J. Korean Soc. Math. Ed. Ser. D. (2023) 26(3), 235–252
https://doi.org/10.7468/jksmed.2023.26.3.235
ISSN 1226-6191
Online ISSN 2287-9943

RESEARCH ARTICLE

Investigating Students' Profiles of Mathematical Modeling: A Latent Profile Analysis in PISA 2012

SeoJin Jeong[1], Jihyun Hwang[2], Jeong Su Ahn[3]

[1]Graduate Student, Department of Mathematics Gifted Education, Korea National University of Education

[2]Assistant Professor, Department of Mathematics Education, Korea National University of Education

[3]Graduate Student, Department of Mathematics Education, Korea National University of Education

Received: June 3, 2023 / Accepted: June 17, 2023 / Published online: September 28, 2023

Abstract

We investigated the classification of learner groups for students' mathematical modeling competency and analyzed the characteristics in each profile group for each country and variable using PISA 2012 data from six countries. With a perspective on measuring sub-competency, we applied the latent profile analysis method to student achievement for mathematical modeling variables - Formulate, Employ, Interpret. The findings showed the presence of 4-6 profile groups, with the variables exhibiting high and low achievement within each profile group varying by country, and a hierarchical structure was observed in the profile group distribution in all countries, interestingly, the Formulate variable showed the largest difference between high-achieving and low-achieving profile groups. These results have significant implications. Comparison by country, variable, and profile group can provide valuable insights into understanding the various characteristics of students' mathematical modeling competency. The Formulate variable could serve as the most suitable predictor of a student's profile group and the score range of other variables. We suggest further studies to gain more detailed insights into mathematical modeling competency with different cultural contexts.

Keywords: latent profile analysis, mathematical modeling, PISA 2012, International comparison

• Corresponding Author: SeoJin Jeong, email: seojin-jeong@naver.com

I. INTRODUCTION

Mathematical modeling is an essential component of mathematics education, as it enables students to connect mathematical concepts to real-world situations. It involves using mathematics to solve problems and make decisions across a wide range of fields, including science, engineering, economics, and social sciences (Blum & Niss, 1991). Moreover, by developing robust mathematical modeling skills, students can increase their opportunities for success in various fields, including engineering, finance, and data analysis. Various efforts have been made to evaluate students' mathematical modeling competency, adopting holistic views that emphasize an epistemological perspective and approaches that focus on the sub-competencies of mathematical modeling from a cognitive perspective (Borromeo Ferri, 2006; Kaiser & Sriraman, 2006).

Latent Profile Analysis (LPA) has been widely employed as an analytical tool in numerous mathematics education studies, both domestically and internationally. Notable examples include research examining the various types of motivation in mathematics learning (Hwang & Son, 2021), explorations into the characteristics of achievement related to attitudes towards mathematics (Chon & Kim, 2019), and studies leveraging PISA 2012 data to investigate the pedagogical environments for mathematics in South Korea and Singapore (Yi & Lee, 2017). However, there is no previous studies were identified that focused on profiling mathematical modeling competency specifically.

The significance of profiling students' mathematical modeling competency cannot be overstated. Such profiling enables educators to gain a more nuanced understanding of student performance in this essential aspect of mathematics, discerning patterns of strengths and weaknesses. This knowledge can subsequently guide the development of tailored instructional practices for individual students and groups. Moreover, competency profiling can inform curriculum development and assessment strategies by underscoring the specific skills and abilities requisite for successful mathematical modeling.

The purpose of this paper is to classify learner groups related to students' mathematical modeling competency and to examine the characteristics of each profile by country, variable, and profile group. We assumed that students' mathematical modeling competency can be measured through their response results, and that the student population comprises qualitatively diverse groups. To compare and analyze the characteristics of a broader population, we utilized PISA 2012 data – which focused on mathematics – from six countries. Latent profile analysis was applied to distinguish students with qualitatively diverse characteristics, followed by a comparative analysis of characteristics by country, variable, and profile group. The various characteristics of profile groups analyzed in this study are expected to contribute to a wide range of research aimed at understanding and developing students' mathematical modeling competency, and improving policy. The research questions are:

1. What are the profile models of students' mathematical modeling competency in each country?
2. What are the characteristics of the profiles for students' mathematical modeling competency, by country, variable and profile group?

II. LITERATURE REVIEW

Mathematical modeling is a process that involves using mathematical concepts, techniques, and structures to represent, analyze, and solve real-world problems. It is a powerful tool that allows researchers, educators, and professionals to describe complex systems, make predictions, and identify optimal solutions for practical challenges (Blum & Ferri, 2009). Mathematical modeling encompasses various mathematical branches, such as statistics, calculus, and algebra, and is widely used in diverse fields like economics, engineering, biology, and social sciences.

The process of mathematical modeling has been discussed since the 19th century in Europe and North America, and has been defined from a variety of perspectives. For example, mathematical modeling has been defined as a circular process of converting real-life problems into mathematical language, solving them within a symbolic system, and reexamining the results in light of real life (Verschaffel et al., 2002). It has also been defined as a process of using mathematics to represent real-life situations and contextual objects (Haines & Crouch, 2007), and as a powerful tool for researchers, educators, and professionals to describe complex systems, predict them, and provide optimal solutions to real-world problems (Blum & Ferri, 2009).

The mathematical modeling cycle is a systematic approach to problem-solving that typically involves problem formulation, model construction, analysis and computation, model validation, and interpretation of the results (Galbraith & Stillman, 2006). From a similar perspective, PISA 2012 integrated mathematical modeling into the definition of mathematical literacy by establishing a cyclic process that includes four different competencies: Employ (problem-solving), Formulate (model construction), Interpret (analyzing results), and Evaluate (validating the model). Figure 1 illustrates the mathematical modeling cycle in PISA 2012 as a part of mathematical literacy (OECD, 2013, p. 26), and Table 1 presents the activities included in the PISA 2012 assessment items related to the mathematical modeling variables presented by OECD (2013). The activities categorized by the mathematical modeling variable in Table 1 include those described by Galbraith & Stillman (2006), and the Formulate involves mathematical aspects in a real-world context, the Employ encompasses activities related to mathematical situations, and the Interpret relates back to the real-world context.

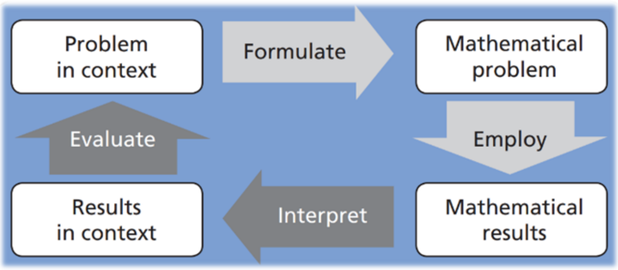

Figure 1. Mathematical modeling cycle in PISA 2012 (OECD, 2013, p. 26)

Both the mathematical modeling process and the mathematical literacy framework of PISA 2012 share a common starting point, which is a real-world context, and create models to solve mathematical problems (Blum & Leiß, 2007). This highlights the importance of students' hypothetical and creative reasoning. In the process of mathematical modeling, speculative thinking is inevitable, leading to cases that are valid explanations based on the understanding of phenomena and rules called mathematical concepts (Baek & Lee, 2018). This allows students to abstract mathematical models from real-world models and modify the mathematical models they create (Kim & Kim, 2004). In PISA 2012, this kind of speculative thinking is also represented by 'Formulate,' which is the ability to identify opportunities to apply and use mathematics, transform a presented situation into a form that can be processed mathematically, provide a mathematical structure and representation, and make assumptions to help solve the problem.

Table 1. Activities included in mathematical modeling variable items (OECD, 2013, pp. 28-31)

Variable	Activities included in the item
Formulate	• Identifying mathematical aspects in a real-world problem, including variables, constraints and assumptions behind any mathematical modelling • Recognizing mathematical structure and connections to mathematical concepts; understanding relationships between real-world and symbolic math language • Simplifying a situation or problem for mathematical analysis • Representing situations with variables, symbols, diagrams, and models, and using tech to portray a mathematical relationship in a context
Employ	• Manipulating numbers, graphs, statistical data, algebraic and geometric forms • Using math tools and tech for problem-solving, switching between representations • Making diagrams, graphs, constructions, generalizations and extracting math information • Applying mathematical facts, rules, and structures; devising and implementing strategies for math solutions; reflecting and justifying on mathematical arguments
Interpret	• Interpreting a mathematical result back into the real-world context • Explaining why a mathematical result is sensible or not within problem's context • Understanding real-world impacts on math procedure or models for contextual judgments • Evaluating the reasonableness of a mathematical solution in a real-world context; critiquing and identifying the limits of mathematical concepts, solutions, model used to solve a problem

Both PISA's mathematical literacy framework and the mathematical modeling cycle share similarities in their structure and process (OECD, 2013). Both approaches start from real-world contexts, generate models to solve mathematical problems, and interpret the calculated results within the context of the problem (Blum & Leiß, 2007). These similarities highlight the importance of connecting mathematical knowledge with real-world situations and the need for students to develop problem-solving and critical thinking skills (Jonassen, 2011).

However, there are differences between PISA's mathematical literacy framework and the mathematical modeling cycle, particularly in the competencies they emphasize (Maaß & Gurlitt, 2011). The mathematical modeling cycle focuses on the technical aspects of model construction and validation, while PISA's framework emphasizes the broader understanding and application of mathematics in real-life situations (Sriraman & Lesh, 2006). Despite these differences, both processes involve reinterpreting calculated results in the context of the problem, emphasizing the importance of connecting mathematical knowledge to real-world situations and the need for students to develop problem-solving and critical thinking skills (Jonassen, 2011).

Furthermore, mathematical modeling profiles examined through PISA 2012 refer to the unique ways individual approach and solve problems through mathematical modeling (Lesh & Lehrer, 2003). These profiles can be characterized by different levels of mastery and proficiency in the various competencies involved in the mathematical modeling cycle or PISA's mathematical literacy framework (Kaiser & Sriraman, 2006). By understanding and identifying these profiles, educators can tailor their instruction and assessment methods to better address the diverse needs and strengths of their students, fostering a more effective learning environment for mathematical modeling and its applications (Galbraith, 2012).

Perspectives on Measuring Mathematical Modeling Competency

Two main perspectives on measuring mathematical modeling competency are the holistic view and the sub-competency view. The holistic view considers mathematical modeling as a complex process that involves multiple competencies, including problem formulation, data analysis, mathematical modeling, and model validation. This view emphasizes the integration of these competencies and their effective coordination to solve real-world problems (Blum & Ferri, 2009; Lesh & Zawojewski, 2007). The evaluation criteria for this perspective may include the accuracy and validity of the model, the relevance of the model to the problem at hand, and the effectiveness of the model in addressing the problem (Galbraith & Stillman, 2006).

In contrast, the sub-competency view breaks down the mathematical modeling process into individual competencies or sub-skills, such as formulating the problem, selecting appropriate mathematical techniques, analyzing data, and validating the model. This perspective emphasizes the mastery and proficiency of each sub-skill and the ability to perform them independently, with the ultimate goal of combining them in a cohesive manner to solve a real-world problem (Gravemeijer & Doorman, 1999; English & Watters, 2005; Yoon et al., 2011).

To investigate students' profiles of mathematical modeling, it is necessary to adopt a sub-competency view of mathematical modeling. Students' profiles of mathematical modeling cannot represent their overall performance of mathematical modeling but they can reveal their strengths and weaknesses when they engage in the mathematical modeling process. However, it should be noted that students' engagement in mathematical modeling tasks depends on real contexts, which can make it difficult to compare students' mathematical modeling competency.

III. METHODS

Data Description

The research subjects are students aged 15 from six countries among the 65 countries that participated in PISA 2012. PISA 2012 results are the latest published data specifically focus on mathematics as a main domain, although PISA evaluates this subject in each of its assessments. The selected countries included South Korea, Singapore, Japan (Asia), Germany, Finland (Europe), and Australia (Oceania). The age of 15 corresponds to the time when compulsory education ends in most countries (OECD, 2013) and aligns with the third year of secondary school in South Korea. According to OECD (2013), the data was collected by sampling a minimum of 150 schools in each country, with a range of 4,500 to 10,000 students, providing a good sampling basis to analyzing the results according to various student characteristics. Table 2 represents the number of data collected for six countries, and there are no missing values.

Table 2. Data description

Country	Australia	Germany	Finland	Japan	Singapore	South Korea	Sum
No. of students	14,481	5,001	8,829	6,351	5,546	5,033	45,241

Variables

Three plausible values (PV1MAPF, PV1MAPE, PV1MAPI) representing student achievement of mathematical modeling variables were collected from the PISA 2012 database. The PV1MAPF score represents student achievement on the Formulate variable, the PV1MAPE score represents student achievement on the Employ variable, and the PV1MAPI score represents student achievement on the Interpret variable.

Data Analysis

In this study, we applied latent profile analysis using Mplus8.8 (Muthén & Muthén, 1998) to classify the types of learner groups for mathematical modeling variable and analyze their characteristics. Latent profile analysis (LPA) is a statistical approach for classifying and investigating qualitatively distinct profile groups based on the averages and patterns of learner response (observed continuous variables), which is one of the mixed modeling methods that estimate unobserved variables from observed data (Hwang & Ko, 2018). Latent profile analysis offers the advantage of a person-centered statistical approach. It allows for the exploration of qualitative differences among individuals and uncovers previously unobservable sub-profile groups (Iverson et al., 2018). The application of LPA involves an exploratory process where the researcher sequentially increases the number of profiles, comparing each to derive the most appropriate final model.

In order to optimal determine the number of profiles by country, 7 indices (AIC, BIC, SABIC, LMRT, BLRT, Entropy, the number of profile groups with a case ratio of less than 5%) were comprehensively considered, and the final model was selected. First, the fit indices such as AIC (Akaike Information Criterion; Akaike, 1987), BIC (Bayesian Information Criterion; Schwarz, 1978), and SABIC (Sample-size Adjusted BIC; Sclove, 1987) indicate that the smaller the value, the better the fit of the model. This means that as the number of profiles increases, the profiles are more differentiated and the model fit improves. Second, the statistical significance of LMRT (Lo-Mendell-Rubin adjusted likelihood ratio Test; Lo et al., 2001) and BLRT (Bootstrapped Likelihood Ratio Test; Arminger et al., 1999) is conducted. Specifically, if the p-value generated by LMRT and BLRT is less than the significance level 0.05 ($p < 0.05$), it suggests that the model is a better fit compared to a model with one less profile. Third, Entropy (Celeux & Soromenho, 1996), which represents the classification accuracy of the profiles, ranges between 0 and 1. A value closer to 1 signifies the higher the classification accuracy, reported cut-off of 0.8 or higher (Clark & Muthén, 2009). Fourth, it has been checked whether there are profile groups with a case ratio of less than 5% in each profiled group. Profiled groups with a case ratio of less than 5% are within the margin of error, therefore the fit of the profile may decrease (Marsh et al., 2009), and it's note that the small groups (less than 5% of cases) are considered spurious (Hipp & Bauer, 2006).

After determining the optimal number of profiles for each country using latent profile analysis, we analyzed the characteristics of the achievement distribution for each profile group and country, with respect to the mathematical modeling variables.

IV. RESULTS

Analysis of Student Achievement on Mathematical Modeling Variables

Table 3 shows the descriptive statistics of student achievement on mathematical modeling variables, by country. The range of means for the Formulate variable was 486.9 to 575.3, with a standard deviation range of 101.7 to 121.8. The range of means for the Employ variable was 489.4 to 569.4, with a standard deviation range of 84.8 to 98.1. The range of means for the Interpret variable was 503.6 to 550.1, with a standard deviation range of 92.4 to 105.7.

Through Table 3, the following characteristics of the distribution of student achievement by variable and country can be confirmed. First, the country with the highest average of student achievement for the three variables is Singapore, and the country with the lowest average is Australia. Second, the variable associated with either the highest or lowest average of student achievement differed by country. Specifically, The Interpret variable is associated with the highest mean of student achievement in Australia, Finland, and Germany, the Formulate variable is associated with the highest mean of student achievement in Japan, Singapore, and South Korea. In contrast The Formulate variable in Australia and Germany, the Employ variable in Finland and Japan, and the Interpret variable in Singapore and South Korea are associated with the lowest mean of student achievement. Third, the country with the highest standard deviation for each variable is South Korea (Formulate and Employ variables) and Germany (Interpret variable), whereas the country with the lowest standard deviation is Finland (Formulate and Employ variables) and Japan (Interpret variable). This means that South Korea and Germany students' achievement is more widely dispersed from the average compared to other countries. Fourth, the Formulate variable shows the highest standard deviation of student achievement in all countries, signifying that the students' achievement in the Formulate variable is most widely dispersed from the average. Conversely, the Employ variable shows the lowest standard deviation of student achievement in all countries, indicating that it is the least dispersed from the average.

Table 3. Descriptive statistics of mathematical modeling variables

Country	Variable	Mean	Variance	Standard Deviation	Minimum	Maximum	Median
Australia (n=14,481)	Formulate	486.9	12,427.1	111.5	10.5	901.4	483.9
	Employ	489.4	9,426.6	97.1	21.7	810.3	489.5
	Interpret	503.6	10,802.6	103.9	109.3	827.4	503.6
Finland (n=8,829)	Formulate	506.7	10,335.6	101.7	138.1	841.4	508.0
	Employ	505.4	7,185.4	84.8	152.9	764.2	506.6
	Interpret	516.3	8,870.2	94.2	103.8	851.6	520.3
Germany (n=5,001)	Formulate	510.5	11,220.3	105.9	158.7	850.1	512.5
	Employ	515.5	9,030.2	95.0	207.0	794.0	519.7
	Interpret	517.1	11,178.7	105.7	119.7	852.4	523.1
Japan (n=6,351)	Formulate	553.9	12,026.1	109.7	130.2	932.3	556.1
	Employ	530.5	8,195.4	90.5	182.7	832.1	533.9
	Interpret	531.1	8,529.7	92.4	146.9	863.3	534.0
Singapore (n=5,546)	Formulate	575.3	12,428.5	111.5	79.8	974.1	579.8
	Employ	569.4	9,050.1	95.1	195.1	873.9	575.8
	Interpret	550.1	9,548.0	97.7	118.7	918.3	553.3
South Korea (n=5,033)	Formulate	562.1	14,826.7	121.8	184.8	916.9	564.6
	Employ	553.1	9,626.5	98.1	198.1	864.5	554.9
	Interpret	540.1	10,925.9	104.5	152.6	845.0	544.0

Decision on the Profile Model by Country

In this study, the number of profiles by country was determined comprehensively based on 7 indices: AIC, BIC, SABIC, pLMRT, pBLRT, Entropy, and the number of profile groups with a ratio of less than 5%. Table 4 presents the model fit index used to determine the number of profiles by country. First, in all countries, AIC, BIC, and SABIC consistently decrease as the number of profiles increases. This indicates that the model fit improves and becomes more refined with an increasing number of profiles. Second, in all countries, the Entropy index exceeds 0.8. This value suggests that the classification accuracy of the profiles is high, with values closer to 1 indicating even better classification accuracy. Third, pLMRT, pBLRT, and the number of profile groups with a ratio of less than 5% exhibit differences in each country's profile model. Therefore, the optimal number of profiles was determined by considering these indices for each country's profile model as follows.

As the most appropriate model for each country, Singapore and Korea have 4 profiles, Finland, Germany, and Japan have 5 profiles, and Australia has 6 profiles. Australia was determined to have 6 profiles as the most appropriate model. The pLMRT value was found to be less than 0.05 for 6 profiles, while for 7 profiles, the pLMRT value exceeded 0.05. It indicates that 6 profiles are more a significant fit than 5, but there is no significant difference between 6 and 7 profiles. Additionally, with 7 profiles, there were 2 profile groups with less than 5% of the case ratio, whereas with 6 profiles, no such groups were present. Furthermore, the Entropy was high for the model with 6 profiles.

Finland was determined to have 5 profiles as the most appropriate model. Considering only the pLMRT values, since the pLMRT value was found to be less than 0.05 for 6 profiles, but the pLMRT value exceeded 0.05 for 7 profiles, so a model with 6 profiles may be more suitable. However, for the model with 6 or more profiles, there is at least one profile group with a case ratio of less than 5%: one profile group with a case ratio of less than 5% for the 6-profile model, two profile groups for the 7 and 8-profile models, and three profile groups for the 9 and 10-profile models. Since profile groups with a case ratio of less than 5% may decrease the fit of the profile model (Marsh et al., 2009) and are considered spurious (Hipp & Bauer, 2006), it has been determined that the most appropriate model is the one with 5 profiles. Furthermore, the entropy was high for the model with 5 profiles.

For Germany, a model with 5 profiles is the most appropriate. The pLMRT value was found to be less than 0.05 for 5 profiles, but for 6 profiles, the pLMR value exceeded 0.05. This indicates that 5 profiles are a more significant fit than 4 profiles, but there is no significant difference between 5 and 6 profiles. Additionally, with 5 profiles, there were no profile groups with less than 5% of the case ratio, and the Entropy was the highest.

For Japan, the most appropriate model is 5 profiles. The pLMRT value was found to be less than 0.05 for 5 profiles, but for 6 profiles, the pLMRT value exceeded 0.05. This suggests that having 5 profiles is a more significant fit than having 4 profiles, but there is no significant difference between having 5 and 6 profiles. Additionally, with 6 profiles, there was 1 profile group with a case ratio of less than 5%, whereas with 5 profiles, no such groups were present. Furthermore, the Entropy was high for the model with 5 profiles.

Table 4. Indicators of profile model fit by country

Country	Model	AIC	BIC	pLMRT	pBLRT	Entropy	No. of profile group <1%	No. of profile group <5%
Australia	1 profile	526,829.2	526,874.7				0	0
	2 profiles	503,618.6	503,694.4	<0.001	<0.001	0.845	0	0
	3 profiles	490,701.0	490,807.1	<0.001	<0.001	0.869	0	0
	4 profiles	482,133.3	482,269.8	<0.001	<0.001	0.885	0	0
	5 profiles	476,824.4	476,991.2	0.001	<0.001	0.886	0	0
	6 profiles	473,254.9	473,452.0	0.030	<0.001	0.884	0	0
	7 profiles	470,807.5	471,034.9	0.052	<0.001	0.878	0	2
Finland	1 profile	315,444.8	315,487.3				0	0
	2 profiles	301,657.1	301,727.9	<0.001	<0.001	0.836	0	0
	3 profiles	292,717.0	292,816.2	<0.001	<0.001	0.884	0	0
	4 profiles	287,238.7	287,366.2	<0.001	<0.001	0.892	0	0
	5 profiles	283,694.9	283,850.8	0.038	<0.001	0.891	0	0
	6 profiles	281,035.8	281,220.0	0.002	<0.001	0.893	0	1
	7 profiles	279532.5	279745.0	0.293	<0.001	0.882	0	2
Germany	1 profile	181,391.3	181,430.4				0	0
	2 profiles	173,157.8	173,223.0	<0.001	<0.001	0.849	0	0
	3 profiles	168,359.9	168,451.1	<0.001	<0.001	0.880	0	0
	4 profiles	165,683.2	165,800.5	0.012	<0.001	0.884	0	0
	5 profiles	163,807.2	163,950.6	<0.001	<0.001	0.886	0	0
	6 profiles	162,806.7	162,976.1	0.325	<0.001	0.874	0	0
Japan	1 profile	228,461.5	228,502.0				0	0
	2 profiles	218,568.5	218,636.0	<0.001	<0.001	0.841	0	0
	3 profiles	212,969.6	213,064.2	<0.001	<0.001	0.874	0	0
	4 profiles	209,586.2	209,707.8	<0.001	<0.001	0.879	0	0
	5 profiles	207,570.1	207,718.7	0.005	<0.001	0.875	0	0
	6 profiles	206,200.9	206,376.5	0.061	<0.001	0.876	0	1
Singapore	1 profile	202,931.5	202,971.2				0	0
	2 profiles	193,707.5	193,773.8	<0.001	<0.001	0.856	0	0
	3 profiles	188,771.3	188,864.0	<0.001	<0.001	0.871	0	0
	4 profiles	185,756.1	185,875.3	<0.001	<0.001	0.885	0	0
	5 profiles	183,965.7	184,111.3	0.113	<0.001	0.878	0	0
South Korea	1 profile	182,284.2	182,323.3				0	0
	2 profiles	174,377.1	174,442.4	<0.001	<0.001	0.841	0	0
	3 profiles	169,851.3	169,942.6	<0.001	<0.001	0.871	0	0
	4 profiles	167,013.0	167,130.4	0.001	<0.001	0.882	0	0
	5 profiles	165,251.9	165,395.4	0.066	<0.001	0.882	0	0

* The highlighted cell indicates the most suitable profile model for each country.

Table 5. Distribution of students of each country's profile group and the average student achievement for each variable

Country	Profile group	No. of students	Proportion (%)	Achievement average		
				Formulate	Employ	Interpret
Australia (n=14,481)	A	891	6.2	704.4	670.8	693.9
	B	2,533	17.5	604.6	592.7	611.6
	C	3,804	26.3	522.7	523.5	539.2
	D	**3,966**	**27.4**	444.1	454.4	467.7
	E	2,551	17.6	368.4	383.2	390.9
	F	736	5.1	273.6	293.5	299.3
Finland (n=8,829)	A	982	11.1	673.5	642.8	664.0
	B	2,259	25.6	579.0	566.9	583.8
	C	**2,938**	**33.3**	499.2	500.0	511.9
	D	2,009	22.8	414.4	428.6	433.1
	E	641	7.3	318.7	342.6	331.2
Germany (n=5,001)	A	614	12.3	676.8	660.5	676.4
	B	1,370	27.4	580.6	580.9	587.9
	C	**1,522**	**30.4**	499.1	508.6	510.3
	D	1,080	21.6	413.9	428.3	421.5
	E	415	8.3	326.3	338.1	320.9
Japan (n=6,351)	A	672	10.6	727.9	672.6	676.2
	B	1,693	26.7	634.1	597.6	597.6
	C	**2,017**	**31.8**	546.8	526.2	527.0
	D	1,452	22.9	459.6	452.3	451.8
	E	517	8.1	353.8	359.2	359.5
Singapore (n=5,546)	A	1,103	19.9	736.1	697.1	686.2
	B	**2,108**	**38.0**	616.1	604.0	583.3
	C	1,625	29.3	500.2	510.6	487.8
	D	710	12.8	379.0	405.1	384.0
South Korea (n=5,033)	A	1,007	20.0	711.5	679.1	665.1
	B	**1,883**	**37.4**	597.8	584.1	572.0
	C	562	11.2	491.6	494.5	482.1
	D	1,581	31.4	374.3	389.4	373.2

* Bold font indicates the profile group with the largest number of students in each country.

Singapore and South Korea were determined to have 4 profiles as the most appropriate model. The pLMRT value was found to be less than 0.05 for 4 profiles, while for 5 profiles, the pLMRT value exceeded 0.05. It indicates that 4 profiles are a more significant fit than 3 profiles, but there are no significant differences between 4 and 5 profiles. Furthermore, with 4 profiles, there were no profile groups with less than 5% of the case ratio, and the Entropy was the highest.

Analysis of Results for the Profile Model by Country

Utilizing the established profile model, Table 5 shows the average student achievement for each mathematical modeling variable, as well as the number and ratios of students in each profile group for each country. For example, in Australia, students in profile group A make up 6.2% of the total, and their average achievement is 704.4 points for the Formulate variable, 670.8 points for the Employ variable, and 693.9 points for the Interpret variable. In the same way, the data can be interpreted for other countries and profile groups.

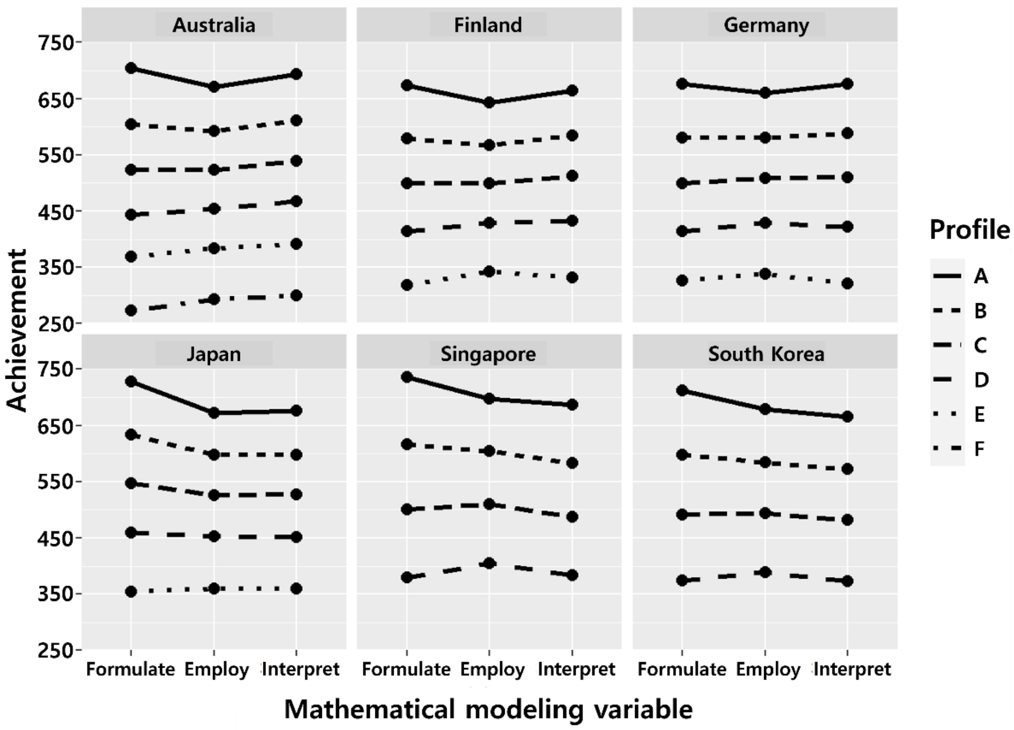

Figure 2. Distribution of profiles in each country for each mathematical modeling variable

Based on the average student achievement data presented in Table 5, Figure 2 represents the distribution of profile groups in each country for each mathematical modeling variable. Figure 3 shows the distribution of countries in each profile group for each mathematical modeling variable. Through the analysis of country-specific profiles for mathematical modeling variables, several characteristics can be identified in five aspects:

similarities and differences in the distribution of profile groups for each country, student achievement by profile groups and variables, and achievement gaps.

First, in all countries, the profile groups within each country have a hierarchical structure, where the average student achievement for mathematical modeling variables tends to decrease as the hierarchy level decreases. In other words, as the country-level profile groups move from A to F, the average student achievement for the Formulate, Employ, and Interpret variables also decrease. Additionally, the lower average achievement for the Formulate variable was also indicated by lower average achievement for the Employ and Interpret variables in the profile group. This can be confirmed by the fact that the cross points between profile groups within each country do not appear in the line graph Figure 2.

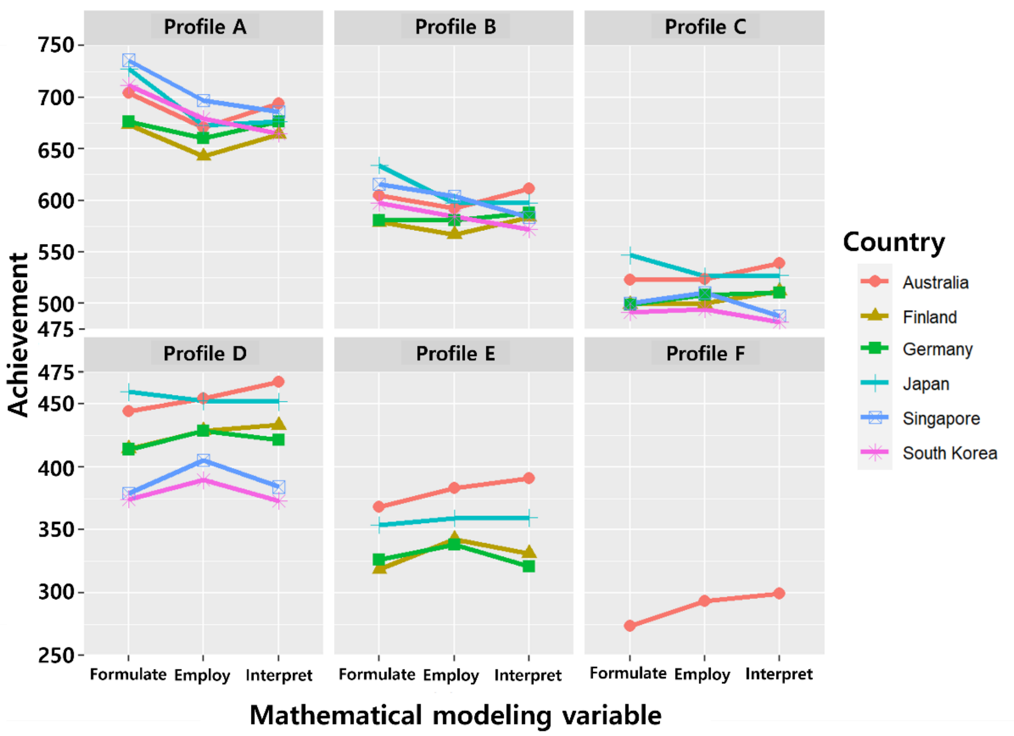

Figure 3. Distribution of countries in each profile for each mathematical modeling variable

Second, the hierarchical structure of profile groups within a country is roughly established for profile groups in six countries (see Figure 3). This hierarchy was particularly evident in the high-achievement profile groups, while it was less pronounced in the low-achievement profile groups. In other words, there was no overlap in the distribution of high-achievement profile groups A, B, and C, while some overlap did occur in the low-achievement profile groups D and E among the countries. This is due to the number of profile groups in Australia being one to two more than other countries, whereas

Singapore and South Korea have the fewest profile groups at four, resulting the average achievement of Australia's profile group E was higher than that of Singapore and South Korea's profile group D. However, it should be noted that we used names from profile group A to F to differentiate groups by country, and using the same name in different countries does not imply homogeneity of the profile group.

Third, among the mathematical modeling variables, the Formulate variable demonstrated the largest difference between the high-achievement and the low-achievement profile groups (see Figure 2, 3). Specifically, in profile group A, which has high achievement, the average achievement for the Formulate variable was higher compared to other variables. In contrast, in the low-achievement profile groups D, E, and F, the average achievement for the Formulate variable was generally lower compared to other variables. The line graph for the high-achievement profile group A has its highest point on the left end, while the line graphs for the low-achievement profile groups D, E, and F generally have their lowest points on the left end. This is related to the Formulate variable having the highest standard deviation across all countries (see Table 3).

Fourth, it's notable that in Singapore and South Korea, the proportion of high achieving profile groups A and B is larger compared to other countries, while the proportion of low achieving profile groups D and E is larger in other countries. In particular, Singapore stands out as the country with the highest average achievement for two out of three variables in the highest-achieving profile group A (see Figure 3). This is not unrelated to the result that Singapore and South Korea occupied the top ranks in PISA 2012.

Fifth, the mathematical modeling variables associated with high or low average achievements manifested differently in each profile group for each country. Specifically, while Australia consistently exhibited high average achievement for the Interpret variable across all profile groups, the countries demonstrating low average achievement for the Interpret variable varied. Furthermore, countries with high or low average achievements for the Formulate and Employ variables varied across each profile group. Indeed, identifying the average achievement of each variable in each profile group is valuable for pinpointing potential areas of improvement in each country's education system.

V. DISCUSSION AND CONCLUSION

In this study, we classified types of learner groups based on student achievement in mathematical modeling competency and analyzed their characteristics in each variable for each country. We applied latent profile analysis method to mathematical modeling variables (Formulate, Employ, and Interpret), using student achievement data collected from 15-year-old students in six of the countries that participated in PISA 2012 (Australia, Finland, Germany, Japan, Singapore, Korea). The results of the analysis showed that first, the number of profile groups and the variables with high and low achievement in the profile groups varied by country. Second, distribution of profile groups for student achievement indicated a hierarchical structure in all countries. Third, the largest difference between high-achievement and low-achievement profile groups was found to be the Formulate variable. Based on the analysis of students' mathematical modeling competency, several implications were derived.

First, comparison results by country, variable, and profile group can provide insight into understanding various characteristics of students' mathematical modeling competency. These results can be particularly valuable for teachers, education researchers, and policymakers. By providing comparison results by country, they can identify the relative position of South Korean students' mathematical modeling competency and ascertain the strengths and weaknesses of each variable. By offering a hierarchical structure in the comparison results by variable and profile group, teachers can monitor students' progress in mathematical modeling competency. Consequently, teachers can provide students with feedback to help them advance to the next profile group. According to Research Result 1, the variable with the highest average and variance of student achievement varied by country. In Australia, Finland, and Germany, the Interpret variable had the highest average, while in Japan, Singapore and South Korea, the Formulate variable had the highest average. Formulate and Employ variables had the highest standard deviation in South Korea, and Interpret variable had the highest standard deviation in Germany for student achievement. This indicates that in South Korea, the education gap is highest for Formulate and Employ variables, despite the Formulate variable having the highest average. Furthermore, Research Result 2 showed that the number and proportion of profile groups differed by country. That is, Singapore and Korea had 4 profile groups, Finland, Germany, and Japan had 5 profiles, and Australia had 6 profiles. Notably, In South Korea, the proportion of high-achieving profile groups A and B was higher than in other countries. Finally, according to Research Result 3, the high-achieving and low-achieving variables within each profile group varied by country. Specifically, in Australia, it is characteristic that the Interpret variable showed high achievement in all profile groups, but in other countries and across variables, it was diverse by profile groups. In all variables and all countries, the profile groups with the most outliers were found to be profile group A, which had the highest achievement, and profile group D, E or F which had the lowest achievement. Through these findings, it is anticipated that various insights into the characteristics of country-specific profile groups can be comprehensively considered and utilized in the development of mathematics education policies and systems.

Second, it's possible to predict a student's profile group and the score range of other variables based on the student's achievement for one of the mathematical modeling variables. In this regard, the Formulate variable could serve as the most suitable scale. According to research result 3, the distribution of profile groups for mathematical modeling variables exhibits a hierarchical structure in all countries, suggesting a high correlation among these variables. This indicates that a student's profile group could potentially be predicted based on their achievement in a single variable. Additionally, according to Research Results 1 and 3, the Formulate variable shows the largest standard deviation among mathematical modeling variables across all countries. It also exhibits the largest difference between the high and low achievement profile groups in the Formulate variable.

This is likely due to the wide distribution range of student achievement results for the Formulate variable, which suggests that it can provide a more sensitive scale compared to other variables. As a result, student achievement in the Formulate variable could be efficiently utilized to predict a student's profile group and achievement in other mathematical modeling variables, potentially leading to reductions in time and cost. Specifically, in the analysis of this study, activities associated with the established Formulate variable include identifying problems in real-world contexts, recognizing mathematical structures, simplifying problems, and mathematically representing situations. Hence, the study suggests the need to consider the Formulate variable, which is related to the achievement of these activities, as a measure in the achievement analysis of mathematical modeling.

On the other hand, this study takes an approach from the perspective that achievement for mathematical modeling variables can be measured through individual evaluation items. And it has limitations as it does not consider the cultural context of each country. It is suggested that further analysis studies, encompassing a variety of contexts and providing detailed insights into mathematical modeling competency and cultural factors, be conducted. Additionally, we propose a follow-up study to explore more in-depth the influencing factors that show the greatest difference in the Formulate variables among mathematical modeling variables.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*(3), 317–332. https://doi.org/10.1007/BF02294359

Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of condition mean- and covariance-structure models. *Psychometrika, 64,* 475–494. https://doi.org/10.1007/BF02294568

Baek, D., & Lee, K. (2018). Role and significance of abductive reasoning in mathematical modeling. *Journal of Educational Research in Mathematics, 28*(2), 221–240. https://doi.org/10.29275/jerm.2018.05.28.2.221

Blum, W., & Ferri, R. B. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application*, *1*(1), 45–58.

Blum, W., & Leiß, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling (ICTMA 12): Education, engineering and economics* (pp. 222–231). Horwood. https://doi.org/10.1533/9780857099419.5.221

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects—State, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*(1), 37–68. https://doi.org/10.1007/BF00302716

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *Zentralblatt für Didaktik der Mathematik, 38*(2), 86–95. https://doi.org/10.1007/BF02655883

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification, 13*, 195–212.

https://doi.org/10.1007/BF01246098

Chon, K., & Kim, S. (2019). A latent profile analysis of math attitudes. *Journal of Curriculum and Evaluation, 22*(2), 319–339. https://doi.org/10.29221/jce.2019.22.2.319

Clark, S. L., & Muthén, B. (2009). *Relating latent class analysis results to variables not included in the analysis*. Retrieved from http://www.statmodel.com/download/relatinglca.pdf.

English, L., & Watters, J. (2005). Mathematical modelling in the early school years. *Mathematics Education Research Journal, 16*(3), 58–79. https://doi.org/10.1007/BF03217401

Galbraith, P. (2012). Models of modelling: Genres, purposes or perspectives. *Journal of Mathematical Modelling and Application, 4*(8), 3–16.

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *Zentralblatt für Didaktik der Mathematik, 38*(2), 143–162. https://doi.org/10.1007/BF02655886

Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: A calculus course as an example. *Educational Studies in Mathematics 39*, 111–129. https://doi.org/10.1023/A:1003749919816

Haines, C., & Crouch, R. (2007). Mathematical modelling and applications: Ability and competence frameworks. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI Study* (pp. 417–424). Springer. https://doi.org/10.1007/978-0-387-29822-1_46

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods, 11*(1), 36–53. https://doi.org/10.1037/1082-989x.11.1.36

Hwang, J., & Ko, E. (2018). Investigating Korean students' different profiles of affective constructs and engagements: a latent profile analysis on TIMSS 2015. *Journal of the Korean School Mathematics Society, 21*(3), 207–225. http://dx.doi.org/10.30807/ksms.2018.21.3.001

Hwang, S., & Son, T. (2021). Exploring middle school students' mathematics learning motivation and influential factors using latent profile analysis. *Educational Research, 80*, 275–294. https://doi.org/10.17253/swueri.2021.80..013

Iverson, A., French, B. F., Strand, P. S., Gotch, C. M., & McCurley, C. (2018). Understanding school truancy: Risk–need latent profiles of adolescents. *Assessment, 25*(8), 978–987. https://doi.org/10.1177/1073191116672329

Jonassen, D. (2011). Supporting problem solving in PBL. *Interdisciplinary Journal of Problem-Based Learning, 5*(2). 95–119. https://doi.org/10.7771/1541-5015.1256

Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modelling in mathematics education. *Zentralblatt für Didaktik der Mathematik, 38*(3), 302–310. https://doi.org/10.1007/BF02652813

Kim, S., & Kim, K. (2004). Analysis on types and roles of reasoning used in the mathematical modeling process. *School Mathematics, 6*(3), 283–299.

Lesh, R., & Lehrer, R. (2003). Models and modeling perspectives on the development of students and teachers. *Mathematical Thinking and Learning, 5*(2–3), 109–129.

https://doi.org/10.1080/10986065.2003.9679996

Lesh, R., & Zawojewski, J. (2007). Problem solving and modeling. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 763–804). Information Age Publishing.

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

Maaß, K., & Gurlitt, J. (2011). LEMA–Professional development of teachers in relation to mathematical modelling. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. A. Stillman (Eds.), *Trends in teaching and learning of mathematical modelling* (pp. 629–639). Springer. https://doi.org/10.1007/978-94-007-0910-2_60

Marsh, H. W., Luedtke, O., Trautwein, U., & Morin, J.S.A. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling, 16,* 191–225. https://doi.org/10.1080/10705510902751010

Muthén, L. K., & Muthén, B. O. (1998). Mplus (version 8.8) [Computer software]. Muthén & Muthén.

OECD. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. OECD Publishing. https://doi.org/10.1787/9789264190511-en

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika,* *52*(3), 333–343. https://doi.org/10.1007/BF02294360

Sriraman, B., & Lesh, R. A. (2006). Modeling conceptions revisited. *Zentralblatt für Didaktik der Mathematik, 38*(3), 247-254. https://doi.org/10.1007/BF02652808

Verschaffel, L., Greer, B., & De Corte, E. (2002). Everyday knowledge and mathematical modeling of school word problems. In K. Gravemeijer, R. Lehrer, B. Van Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 257–276). Springer Netherlands. https://doi.org/10.1007/978-94-017-3194-2_16

Yi, H., & Lee, Y. (2017). A latent profile analysis and structural equation modeling of the instructional quality of mathematics classrooms based on the PISA 2012 results of Korea and Singapore. *Asia Pacific Education Review, 18*, 23–39. https://doi.org/10.1007/s12564-016-9455-4

Yoon, C., Thomas, M., & Dreyfus, T. (2011). Grounded blends and mathematical gesture spaces: developing mathematical understandings via gestures. *Educational Studies in Mathematics*, *78*(3), 371–393. https://doi.org/10.1007/s10649-011-9329-y