersections of Scholarly Communication and Information Literacy. (2013). *Intersections of scholarly communication and information literacy: Creating strategic collaborations for a changing academic environment.*

Bae, H. (2019). Educational implications of data literacy in social studies. *Theory and Research in Citizenship Education, 51*(1), 95-120. https://doi.org/10.35557/trce.51.1.201903.004

Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15). Springer. https://doi.org/10.1007/1-4020-2278-6_1

Calzada Prado, J., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri, 63*(2). https://doi.org/10.1515/libri-2013-0010

Choi, I., Si, K., & Park, S. B. (2023). A study on how to activate the use of educational data. *Journal of Curriculum and Evaluation, 26*(2), 217-236. https://doi.org/10.29221/jce.2023.26.2.217

Data Quality Campaign. (2014, June 30). Teacher data literacy: It's about time. *Data Quality Campaign*. https://dataqualitycampaign.org/resource/teacher-data-literacy-time/

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5*(4), 194-197. https://doi.org/10.1016/S0897-1897(05)80008-4

Deahl, E. (2014, May 9) Better the data you know: Developing youth data literacy in schools and informal learning environments. http://dx.doi.org/10.2139/ssrn.2445621

Gill, B., Borden, B. C., & Hallgren, K. (2014). *A conceptual framework for data-driven*

decision making. Mathematica Policy Research.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. National Comprehensive Center for Teacher Quality.

Gummer, E. S., & Mandinach, E. B. (2015). Building a conceptual framework for data literacy. *Teachers College Record, 117*(4), 1-22. https://doi.org/10.1177/016146811511700401

Henderson, J., & Corry, M. (2021). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning, 14*(2), 232–244. https://doi.org/10.1108/jrit-11-2019-0074

Hong, S., Nam, M., Lee, Y., Lee, D., & Park, S. (2017). *Exploring teacher roles and competencies in intelligence information society*. (ORM 2016-26-9). Korea Institute for Curriculum and Evaluation. https://nkis.re.kr/subject_view1.do?otpId=KICE00052483&otpSeq=0

Johnston, B., & Webber, S. (2003). Information literacy in higher education: A review and case study. *Studies in Higher Education, 28*(3), 335-352.

Kim, J., & Kim, J. (2021). Exploring the influential factors of peer relations in the first graders of middle and high school with random forest regression. *The Korean Journal of Educational Methodology Studies, 33*(3), 589-612.

Kim, K. (2019). *A developmental study of an instructional model for maker education using Single-Board Computer(SBC)* [Unpublished master's thesis]. The Graduate school of Seoul National University.

Kim, S., & Kim, T. (2022). The development and application of the big data analysis course for the improvement of the data literacy competency of teacher training college students. *Journal of The Korean Association of Information Education, 26*(2), 141–151. https://doi.org/10.14352/jkaie.2022.26.2.141

Kim, Y., & Suh, H. (2022). The direction of literacy education in the context of 'datafication': Reviewing alternative discussions on data Literacy. *Korean Association for Literacy, 13*(5), 379-411. https://doi.org/10.37736/kjlr.2022.10.13.5.13

Kippers, W. B., Poortman, C. L., Schildkamp, K., & Visscher, A. J. (2018). Data literacy: What do educators learn and struggle with during a data use intervention?. *Studies in Educational Evaluation, 56*, 21-31. https://doi.org/10.1016/j.stueduc.2017.11.001

Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science, 49*(1), 3-14.

Kwon, Y. (2019). Education data and analytics: A review of the state of the art. *The Korea Journal of BigData, 4*(1), 73-81. https://doi.org/10.36498/kbigdt.2019.4.1.73

Lee, M. (2023, June 6). [Exploring Future Education in Gyeongnam] (1) Big Data and AI Platform "iTalkTalk.". *The Kyung Nam daily news*. http://www.knnews.co.kr/news/articleView.php?idxno=1405817

Li, Z., Lim, E., & Lim, C. (2023). Development of teaching strategies for the effective use of automated writing evaluation(AWE) in English writing. *Korean Association for Educational Information and Media*, 29(2), 395-423. https://doi.org/10.15833/kafeiam.29.2.395

Liew, A. (2007). Understanding data, information, knowledge and their inter-relationships. *Journal of Knowledge Management Practice, 8*(2), 1-16.

Long, D., & Magerko, B. (2020), What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16.

Luckin, R., Holmes, W., Griffiths, M. & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher, 42*(1), 30-37. https://doi.org/10.3102/0013189x12459803

Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, *60,* 366-376. https://doi.org/10.1016/j.tate.2016.07.011

Mandinach, E. B., Parton, B. M., Gummer, E. S., & Anderson, R. (2015). Ethical and appropriate data use requires data literacy. *Phi Delta Kappan, 96*(5), 25-28.

Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. Office of Planning, Evaluation and Policy Development, US Department of Education.

Miller-Bains, K. L., Cohen, J., & Wong, V. C. (2022). Developing data literacy: Investigating the effects of a pre-service data use intervention. *Teaching and Teacher Education, 109*, 103569.

Oh, J. A., & Oh, H. (2018). The predictors of factors related to career decision making amongst adolescents who drop out of school, using decision-making tree analysis. *Studies on Korean Youth, 29*(1), 145-177.

Park, J. (2020, October 30). Education data system (EDS) is launched... strengthening education policies based on data. *Future Korea Weekly*. https://www.futurekorea.co.kr/news/articleView.html?idxno=142336

Pentland, A. S. (2013). The data-driven society. *Scientific American, 309*(4), 78-83.

Piro, J. S., Dunlap, K., & Shutt, T. (2014). A collaborative data chat: Teaching summative assessment data use in pre-service teacher education. *Cogent Education, 1*(1), 968409. https://doi.org/10.1080/2331186X.2014.968409

Reeves, T. D., & Honig, S. L. (2015). A classroom data literacy intervention for pre-service teachers. *Teaching and Teacher Education, 50*, 90-101. https://doi.org/10.1016/j.tate.2015.05.007

Reeves, T. D. (2017). Equipping preservice elementary teachers for data use in the classroom. *Action in Teacher Education, 39*(4), 361-380. https://doi.org/10.1080/01626620.2017.1336131

Rha, I. J., & Chung, H. M. (2001). Developing an action model for WBI design. *Journal of Educational Technology, 17*(2), 27-52. https://doi.org/10.17232/kset.17.2.27

Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Routledge.

Schildkamp, K., & Ehren, M. (2013). From "Intuition"- to "Data"-based decision making in Dutch secondary schools?. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 49-67). Springer.

Seok, H. (2023). *Educational data literacy program development and effectiveness verification for pre-service Teachers* [Unpublished master's thesis]. The Graduate school of Ewha Womans University.

Shields, M. (2005). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly, 28*(2), 6. https://doi.org/10.29173/iq790

Shin, C., Wi, E., & An, Y. (2021). The middle school achievement gap before and after covid-19 in urban and rural areas. *The Korea Association of Yeolin Education, 29*(5), 47–71. https://doi.org/10.18230/tjye.2021.29.5.47

Song, Y. (2021). *The data-driven debate (DDD) instructional model for improving data literacy* [Unpublished master's thesis]. The Graduate school of Seoul National University.

Song, Y., Song, S., Kim, Y., & Lim, C. (2021). A developmental study of an instructional model and strategies for data-driven debate (DDD) to improve data literacy. *Journal of Educational Technology, 37*(4), 943-982. https://doi.org/10.17232/kset.37.4.943

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan, 72*(7), 534-539.

Suh, H., & Han S. (2021). Changes in statistical knowledge and experience of data-driven decision-making of pre-service teachers who participated in data analysis projects. *Communications of Mathematical Education, 35*(2), 153-172.

Huijin SEOK

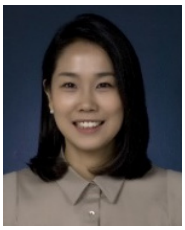
Master, Educational Technology, Ewha Womans University
Interests: Learning with technology, Skill-based learning, HRD
E-mail: joyseok@ewhain.net

Jiwon LEE

Master student, Educational Technology, Ewha Womans University
Interests: Technology-facilitated Learning Design, AI Learning
E-mail: leejiwon@ewhain.net

Eunjeong SONG

Adjunct Professor, AI Convergence Education, Dongguk University
Interests: Machine learning analysis of Educational data, Generative AI in education
E-mail: songej@songej.com



Jeongmin LEE

Professor, Educational Technology, Ewha Womans University
Interests: Technology-facilitated Learning Design, AI Learning
E-mail: jeongmin@ewha.ac.kr

Received: September 12, 2023 / Peer review completed: September 25, 2023 / Accepted: October 9, 2023