



RELATIONSHIP BETWEEN MATHEMATICAL LITERACY AND OPPORTUNITY TO LEARN WITH DIFFERENT TYPES OF MATHEMATICAL TASKS

Jihyun Hwang¹, Yeajin Ham²

¹Kangwon National University, Chuncheon-Si, Gangwon-Do, South Korea

²University of Iowa, Iowa City, Iowa, USA

Email: yeajin-ham@uiowa.edu

Abstract

We investigated how the opportunity to learn (OTL) with different types of mathematics tasks are related to mathematical literacy and the role of perceived control in the relationship between OTL and mathematical literacy. The structural equation modeling was applied to the data of 1,649 Korean students from the PISA 2012 database. OTL with the four different types of tasks – algebraic word problems, procedural tasks, pure mathematics reasoning, and applied mathematics reasoning – were measured via student survey on how often they have encountered each type of task in their mathematics lessons and tests. The results showed that OTL with the procedural tasks was likely to increase mathematical literacy directly and indirectly through internal perceived control. Engaging in the applied reasoning tasks is positively related to external perceived control, but negatively to mathematical literacy.

Keywords: Opportunity to Learn, Mathematical Tasks, Mathematical Literacy, Perceived Control, PISA 2012

Abstrak

Kami menyelidiki tentang bagaimana Kesempatan Belajar (KB) siswa dengan berbagai jenis tugas matematika yang terkait dengan literasi matematika dan peran *perceived control* dalam hubungan antara KB dan literasi matematika. Pemodelan persamaan struktural diterapkan pada data 1.649 siswa Korea dari database PISA 2012. KB dengan empat jenis tugas, yaitu soal cerita aljabar, tugas prosedural, penalaran matematika, dan penalaran matematika terapan, yang diukur melalui survei siswa tentang seberapa sering mereka menjumpai setiap jenis tugas dalam pelajaran dan tes matematika mereka. Hasil penelitian menunjukkan bahwa KB dengan tugas prosedural cenderung meningkatkan literasi matematika secara langsung dan tidak langsung melalui *perceived control* internal. Keterlibatannya dalam tugas penalaran yang diterapkan, bernilai positif terhadap *perceived control* eksternal, namun bernilai negatif terhadap literasi matematika.

Kata Kunci: Kesempatan Belajar, Soal Matematika, Literasi Matematika, *Perceived Control*, PISA 2012

How to Cite: Hwang, J., & Ham, Y. (2021). Relationship Between Mathematical Literacy and Opportunity to Learn with Different Types of Mathematical Tasks. *Journal on Mathematics Education*, 12(2), 199-222. <http://doi.org/10.22342/jme.12.2.13625.199-222>

Since Carroll (1963) introduced the concept of opportunity to learn (OTL), OTL has been conceptualized as the inputs and processes that are needed to produce student achievement of intended outcomes (Elliott & Bartlett, 2016). Building on the early conceptualization of OTL as allocation of learning time (Carroll, 1963; Cogan & Schmidt, 2015), studies on mathematical practices and OTL are concerned with the processes through which individuals come to know mathematics content (Barnard-Brak et al., 2018). A contemporary definition of OTL is comprised of factors that have a significant influence on teachers' instructional practices and students' learning (Stevens & Grymes, 1993); these factors include content coverage and emphasis. Content coverage refers to which concepts and cognitive skills of curricula are covered during classroom learning, whereas content emphasis is related to activities and tasks that engage students (Stevens & Grymes, 1993).

In this study, OTL in mathematics classrooms is conceptualized as mathematical tasks that allow learners to have actual experiences with mathematics, which focuses on OTL as content emphasis in learning tasks (Schoenfeld, 1992). Learners' cognitive processes are shaped through the experiences with learning tasks. Further, by engaging in the mathematical tasks, students develop their understanding of what it means to *do mathematics* (Schoenfeld, 1994). By tackling various types of mathematical tasks, students can do mathematics and construct an epistemological understanding of what doing so means. Therefore, we conceptualize OTL as cognitive processes that learners engage in while doing mathematics through engaging in different types of tasks.

OTL, which is conceptualized around mathematical tasks, is related to students' learning outcomes, according to the framework suggested by Stein et al. (1996). However, a large body of literature based on this framework has focused on teachers' instructional practices, such as teachers' implementation of mathematical tasks that require high cognitive demands. There are few previous research studies that exist on the relationship between students' OTL and their learning outcomes, and these studies' conceptualization of OTL did not include mathematical tasks used in the mathematics instruction. For example, OTL in the study of Ottmar et al. (2014) consisted of two dimensions: instructional quality (teachers' efforts to promote reasoning and understanding of concepts via teacher-student interaction) and exposure to mathematics instruction (how long students were exposed to mathematics instruction). To bridge this research gap, we question whether there is a relationship between OTL (students' exposure to different types of mathematical tasks) and mathematics achievement.

Due to the complex nature of learning environments (Berliner, 2002; Jacobson et al., 2019) we hypothesize that the relationship between our conceptualization of OTL (frequency and type of mathematics tasks) and mathematics achievement is not only a direct, but also an indirect relationship through other factors. According to the framework of mathematical instructional tasks (Stein et al., 1996), the mathematical tasks that are set up by the teacher interact with and shape students' dispositions, including attitudes, beliefs, and motivation. This interaction between mathematical tasks and learning disposition in turn, influences students' cognitive processes and learning behaviors (Henningsen & Stein, 1997). Finally, the overall processes that involve mathematical tasks and students' perception are reflected in the students' learning behaviors and outcomes. For this reason, we consider students' perceived control as a mediating factor between OTL and achievement. Students who believe that academic outcomes are under their own control are predicted to be more actively engaged in mathematical tasks and earn better academic outcomes. This interaction between learning tasks and students' disposition (specifically, perceived control in this study) has not been found in previous literature.

This study is a secondary analysis that uses the database from the Program for International Student Assessment (PISA) 2012. Using extensive data from the PISA 2012 database, we investigated the relationship between OTL, mathematical literacy, and perceived control using structural equation

modeling (SEM) approach. In this study, we focused on one educational context, South Korea, rather than examining different contexts of multiple countries. Before making international comparisons, an exploratory study to understand the phenomenon in a single context would be required for meaningful conclusions. Moreover, the rationale for selecting South Korea is that it is one of the high achieving countries, and that it has not been fully investigated in terms of OTL (Son, 2012). Furthermore, the OECD working paper (Schmidt, Zoido, & Cogan, 2014) showed that in each country, there is linear or quadratic relationship between exposure to different types of mathematical tasks and mathematical literacy. Since we focus on Korean student data and include perceived control in SEM analysis, our study can provide a broader picture of the relationship among different types of mathematics tasks, perceived control, and mathematical literacy.

The purpose of this study is to explore the relationship between OTL – a combination of exposure and types of mathematical tasks – and mathematical literacy measured in the PISA 2012. The nature of the relationship between OTL and mathematical literacy is not direct, and rather, the relationship is mediated by perceived control. OTL is hypothesized to be related to mathematical literacy via students' perceived control.

Opportunity to Learn

Within an educational context, OTL refers to the inputs and processes that are provided to students for intended learning outcomes (Elliott & Bartlett, 2016). One of the first conceptualizations of OTL focused on sufficient time and adequate instruction to learn (e.g., Carroll, 1963; Schmidt, 1992). With growing interest in the concept of OTL in relation to the demand for curricular validity, the concept of OTL has been expanded to accommodate a multi-dimensional construct that encompasses both the quality of instruction and its alignment with the assessment of learning outcomes (Abedi & Herman, 2010). Specifically, Stevens (1996) proposed a comprehensive conceptual framework of OTL that includes four elements: content coverage, content exposure, content emphasis, and quality of instructional delivery. As such, OTL, as a comprehensive and multi-dimensional concept, offers a basis for investigating students' learning in the mathematics classroom (e.g., Abedi & Herman, 2010).

When considering how these different dimensions of OTL are realized in the mathematics classroom, it is clear that mathematical tasks serve as a critical learning space that provide students with experiences of mathematical practice. In other words, mathematical tasks that comprise different dimensions of OTL (e.g., content coverage, content exposure, content emphasis, and quality of instructional delivery) interact with and, in turn, shape students' learning processes, both cognitive and non-cognitive. For example, in the studies of Watson (2003) and Törnroos (2005), class tasks, in addition to the curriculum and the textbook, were identified as one of the critical aspects of OTL.

In the PISA 2012, OTL is conceptualized as a constellation of three constructs that describe classroom learning environments: (1) measurement of content, (2) teaching practices, and (3) teaching quality (OECD, 2013). According to Schmidt et al. (2014), OTL in the PISA 2012 refers to the content

students learn, as well as the cohesiveness that exists between what is taught and what they actually learn. Also, students' experiences with mathematical content are shaped by instructional practices, including student-centered instruction and lectures. Students' OTL is characterized by the factors underlying the quality of instructional practices, such as classroom organization, emotional support, and cognitive activation (OECD, 2013).

In the frameworks that are used in previous studies and PISA 2012, the concept of OTL includes specific content that is covered in mathematics classrooms, as well as mathematical tasks that deliver mathematics content. On one hand, the commonality of these frameworks is that mathematical tasks are an important factor of mathematics learning, and that teachers can affect students' cognitive and motivational processes of learning by designing these tasks. On the other hand, we also recognize differences among the frameworks that conceptualized OTL. One difference between the PISA 2012 framework and other literature on OTL is that in the PISA, OTL is operationalized as students' judgment on whether and how often they have encountered different mathematical tasks. This operationalization for measurement is partly limited in covering depth of teaching or quality of instructional delivery variables (E.g., Stevens, 1993), which is also recognized in the PISA 2012 framework (OECD, 2013, p.187). As such, we do recognize the multifaceted characteristic of OTL, but also acknowledge that large-scale assessment would not be enough to fully understand OTL that students experience in mathematics classrooms as reported in the PISA 2012 framework. In this study, we assume the operationalization of OTL in the PISA 2012, student-reported frequency of being exposed to different types of mathematical tasks.

Mathematical Tasks

Among the multiple aspects of OTL, we highlight students' exposure to different types of mathematical tasks in lessons and tests, as the tasks themselves reflect what content the students learn and what doing mathematics entails (Stein et al., 1996). In other words, mathematical tasks are essential tools for ensuring that students can understand mathematical concepts more fully, as well as to develop cognitive processes of mathematical reasoning via their experience with the tasks (Martin & Gourley-Delaney, 2014).

With regard to the cognitive processes of learning, the students' experience of mathematics depends on the level of cognitive demands, how the tasks are presented, and how the tasks are implemented. Adopting the conceptual framework regarding the relationship between variables that are related to tasks and students' learning outcomes (Stein et al., 1996), many studies have shown that cognitive demands of mathematical tasks can change as they are implemented (e.g., Boston & Smith, 2009; Henningsen & Stein, 1997). When students engage in mathematics, their reasoning differs according to what type of mathematical tasks are being offered (see *Potential of the Task* in Boston & Smith, 2009). Mathematical thinking processes that students employ are closely related to the mathematical tasks that are embedded in the learning context (Henningsen & Stein, 1997). Certainly,

the elements and characteristics of mathematical tasks require students to engage in different cognitive processes. Hanna and Jahnke (2007) provided a good example by comparing activities that involve either pure mathematics or a real-life situation. Proving statements is a combination of two processes: “(1) finding the ‘right premises’ and (2) devising the chain of deductive steps leading from the premises to the statement” (p. 149). Mulnix (2012) labeled these as the process of *searching for reasons* (e.g., abduction/induction) and the process of *giving reasons* (e.g., deduction). According to Hanna and Jahnke (2007), the process of giving reasons is more emphasized in tasks that involve pure mathematical reasoning, which is why the process of searching for reasons has usually been downplayed. In contrast, mathematical tasks with real situations require setting up the premise first (searching for reasons), which is followed by the process of building logical connectedness (giving reasons).

Previous studies have scrutinized mathematical tasks set up by teachers and teachers’ actual implementation of the tasks based on the framework developed by Stein et al.’s (1996) (e.g., Arbaugh & Brown, 2005; Boston & Smith, 2009). However, few studies have been conducted to investigate the link between mathematical tasks and students’ learning outcomes, particularly measured by large-scale assessments. This is possibly because of the assumption that large-scale assessments are designed to evaluate students’ content knowledge, not their mathematical practices (Lane, 2004). However, we argue that students formulate and utilize epistemic and cognitive resources to reason through OTL with mathematical tasks (Hammer, 2000), and students utilize some of those resources to solve problems in assessments (Bailin & Siegler, 2003; Hwang et al., 2020). The common cognitive resources used while engaging in mathematics tasks and assessment settings can help us to understand how students’ OTL is connected to achievement scores in large-scale assessments.

The relationships between OTL and achievement can also be influenced differ by what mathematics tasks are involved in students’ OTL. Individual differences in mathematics learning can be understood as interaction between features of tasks and students’ inputs (i.e., cognitive resources, affectivity; Bornemann et al., 2010; Muis et al., 2015). As discussed, students’ mathematical reasoning, as one of the critical components of doing mathematics, differs by what type of mathematical tasks are offered to them. The emotional components, such as task valuing and perceived control (Muis et al., 2015), can motivate them to either continue reasoning or terminate the reasoning process (McLeod, 1992). Particularly, students’ perceived control – “the tendency of people to perceive that outcomes in a particular arena were either within or outside of their control” (McNabb, 2003, p. 418) – influences students’ approaches to solving mathematical tasks. For example, students are likely to engage more actively when they believe that the outcomes from engagement in tasks are under their control (Hrbáčková et al., 2012).

Perceived Control

Control beliefs refer to an overall set of beliefs about how effective one’s process of producing expected outcomes can be (Skinner et al., 1998). In academic settings, perceived control is understood

as a critical psychological disposition that affect students' behavior, emotion, and achievement (d'Ailly, 2003; Schunk, 1984; Murayama et al., 2013). According to the previous frameworks of perceived control (e.g., Skinner et al., 1998; Rotter, 1966; Rotter & Mulry, 1965), perceived control over learning is constituted of two types of beliefs: strategy beliefs (what it takes to do well) and capacity beliefs (whether I believe I have the strategies; Skinner et al., 1998). According to Rotter (1966), people differ in their beliefs whether outcomes occur independently of how one behaves (external control) or are highly contingent on one's behavior (internal control). The construct of locus of control assume that internal and external causes are inversely related to each other and thus, can be assessed as a single, bipolar dimension (Skinner et al., 1990). Though perceived control has been shown to be an important indicator of students' motivation in learning (Patrick et al., 1993), previous studies on perceived control rarely examined it in the relation to the success in academic tasks (Skinner et al., 1990; Lipnevich et al., 2016).

In PISA 2012, the conceptualizations of perceived control and other self-perceptions are based on the *planned behavior theory* of Ajzen (2002). According to Ajzen (2002), perceived control belief is conceptualized as a person's belief about the ease or difficulty of performing a behavior and this belief forms a behavioral intention that directly increases the likelihood of a desired behavior. In this study, we viewed locus of control as having two types and identified student survey items that ask about their locus of control over mathematics learning and categorized them into internal and external perceived control. It is assumed that a learner's strong internal perceived control does not necessarily lead to weak external perceived control, and furthermore, they are qualitatively different with various sources of beliefs.

Mathematical Literacy in the PISA

The relationships between OTL and achievement can differ by how achievement is defined, and with what measure it is assessed. In this study, achievement scores represent students' mathematical literacy measured with the PISA 2012. According to the PISA 2015 framework, mathematical literacy "explains the processes content knowledge, and contexts reflected in the assessment's mathematics problems", and this shows how students perform in mathematics (OECD, 2017). The construct of mathematical literacy describes competency of individuals to reason mathematically and use math concepts, procedures, facts, and tools to describe, explain, and predict phenomena (OECD, 2017). This conceptualization of mathematical literacy supports the importance of students' engagement in pure mathematics tasks (reason mathematically) and their exploration in the abstract world of mathematics (use math concepts, procedures, facts, and tools; OECD, 2017).

When contemplating PISA's definition of mathematical literacy, it also emphasizes the capacities to formulate problem situations, employ mathematical problems, and interpret mathematics results in various contexts. In other words, rich experiences of real-world tasks in math classrooms are essential in developing these capacities. Accordingly, having experiences of doing mathematics in real world

contexts (personal, societal, occupational, and scientific situations) contributes to the development of mathematical literacy.

Hypothesized Model

According to the literature review, we suggest the hypothesized model in Figure 1 representing the relationships among OTL by different types of tasks, perceived control, and mathematical literacy. This model includes a latent variable for each type of mathematical tasks (word problems, procedural tasks, pure mathematics reasoning, and applied mathematics reasoning tasks). In addition, there are two latent variables for internal and external perceived control in the model. By evaluating the appropriateness of the hypothesized model, we aimed to answer the following questions: (1) what are the relationships between opportunities to learn with different types of tasks and perceived control? (2) what are the relationships between opportunities to learn with different types of tasks and mathematical literacy measured in the PISA 2012?

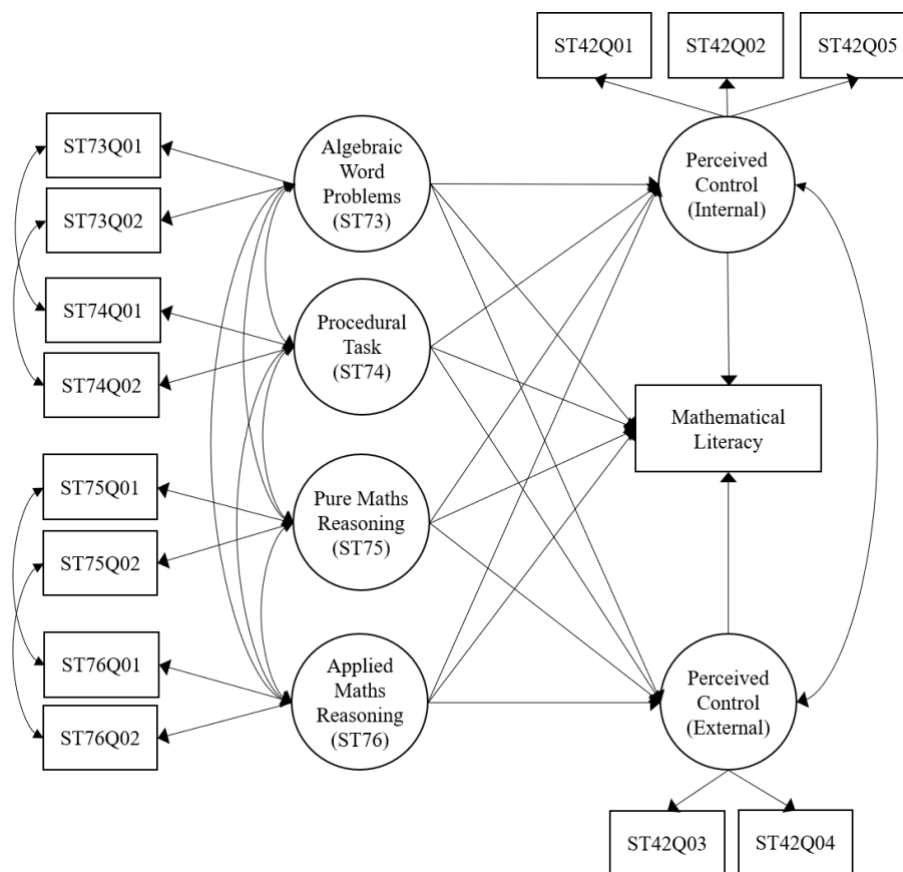


Figure 1. Full Model of the Relationship among OTL, Perceived Control, and Mathematical Literacy

It should be noted that each of the latent variables of OTL and EPC was estimated using two indicators due to the structure of the PISA 2012 data. It is commonly recommended to have more than two indicators per latent variable. However, some researchers argued that one or two indicators could

be sufficient (Hayduk & Littvay, 2012). Furthermore, the analysis did not yield any errors that are very likely to happen with two indicators (e.g., negative residual variances known as Haywood cases), and the number of indicators showed little effect of bias (Little et al., 1999). For those reasons, the estimated model is valid to interpret the relationship between OTL and mathematical literacy.

Thus, we are interested in both direct relationship and indirect relationships through perceived control, between mathematical tasks and mathematical literacy. We attempted to test the positive relationships between OTL with different tasks and mathematical literacy by examining the hypothesized model with the structural equation modeling. When students learn mathematics through OTL with mathematical tasks, the positive relationships between the tasks and mathematical literacy are somewhat expected. Particularly, we expected that students' opportunity to engage in applied mathematics reasoning tasks would be more strongly related to mathematical literacy based on the definition of mathematical literacy given by the PISA 2015.

METHOD

Participants

We utilized the PISA 2012 international database, which is open to the public. The rationale to use this database instead of the PISA 2015 or 2018 was that the focus subjects of these recent PISA studies were not mathematics. The variables included to address our research questions were available only in the PISA 2012. Among 5,033 Korean students in the original PISA 2012 database we collected responses of 1,649 Korean students who participated in both student questionnaire and mathematical literacy assessment. The PISA 2012 student context questionnaires in a rotation design, which consisted of the 'common' question (answered by all students) and 'rotated' questions (answered by two thirds of the student sample; OECD, 2014, p. 59). Because all of the survey items used in this research (ST43, questions asking 'Thinking about your mathematics lessons: to what extent do you agree with the following statements?' and ST73-76, mathematical tasks) were included together in the form A, the rotation questionnaires design allowed us to observe only students taking this form, a third of the Korean students participating in the PISA 2012 (See figure 3.9 in OECD, 2014, p. 61).

Variables

Mathematical literacy. As seen in the hypothesized model, we collected all sets of plausible values representing students' mathematical literacy scores provided in the PISA 2012. Large-scale international studies such as TIMSS and PISA do not provide one value for each student's achievement in mathematics. Rather, as Foy, Brossman, and Galia (2012) argued, plausible values are provided through the process called "conditioning" with all background variables, for which relationships between background variables and mathematics achievement can serve as a satisfactory explanation.

Table 1. Weighted Means and Standard Deviations of Mathematical Literacy

Plausible Value	PV1	PV2	PV3	PV4	PV5
Weighted Mean	553.57	553.53	555.08	553.56	553.44
Standard Deviation	99.61	100.22	99.53	100.52	101.30

Furthermore, we highlighted that “[p]lausible values are not test scores and should not be treated as such” (OECD, 2014, p. 147) and the plausible values should be analyzed in a correct way. According to von Davier, Gonzalez, and Mislevy (2009), averaging plausible values themselves to have one value representing students’ mathematical literacy could lead to biased estimates. Chaney et al. (2001) suggested conducting separate analysis with each set of plausible values and average the estimated parameters. We also applied some formulas that Chaney and his colleagues provided to compute the standard errors for calculated estimates. Lastly, mathematical literacy was standardized in the SEM analysis. Table 1 shows the weighted mean and the standard deviation of each set of the plausible values.

Perceived control. We collected students’ responses to the question, given the code, ST43, asking students’ degrees of agreements to six statements in Table 2. For data analysis, we assigned “4” to students’ strong agreement to each statement, “3” to moderate agreement, “2” to moderate disagreement, and “1” to strong disagreement. Though the way of assigning numbers to students’ responses is different from the way used in the PISA 2012, our method allows us to interpret that higher numbers of students’ responses indicate stronger agreement to the statements about perceived control.

After selecting the data of the question ST43, we categorized the six statements into the two: internal (IPC) and external perceived control (EPC) based on the discussion to build the hypothesized model. Internal perceived control was measured through the three statements – ST43Q01, ST43Q02, and ST43Q05. Other two statements – ST43Q03 and ST43Q04 – were used to measure external perceived control. We excluded the statement, ST43Q06 that asked about test preparation because it is neither internal nor external perceived control. This statement can imply that students’ performance on mathematics exams is irrelevant to their efforts, but the statement itself does not allow us to identify the statement as either internal or external perceived control.

As seen in Table 2, more than 85% of Korean students agreed with the three statements ST43Q01, ST43Q02, and ST43Q05, which were used to measure internal perceived control. Simultaneously, most Korean students disagreed with the other statements about external perceived control and test preparation. When ST43Q03 and ST43Q04 were compared, it was interesting that more students strongly agreed that their success/failure is attributed to their teachers.

Table 2. The Number of Students and Weighted Percentage by Response to the Six Statements

Code	Question: Thinking about your mathematics lessons: to what extent do you agree with the following statements? Statement	Frequency			
		Strongly Agree (4)	Agree (3)	Disagree (2)	Strongly Disagree (1)
ST43Q01	If I put in enough effort I can succeed in mathematics.	503 (30.1%)	945 (57.7%)	166 (10.1%)	35 (2.1%)
ST43Q02	Whether or not I do well in mathematics is completely up to me.	612 (37.0%)	912 (55.5%)	93 (5.6%)	32 (1.9%)
ST43Q03	Family demands or other problems prevent me from putting a lot of time into my mathematics work.	66 (3.9%)	311 (18.8%)	940 (57.2%)	332 (20.1%)
ST43Q04	If I had different teachers, I would try harder in mathematics.	116 (7.1%)	333 (20.0%)	912 (55.3%)	288 (17.6%)
ST43Q05	If I wanted to, I could do well in mathematics.	482 (29.0%)	951 (57.9%)	172 (10.5%)	44 (2.6%)
ST43Q06	I do badly in mathematics whether or not I study for my exams.	108 (6.6%)	425 (26.0%)	793 (48.2%)	323 (19.2%)

Opportunity to Learn. Students' OTL data were collected with the responses to four questions (ST73, ST74, ST75, and ST76). Each question included two sub-questions showing different types of tasks: "how often have you encountered these types of problems in your mathematics lessons (ST[73–76]Q01)?"; and "in the tests you have taken at school (ST[73–76]Q02)?" We highlight that the four questions focused on students' perception of *how often* they encountered OTL with the tasks, among other dimensions of OTL (Stevens & Grymes, 1993). Table 3 shows the detailed questions for OTLs labeled with *algebraic word problem* (WP; ST73) and *procedural tasks* (PT; ST74) in the PISA 2012. Table 4 also shows the other two questions used to identify OTL with *pure mathematics reasoning* (PMR) and *applied mathematics reasoning* (AMR).

Table 3. OTL Questions for Algebraic Word Problem and Procedural Task (OECD, n.d.)

Algebraic Word Problem (WP; ST73)	Question. In the box is a series of problems. Each requires you to understand a problem written in text and perform the appropriate calculations. Usually the problem talks about practical situations, but the numbers and people and places mentioned are made up. All the information you need is given. Here are two examples:				
	Example 1 <Ann> is two years older than <Betty> and <Betty> is four times as old as <Sam>. When <Betty> is 30, how old is <Sam>?				
	Example 2 Mr. <Smith> bought a television and a bed. The television cost <\$625> but he got a 10% discount. The bed cost <\$200>. He paid <\$20> for delivery. How much money did Mr. <Smith> spend?				
	# of Students	Frequently (4)	Sometimes (3)	Rarely (2)	Never (1)
	ST73Q01	545 (33.1%)	821 (49.8%)	215 (13.1%)	68 (4.0%)
	ST73Q02	336 (20.5%)	792 (48.2%)	410 (24.7%)	111 (6.6%)

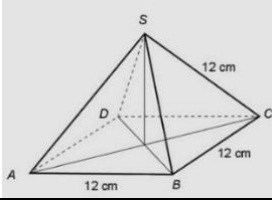
Question. Below are examples of another set of mathematical skills.

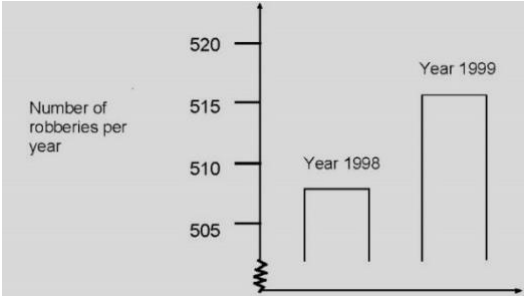
Procedural Tasks (PT; ST74)	Example 1) Solve $2x + 3 = 7$.				
	Example 2) Find the volume of a box with sides 3m, 4m and 5m.				
	# of Students	Frequently (4)	Sometimes (3)	Rarely (2)	Never (1)
	ST74Q01	968 (58.9%)	540 (32.6%)	109 (6.6%)	32 (1.8%)
	ST74Q02	746 (45.5%)	623 (37.8%)	217 (13.1%)	63 (3.7%)

Note. All percentages were weighted.

An interesting finding from Tables 3 and 4 is that more than 70% of students answered that they encountered each of WP, PT, and PMR frequently in their mathematics lessons and tests. However, approximately a half of students reported that they encountered AMR frequently or sometimes. Particularly, 91.4% of students encountered PT at least sometimes, whereas 44.2% of students saw AMR tasks at most rarely in lessons. This indicates that there were substantial gaps in the frequencies of OTL with different tasks that were offered to Korean students; specifically, limited OTL with AMR that requires students to make sense of real problem situations and interpret/explain the solutions. This is a unique mathematical process that AMR offers, while other task types do not.

Table 4. OTL Questions for Pure Mathematics Reasoning and Applied Mathematics Reasoning (OECD, n.d.)

Pure Mathematics Reasoning (PMR; ST75)	Question. In the next type of problem, you have to use mathematical knowledge and draw conclusions. There is no practical application provided. Here are two examples.				
	 <p>Example 1) Here you need to use geometrical theorems: Determine the height of the pyramid. Example 2) If n is any number: can $(n+1)^2$ be a prime number?</p>				
	# of Students	Frequently (4)	Sometimes (3)	Rarely (2)	Never (1)
	ST75Q01	560 (34.0%)	722 (44.1%)	281 (16.8%)	86 (5.0%)
	ST75Q02	511 (31.0%)	719 (44.0%)	308 (18.6%)	111 (6.5%)

Applied Mathematics Reasoning (AMR; ST76)	Question. In this type of problem, you have to apply suitable mathematical knowledge to find a useful answer to a problem that arises in everyday life or work. The data and information are about real situations. Here are two examples.		
	<p>Example 1) A TV reporter says “This graph shows that there is a huge increase in the number of robberies from 1998 to 1999.”</p>  <p>Do you consider the reporter’s statement to be a reasonable interpretation of the graph? Give an explanation to support your answer.</p> <p>Example 2) For years the relationship between a person’s recommended maximum heart rate and the person’s age was described by the following formula:</p>		
$Recommended\ maximum\ heart\ rate = 220 - age$			

Recent research showed that this formula should be modified slightly. The new formula is as follows:

$$\text{Recommended maximum heart rate} = 208 - (0.7 \times \text{age})$$

From which age onwards does the recommended maximum heart rate increase as a result of the introduction of the new formula? Show your work.

# of Students	Frequently (4)	Sometimes (3)	Rarely (2)	Never (1)
ST76Q01	203 (12.4%)	717 (43.6%)	564 (34.1%)	165 (9.8%)
ST76Q02	187 (11.3%)	631 (38.6%)	618 (37.2%)	218 (12.9%)

Note. All percentages were weighted.

Data Analysis

Using the variables described above, we applied the structural equation modeling (SEM) to evaluate the hypothesized model in [Figure 1](#). The SEM approach was utilized with the R package *lavaan.survey* (Obserski, 2016) and maximum likelihood estimation that considered all variables as continuous. The strength of this R package was that the complex PISA 2012 hierarchical design could be fully considered in the SEM analysis using students' weights and balanced repeated replications (BRR). First, because our research interests were solely at the student level, the data analysis required to use student weights in data analysis (Asparouhov & Muthen, 2006). The PISA 2012 provided "final trimmed nonresponse adjusted student weight," which was calculated with the consideration of stratified sampling design. Thus, statistical results such as descriptive statistics and SEM results were weighted. Second, weighting was not enough to make unbiased decisions when multilevel sampling was applied. Particularly, "the variance estimator can be unstable" relying on the sample design (OECD, 2017, p. 123). To resolve this issue, it was recommended to use BRR to estimate sampling variances (OECD, 2017). In this research, we employed Fay's method of BRR by using variables "final student replicate BRR-Fay weights" in the databases.

Considering the complexity of the analysis using the five sets of plausible values and Fay's method of BRR, we applied the three steps of the SEM approach suggested by Byrne (1998): model specification, model assessment, and model respecification. First, as discussed in the previous section, the hypothesized model in [Figure 1](#) was already constructed based on the relevant literatures. Second, the followings were evaluated for the next stage of respecification: the overall model fits, the suitability of parameter estimates, and the statistical significance of parameter estimates. It was checked that the outputs included some error messages like negative variances, correlations greater than 1, and non-positive definite covariance matrices, which all are unreasonable (Byrne, 1998). Furthermore, non-significant parameters could slightly contribute to the power of the model to explain the phenomenon. Thus, we considered the parameters with $p < 0.1$ because we had less concern of Type I error. For the overall model evaluation, all model fits were comprehensively evaluated using criteria summarized by Schreiber, Nora, Stage, Barlow, and King (2006, p. 330) – the comparative fit index (CFI), the Tucker-

Lewis fit index (TLI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA).

On one hand, we removed some non-significant indicators to respecify the hypothesis model after model assessments. On the other hand, we included correlated residuals having large modification indices representing expected changes in model fits. However, we should have theoretical rationales in addition to the statistical evidence to add correlated residuals (Kline, 2011). Accordingly, we considered four pairs of correlated residuals: (1) ST73Q1 and ST74Q1, (2) ST73Q2 and ST74Q2, (3) ST75Q1 and ST76Q1, and (4) ST75Q2 and ST76Q2. These pairs showed statistical evidence of large modification indices. Also, it is noticeable that residuals of the questions about WP and PT, and PMR and AMR were correlated, which could be due to the cognitive demands of mathematics tasks (Boston & Smith, 2009).

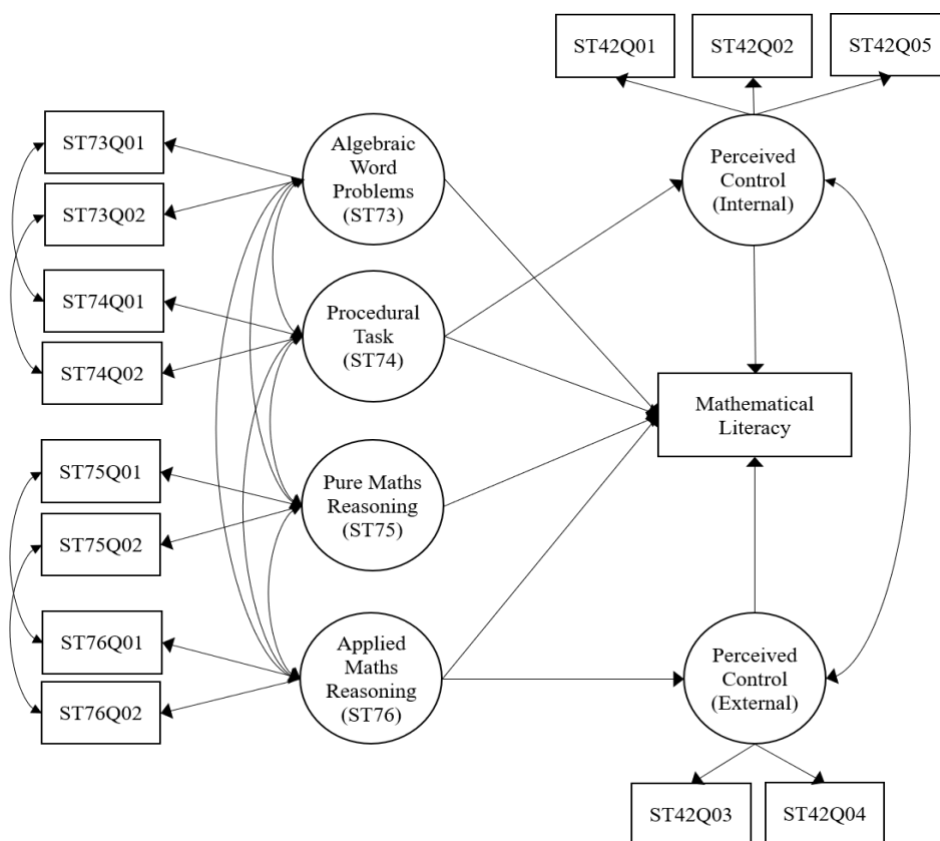


Figure 2. Nested Model Respecified from the Full Model

After the respecification, we compared the full model (see Figure 1) and the nested model (see Figure 2). Using χ^2 tests, AIC, BIC, and sample-size adjusted BIC, the comparison could answer whether there was a significant difference in the goodness of fit between the nested and full models, though the nested model had a smaller number of parameters. If we retain the null hypothesis that there is no significant difference between the models, we prefer the nested model to the full model, because the nest model had a similar goodness of fit with less parameters. Then, the SEM results of the relationships

among OTLs, perceived control, and mathematical literacy would be based on the regression weights in the selected model.

We highlight that estimation methods in SEM (e.g., maximum likelihood in this study) require a normality assumption of endogenous variables. However, it is known that parameter estimates are robust against violation of normality assumption while Type I error rate of hypothesis tests on individual parameters are likely to be inflated. Furthermore, plausible values of mathematical literacy measured in PISA 2015 were constructed based on normal population distributions. Thus, we argued the robustness of the model.

RESULTS AND DISCUSSION

We will report the fit indices to compare the nested and full models to answer our research questions. After the discussion of model selection process, we will report the SEM results to explain the relationships among OTLs, perceived control, and mathematical literacy.

Table 5. Model Fit Indices using Five Plausible Values

		CFI	TLI	Information Index			RMSEA	SRMR	χ^2 test p-value
				AIC	BIC	Adjusted BIC	Point Estimate & 90% Confidence Interval		
PV1	Full	0.977	0.962	43687.9	44036.9	43830.4	0.049 (0.043 0.055)	0.021	0.784
	Nested	0.977	0.966	43679.6	43996.3	43808.9	0.046 (0.040 0.052)	0.022	
PV2	Full	0.977	0.961	43715.9	44064.9	43858.4	0.049 (0.043 0.055)	0.022	0.785
	Nested	0.977	0.965	43707.5	44024.3	43836.8	0.046 (0.041 0.052)	0.023	
PV3	Full	0.976	0.960	43678.0	44027.0	43820.5	0.050 (0.044 0.056)	0.023	0.780
	Nested	0.977	0.964	43669.7	43986.4	43799.0	0.047 (0.041 0.053)	0.023	
PV4	Full	0.978	0.962	43702.3	44051.3	43844.8	0.048 (0.042 0.054)	0.022	0.784
	Nested	0.978	0.966	43693.9	44010.7	43823.3	0.046 (0.040 0.051)	0.023	
PV5	Full	0.977	0.961	43702.6	44051.6	43845.1	0.049 (0.043 0.055)	0.022	0.787
	Nested	0.977	0.966	43694.2	44011.0	43823.5	0.046 (0.040 0.052)	0.023	

Note. Bold numbers indicate the better model between the nested and full models.

Model Comparison

The model fit indices of the full and nested models were estimated across all sets of plausible values of mathematical literacy. Based on the recent criteria (CFI \geq 0.95, TLI \geq 0.95, RMSEA $<$ 0.06 with confidence interval, and SRMR \leq 0.08; Schreiber et al., 2006), all indices of both models were acceptable. When we compared the full and nested models, the χ^2 -test results indicated that there were no significant differences in the goodness of fit between the two models as seen in the last column of Table 5. Additionally, most indices – CFI, TLI, RMSEA – showed that the nested model had slightly better fit indices with the smaller number of the parameters. Less values of information indices (AIC,

BIC, and adjusted BIC) indicated a better model, which led to the same conclusion. Thus, we selected the nest model in Figure 2, which was a more parsimonious model with the similar goodness of fit.

SEM Results

Table 6 reports the estimated measurement model in the standardized metric. Lower factor loadings indicate that the corresponding indicators were conceptually distant from the latent variables. All factor loadings except for that of ST43Q03 were greater than 0.4 with $p < 0.001$, which satisfied previously suggested recommendations (Tabachnick & Fidell, 2007). Although the factor loading of ST43Q03 was 0.387, we argue that this coefficient was acceptable. However, the factor loadings for external perceived control were similar, which means that both statements can reflect conceptually similar distance of different facets of external perceived control – teachers and family.

Table 6. Results from the Measurement Model

Observed Variable	Latent Variables	Coefficient	SE	z-value	p-value
ST73Q01	WP	0.675	0.026	25.502	<0.001
ST73Q02	WP	0.621	0.022	27.759	<0.001
ST74Q01	PT	0.632	0.023	27.447	<0.001
ST74Q02	PT	0.642	0.022	28.933	<0.001
ST75Q01	PMR	0.786	0.022	35.501	<0.001
ST75Q02	PMR	0.751	0.022	33.926	<0.001
ST76Q01	AMR	0.746	0.021	35.454	<0.001
ST76Q02	AMR	0.744	0.021	35.330	<0.001
ST43Q01	IPC	0.540	0.018	29.934	<0.001
ST43Q02	IPC	0.444	0.019	23.347	<0.001
ST43Q05	IPC	0.505	0.017	29.676	<0.001
ST43Q03	EPC	0.387	0.056	6.910	<0.001
ST43Q04	EPC	0.439	0.069	6.332	<0.001

The SEM results included the correlation coefficients between OTLs with different types of tasks as seen in Table 7. Overall, all OTLs were highly correlated with each other, which means that if students have more frequent OTLs with a certain type of task, they were very likely to do so with others. It should be noted that the correlation coefficient between AMR and PT was relatively low, 0.258. This correlation indicates that AMR with PT was somewhat independent compared to other pairs of OTLs.

Table 7. Correlation Coefficients between Latent Variables

	WP	PT	PMR	AMR
WP	1	0.572	0.387	0.396
PT		1	0.426	0.258
PMR			1	0.428
AMR				1

Note. All correlation coefficients are significant with $p < 0.01$

Figure 3 shows all SEM results based on the nested model. Table 8 reports the regression weights, which were all significant at the alpha level of 0.05. First, only the regression weight of AMR was negative (-0.174), the others were positive (0.132 for WP, 0.099 for PT, and 0.104 for PMR). The degree of this negative effects was also larger than others. IPC has much stronger relationship to mathematical literacy than EPC. It is remarkable that EPC is negatively related to mathematical literacy although it is not statistically significant at alpha 0.05. In addition, students' EPC was expected to increase by 0.091 when students had increase of 1 in their AMR. These findings about AMR indicate that encountering AMR frequently had negative influences on their mathematical literacy and positive influences on EPC simultaneously. However, OTL with PT was expected to increase mathematical literacy scores both directly and indirectly through IPC. When students increased 1 in their PT, students were expected to have increase in their IPC by 0.297. This value was remarkably large when considering that IPC was a psychological factor and that this result indicated possible impact of tasks introduced to students on their psychological perception. Because students' increase in IPC by 1 was expected to increase mathematical literacy by 0.358, the indirect effect of PK was $0.106 = 0.297 \times 0.358$ ($SE = 0.016, p < 0.001$). This value was similar with the direct effect of PT, 0.099 ($p = 0.004$).

Table 8. Results from the Structural Equation Modeling

Independent Variable	Dependent Variable	Coefficient	SE	z-value	p-value
PT	IPC	0.297	0.039	7.604	<0.001
AMR	EPC	0.091	0.043	2.120	0.034
IPC	Mathematical Literacy	0.358	0.030	11.970	< 0.001
WP	Mathematical Literacy	0.132	0.035	3.738	< 0.001
PT	Mathematical Literacy	0.099	0.044	2.259	0.024
PMR	Mathematical Literacy	0.104	0.036	2.914	0.004
AMR	Mathematical Literacy	-0.174	0.032	-5.386	< 0.001

This research studied the relationship among OTL, perceived control, and mathematical literacy. OTL itself is conceptualized as the frequency of engagement in four different mathematical tasks that was perceived by students. It is critical to think about the possible explanations for this finding, though verifying the speculated reasons is beyond the scope of this study. Here, we provide possible explanations that may be implemented by subsequent studies and, by extension, identify pathways for future research. First, we investigated the relationship between mathematical tasks as OTL and two types of perceived control: internal and external. On the one hand, students' OTL with procedural tasks was positively related to internal perceived control ($p < 0.001$); on the other hand, OTL with applied mathematics reasoning tasks was related to external perceived control ($p = 0.034$). Although these significant relationships were not necessarily causal, this suggests that students' experiences with different types of tasks in mathematics classrooms are one of the factors that shape students' perceptions of perceived locus of control.

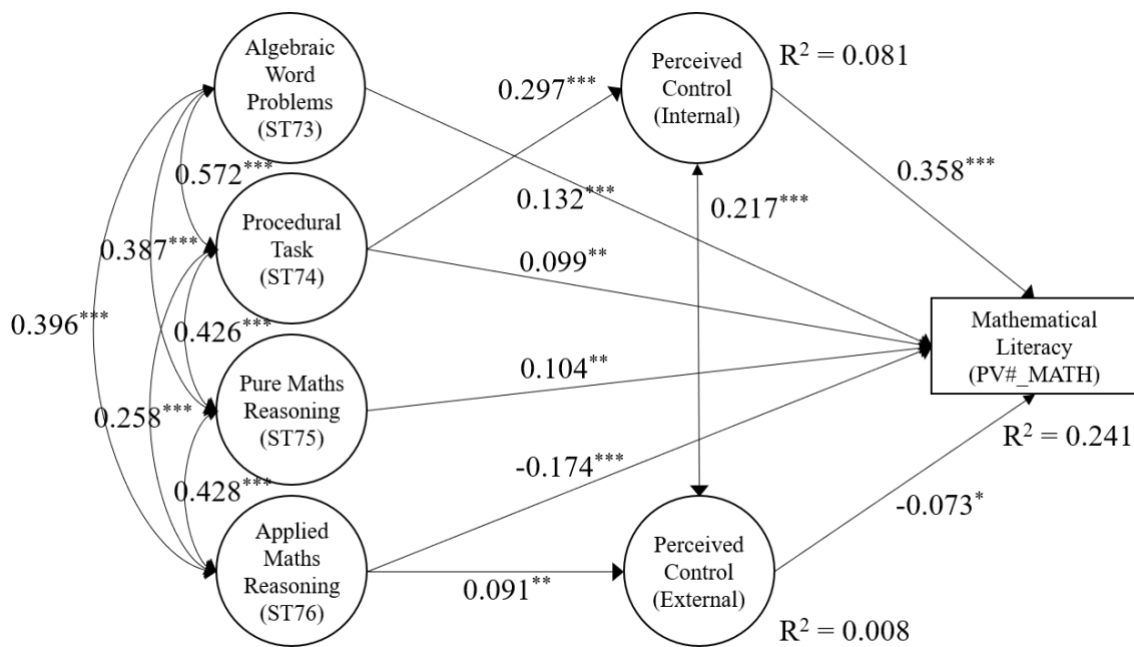


Figure 3. The SEM Results based on the Nested Model

When speculating several reasons for why perceived control and OTL diverged as observed, we put forward how students’ engagement in mathematical tasks can relate to variations in their teachers’ implementation of the tasks. Specifically, there could be larger variances in ways to implement of applied reasoning tasks (involving high cognitive demands) than in procedural tasks (involving low cognitive demands). In this sense, students’ OTL in applied mathematical reasoning tasks can also vary according to how teachers’ implement those tasks, which includes how the tasks are presented and how students’ learning is scaffolded by teachers. Since cognitive demand for tasks varies by the ways of teachers’ task presentation and scaffolding, students might perceive that their success/failure of the learning tasks is subject to teachers’ implementation and that their success/failure is out of their control. In contrast to the applied reasoning tasks, procedural tasks do not give as much space for implementation-variation, and hence, students’ positive/negative experiences with procedural tasks are perceived not to be contingent on how they are implemented by teachers. In this study, students who were frequently engaged in procedural tasks were more likely to think that their success/failure is under their own control, which resulted in a strongly positive relationship between OTL in procedural tasks and internal perceived control.

In this study, we also investigated the relationship between OTL via mathematical tasks and mathematical literacy scores. In constructing the model in Figure 1, we conjectured that all types of OTL would be positively related to mathematical literacy, though the degree of the relationship may vary. Particularly, we expected that applied mathematical reasoning tasks would have a stronger positive relationship to mathematical literacy than the other three task types. The PISA assessment for

mathematical literacy involved items that assessed capabilities in mathematical reasoning, and use of mathematical concepts and procedures in various real contexts. At the beginning of the study, we reasoned that many constituent capacities of mathematical literacy are utilized when students engage in applied mathematical reasoning tasks; however, the findings challenged this conjecture. Indeed, OTL via applied mathematical reasoning was negatively related to mathematical literacy although all other mathematical tasks (word problems, procedural tasks, and pure mathematics reasoning tasks) were positively related to mathematical literacy. This means that those students who were exposed more frequently to applied mathematical reasoning tasks were likely to have lower mathematical literacy scores, whereas those who were more frequently engaged in other types of tasks were likely to have higher mathematical literacy scores.

At this point, we will discuss why OTL via applied reasoning mathematical tasks had a uniquely different relationship with mathematical literacy compared to the other three tasks. We will highlight the specific cognitive processes that are required to successfully engage in each type of task, and how we can characterize such thinking processes. According to Hanna and Jahnke (2007), engagement in mathematical tasks requires students to undergo two reasoning processes: (1) making abductive inferences, such as “an action of *selection*,” to build the correct premise and (2) making deductive inferences between the premise and the conclusion (p. 149). As presented in the example tasks (Tables 3 and 4), pure mathematical reasoning, word problems, and procedural tasks offer the correct premises directly to the students, thereby placing less emphasis on abductive inferences. For applied mathematical reasoning, abductive inference is one of the most important elements when formulating a given real situation and building premises.

Also, theoretically, mathematical literacy is defined as a combination of both abductive and deductive inferences, which includes the abilities of formulating premises, employing mathematical concepts and ideas, and interpreting solutions in real situations. However, the mathematical literacy that was measured in the PISA 2012 might not have captured both abductive and deductive inferences in a balanced way. Even though the OECD reported that PISA mathematics assessments are improved by using computer-based delivery formats, it is still difficult to evaluate students’ inductive/abductive reasoning skills with mathematics test items. Students are asked to answer multiple-choices questions in the assessments, and this type of assessment does not well reflect students’ processes of searching for reasons. Thus, in standardized test settings, often with time limits, students focus more on finding a correct answer from the information that is presented in the test problems, rather than exploring and formulating a real-word problem situation.

The questionnaire in PISA 2012 asked students how often they encountered each type of tasks during their mathematics classes. The negative relationship between applied reasoning mathematical tasks and mathematical literacy encouraged us to rethink how the frequency of OTL via applied reasoning tasks can affect mathematics learning. Departing from the idea of ‘the more, the better,’ we speculated that the implementation of applied reasoning tasks has much to do with how they are

implemented, as opposed to how often they are implemented. Considering that the PISA student survey was about how often students encountered each type of tasks, it is possible that other important aspects of OTL, such as quality and process of OTL, were not taken into consideration. The varying quality of students' OTL of applied mathematical reasoning could be another reason for the negative relationship between applied mathematics reasoning tasks and mathematical literacy. Particularly, some researchers (e.g., Boston & Smith, 2009) have argued that teachers tend to reduce the original cognitive demands of mathematical tasks when implementing them. This means that the result is probably due to the ways in which those reasoning tasks were implemented. It can be challenging for teachers to scaffold students' learning process carefully and successfully by engaging them in applied mathematics reasoning tasks. As such, this may inhibit students from fully taking advantage of OTL via applied mathematical reasoning.

The findings of this study support that allocating more learning time to applied reasoning task is not necessarily beneficial to, or does not guarantee, overall mathematics learning. However, our attempts made so far to explain the *negative* relation between applied reasoning tasks and mathematical literacy still may not seem to be sufficient. Thus, future research on teachers' implementation of applied mathematics reasoning tasks in classrooms should be followed to validate and explain the negative relationship between applied mathematics reasoning task and mathematical literacy.

Another research question of our study was on the role of perceived control in the relationship between OTL with mathematical tasks and mathematical literacy. The results showed that engagement in OTL with procedural tasks is likely to influence mathematical literacy directly and indirectly through internal perceived control. Particularly, the effect of engagement in procedural tasks on mathematical literacy is even greater when the indirect effect through internal perceived control is taken into consideration. Considering the strong positive relationship between internal perceived control and mathematical literacy, students are likely to have high mathematical literacy scores when they believe that being successful in mathematics is under their control. To synthesize, OTL through procedural tasks is likely to promote students' internal perceived control, and in turn, this can have an effect on better mathematical literacy. Though this may suggest the merit of engaging students in procedural tasks in relation with students' perceived control, we are not to argue that procedural tasks should be offered more in mathematics classrooms than other types of tasks. As Yeo (2007) argued, students need to have a variety of OTL, from procedural tasks to mathematizing tasks, and teachers need to be cognizant about different OTL that is afforded by various types of tasks. This is specifically because OTL through different types of tasks may have varying effect on cognitive and non-cognitive processes during mathematics learning, as shown in our study.

CONCLUSION

This research showed that students can improve their mathematical literacy by engaging in various types of tasks from procedural tasks, word problems, to pure and applied mathematics reasoning

tasks. Opposite from our expectation, the results showed that students are likely to have lower mathematical literacy when they have encountered applied mathematics reasoning tasks more frequently. In addition to the discussion of the results, we suggest future studies about how different types of tasks are implemented in classrooms, how such implementation influence students' perceived control, and how students perform on tests based on their classroom experiences.

We suggest several implications based on the findings in this research: First, teachers and curricular developers need to implement various tasks considering their relation to students' perceived control and mathematical literacy. However, it is important to recognize that students' frequent engagement in certain tasks could have unexpected influences on their mathematical literacy, especially when they are not appropriately facilitated. Particularly, Korean students' OTL with applied mathematics reasoning tasks had negative relationship with their achievement. These findings call for more investigation on how to implement such tasks appropriately.

Secondly, it is critical to consider how teachers select OTL with different types of tasks to offer in their teaching practices. When educators emphasize tasks with high cognitive demands, it is sometimes misunderstood that tasks with low cognitive demand are less beneficial to students' higher order thinking processes in mathematics learning. Even worse, tasks with low cognitive demand are considered as something that teachers should avoid. However, our findings show that procedure tasks can help students believe that their success is under their own control, which could lead to better mathematics learning behaviors and higher mathematical literacy through appropriate scaffoldings for other types and levels of mathematical tasks. Then, students will be able to engage in different types of mathematical thinking and perceive that they can succeed in mathematics learning with their own effort.

We did not make direct relation between educational contexts of Korea and the results of this study in our interpretation of the results, which should be noted when attempting to generalize the findings to different educational systems. Moreover, we highlight that mathematical literacy defined and measured by the PISA could be different from achievement measures in other mathematics assessments. Therefore, replication studies using other assessments tools or in other educational contexts are needed.

REFERENCES

- Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record*, 112(3), 723-746. https://www.ncaase.com/docs/Abedi_OTL_TRC_2010.pdf
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior 1. *Journal of Applied Social Psychology*, 32(4), 665-683. <https://doi.org/10.1111/j.1559-1816.2002.tb00236.x>
- Asparouhov, T., & Muthen, B. (2006). *Multilevel modelling of complex survey data*. Paper presented at the Joint Statistical Meeting, ASA Section on Survey Research Methods, Seattle, WA.

- Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education*, 8, 499–536. <https://doi.org/10.1007/s10857-006-6585-3>
- Bailin, S., & Siegler, H. (2003). Critical thinking. In N. Blake, P. Smeyers, R. Smith, & P. Standish (Eds.), *The Blackwell Guide to the Philosophy of Education* (pp. 181–193). Malden, MA: Blackwell Publishing.
- Barnard-Brak, L., Lan, W. Y., & Yang, Z. (2018). Differences in mathematics achievement according to opportunity to learn: A 4pL item response theory examination. *Studies in Educational Evaluation*, 56, 1–7. <https://doi.org/10.1016/j.stueduc.2017.11.002>
- Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher*, 31(8), 18–20. <https://doi.org/10.3102/0013189X031008018>
- Bornemann, B., Foth, M., Horn, J., Ries, J., Warmuth, E., Wartenburger, I., & van der Meer, E. (2010). Mathematical cognition: Individual differences in resource allocation. *ZDM*, 42(6), 555–567. <https://doi.org/10.1007/s11858-010-0253-x>
- Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, 40(2), 119–156. <https://doi.org/10.2307/40539329>
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum.
- Carroll, J. B. (1963). A model for school learning. *Teachers College Record*, 64, 723–733. <http://www.tcrecord.org/books/pdf.asp?ContentID=2839>
- Chaney, B., Jocelyn, L., Levine, D., Mule, T., Rizzo, L., Roey, S., . . . Williams, T. (2001). *User's guide for the Third International Mathematics and Science Study (TIMSS)* (NCES 2001-065). Washington, DC: National Center for Education Statistics.
- Cogan, L. S., & Schmidt, W. H. (2015). The concept of opportunity to learn (OTL) in international comparisons of education. In K. Stacey & R. Turner (Eds.), *Assessing Mathematical Literacy* (pp. 207-216). Springer. https://doi.org/10.1007/978-3-319-10121-7_1
- d'Ailly, H. (2003). Children's autonomy and perceived control in learning: A model of motivation and achievement in Taiwan. *Journal of Educational Psychology*, 95(1), 84-96. <https://doi.org/10.1037/0022-0663.95.1.84>
- Elliott, S., & Bartlett, B. (2016). Opportunity to Learn. In P. Nathan (Ed.), *Oxford Handbook of Education Online*. New York: Oxford Press. <https://doi.org/10.1093/oxfordhb/9780199935291.013.70>
- Foy, P., Brossman, B., & Galia, J. (Eds.). (2012). *Scaling the TIMSS and PIRLS 2011 achievement data*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Hammer, D. (2000). Student resources for learning introductory physics. *American Journal of Physics*, 68(S1), S52–S59. <https://doi.org/10.1119/1.19520>
- Hanna, G., & Jahnke, H. N. (2007). Proving and modelling. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education: The 14th ICMI Study* (New ICMI Study Series ed., pp. 145–152). New York, NY: Springer.

- Hayduk, L., & Littvay, L. (2012). Should researchers use single indicators, best indicators, or multiple indicators in structural equation models? *BMC Medical Research Methodology*, *12*(159), 1-17. <https://doi.org/10.1186/1471-2288-12-159>
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524-549. <https://doi.org/10.5951/jresmetheduc.28.5.0524>
- Hrbáčková, K., Hladík, J., & Vávrová, S. (2012). The relationship between locus of control, metacognition, and academic success. *Procedia - Social and Behavioral Sciences*, *69*, 1805-1811. <https://doi.org/10.1016/j.sbspro.2012.12.130>
- Hwang, J., Choi, K. M., & Hand, B. (2020). Examining domain-general use of reasoning across science and mathematics through performance on standardized assessments. *Canadian Journal of Science, Mathematics and Technology Education*, *20*, 521-537. <https://doi.org/10.1007/s42330-020-00108-4>
- Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a Complex System: Conceptual and Methodological Implications. *Educational Researcher*, *48*(2), 112–119. <https://doi.org/10.3102/0013189x19826958>
- Lane, S. (2004). Validity of high-stakes assessment: Are students engaged in complex thinking? *Educational Measurement: Issues and Practice*, *23*(3), 6–14. <https://doi.org/10.1111/j.1745-3992.2004.tb00160.x>
- Lipnevich, A. A., Preckel, F., & Krumm, S. (2016). Mathematics attitudes and their unique contribution to achievement: Going over and above cognitive ability and personality. *Learning and Individual Differences*, *47*, 70–79. <https://doi.org/10.1016/j.lindif.2015.12.027>
- Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When “good” indicators are bad and “bad” indicators are good. *Psychological Methods*, *4*(2), 192-211. <https://doi.org/10.1037/1082-989X.4.2.192>
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.
- Martin, L., & Gourley-Delaney, P. (2014). Students' images of mathematics. *Instructional Science*, *42*, 595–614. <https://doi.org/10.1007/s11251-013-9293-2>
- McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575–596). New York, NY: Macmillan.
- McNabb, T. (2003). Motivational issues: Potential to performance. In N. Colangelo & G. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 417–423). Boston, MA: Allyn & Bacon.
- Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I., & Chevrier, M. (2015). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology*, *42*, 172–185. <https://doi.org/10.1016/j.cedpsych.2015.06.003>
- Mulnix, J. W. (2012). Thinking critically about critical thinking. *Educational Philosophy and Theory*, *44*(5), 464–479. <https://doi.org/10.1111/j.1469-5812.2010.00673.x>

- Murayama, K., Pekrun, R., Lichtenfeld, S., & vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement: The unique contributions of motivation and cognitive strategies. *Child Development, 84*(4), 1475–1490. <https://doi.org/10.1111/cdev.12036>
- Obserski, D. (2016). *lavaan.survey*: Complex survey structural equation modeling (SEM). R package version 1.1.3.1
- OECD. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris, France: OECD Publishing.
- OECD. (2014). *PISA 2012 technical report*. Paris, France: OECD Publishing.
- OECD. (2017). *PISA 2015 technical report*. Paris, France: OECD Publishing.
- OECD. (n.d.). OECD Programme for International Student Assessment 2012: Student questionnaire - form A. Retrieved from http://www.oecd.org/pisa/pisaproducts/PISA12_StQ_FORM_A_ENG.pdf
- Ottmar, E. R., Decker, L. E., Cameron, C. E., Curby, T. W., & Rimm-Kaufman, S. E. (2014). Classroom instructional quality, exposure to mathematics instruction and mathematics achievement in fifth grade. *Learning Environments Research, 17*(2), 243-262. <https://doi.org/10.1007/s10984-013-9146-6>
- Patrick, B. C., Skinner, E. A., & Connell, J. P. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology, 65*(4), 781–791. <https://doi.org/10.1037/0022-3514.65.4.781>
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1–28. <https://doi.org/10.1037/h0092976>
- Rotter, J. B., & Mulry, R. C. (1965). Internal versus external control of reinforcement and decision time. *Journal of Personality and Social Psychology, 2*(4), 598–604. <https://doi.org/10.1037/h0022473>
- Schmidt, W. H. (1992). The distribution of instructional time to mathematical content: One aspect of opportunity to learn. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. 129–145). New York, NY: Pergamon.
- Schmidt, W. H., Zoido, P., & Cogan, L. (2014). *Schooling matters: Opportunity to learn in PISA 2012*. Paris, France: OECD Publishing.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). New York, NY: MacMillan.
- Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), *Mathematical Thinking and Problem Solving* (pp. 53–69). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99*(6), 323–338. <https://doi.org/10.3200/JOER.99.6.323-338>
- Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist, 19*(1), 48–58. <https://doi.org/10.1080/00461528409529281>

- Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, 82(1), 22–32. <https://doi.org/10.1037/0022-0663.82.1.22>
- Skinner, E. A., Zimmer-Gembeck, M. J., Connell, J. P., Eccles, J. S., & Wellborn, J. G. (1998). Individual differences and the development of perceived control. *Monographs of the Society for Research in Child Development*, 63(2/3), 1–231. <https://doi.org/10.2307/1166220>
- Son, J. W. (2012). A cross-national comparison of reform curricula in Korea and the US in terms of cognitive complexity: the case of fraction addition and subtraction. *ZDM*, 44(2), 161-174. <https://doi.org/10.1007/s11858-012-0386-1>
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, 33(2), 455–488. <https://doi.org/10.3102/00028312033002455>
- Stevens, F. I., & Grymes, J. (1993). *Opportunity to learn: Issues of equity for poor and minority students*. Washington, DC: National Center for Education Statistics.
- Stevens, F. (1996, April). *The need to expand the opportunity to learn conceptual framework: Should students, parents and school resources be included?* Paper presented at the annual meeting of the American Educational Research Association. New York, NY.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics*. Boston: Pearson.
- Törnroos, J. (2005). Mathematics textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation*, 31(4), 315–327. <https://doi.org/10.1016/j.stueduc.2005.11.005>
- von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series. Issues and Methodologies in Large-Scale Assessments*, 2, 9–36.
- Watson, A. (2003). Opportunities to learn mathematics. *Mathematics education research: Innovation, networking, opportunity*, 29–38.
- Yeo, J. B. W. (2007). *Mathematical tasks: Clarification, classification and choice of suitable tasks for different types of learning and assessment*. Technical Report Me 2007-01. National Institute of Education, Nanyang Technological University, Singapore.