Exploration of the linear and nonlinear relationships between learning strategies and mathematics achievement in South Korea using the nominal response model: PISA 2012

Candidate number: 3002

Master of Science in Assessment, Measurement, and Evaluation

30 Credit Points

Center of Educational Measurement in Oslo

Faculty of Educational Science, University of Oslo

May 2022

Table of Contents

Popular Abstract

Learning strategies (e.g., memorization, elaboration, and metacognitive strategies) have been found to be linked with mathematics achievement. Recent studies have found that high-achieving Asian students use mixed learning strategies, primarily metacognitive strategies, rather than relying on rote learning. There is a lack of research exploring how students use learning strategies in a specific East Asian education system, and few studies have applied the nominal response model (NRM) to scoring learning strategies in PISA 2012. Thus, this study fills this research gap by focusing on South Korean students' mathematics achievement in relation to their use of learning strategies. In the analysis, we use the NRM to achieve this purpose. We also explore the linear and nonlinear relationship between learning strategy use and mathematics achievement. Finally, we discuss these findings with respect to the South Korean education system, how Korean students use metacognitive strategies with memorization strategies, and the implications of these findings.

Acknowledgments

I would like to express my deep and sincere appreciation to my supervisors, Chia-Wen and Yi-Jhen. Even though I couldn’t be on campus because of the pandemic, they let me have online meetings, gave me online lectures, and responded to my questions. They encouraged me to keep writing throughout the entire journey and eventually complete this thesis. I really appreciate their dedicated support.

It was unusual beginning this program with the pandemic. We spent one year learning online without seeing teachers and classmates in person. Even so, dedicated support by lecturers and staff at CEMO motivated me to study and helped me find joy in learning. My classmates also supported me mentally and opened my eyes to diverse perspectives. I am enormously grateful to all the amazing members at CEMO, the wonderful teachers, and my dear colleagues for being part of this master’s program journey with me.

Last, but not least, I really appreciate my family for sticking with me. I admit that there are times when I would have just given up, but my family gave me the strength to carry on. My parents have always given me the confidence to know that I can manage. My brother made me laugh and take breathers from my thesis. Many thanks to my family for their support and care.

Abstract

Learning strategies are acknowledged as important predictors of mathematical achievement. Recent studies have found that Asian students use combined learning strategies, primarily including metacognitive strategies, rather than rote memorization. To the best of the authors' knowledge, there is only one prior study including South Korea in investigations of the relationship between learning strategies and mathematics achievement in PISA 2012. In that study, students were classified into groups using specific learning strategies, and their mathematics achievements were compared. There are two research gaps: (1) previous studies insufficiently explored how students use learning strategies in the South Korean education system, and (2) there is little research applying the nominal response model (NRM) to explore the association between learning strategy use and mathematics achievement in PISA 2012. Thus, the present study explores to what extent the NRM fits the data compared to the generalized partial credit model (GPCM). We created a learning strategy score from the NRM for South Korean students in PISA 2012 (N = 5,033). Then, using correlation analysis and quadratic regression analysis, we identified linear and nonlinear relations between learning strategy scores from the NRM and mathematics achievement. The findings indicated that (1) NRM was a better fit to create learning strategy scores than the GPCM, (2) the average correlation coefficient between learning strategy score and mathematics achievement was 0.18 ($p < .05$), and (3) for the curvilinear relationship between learning strategy score and mathematics achievement, the standardized quadratic coefficient was -0.090 ($p < .001$). Overall, the NRM represents an appropriate model for explaining the relationship between learning strategy and mathematical achievement. Additionally, high-performing South Korean students tend to primarily use metacognitive strategies with memorization. The negative quadratic coefficient captured the limited effect of the primary use of metacognitive strategies with memorization. The implications for the South Korean education system are discussed.

Keywords: learning strategy, mathematics achievement, nominal response model, South Korea, PISA 2012.

Introduction

Learning strategies can be defined as behaviors and thoughts in which a learner engages, or which are intended to influence the learner's encoding process. Like goal-oriented activities, learning strategies are used for acquiring, organizing, or transforming information, as well as for reflecting upon and guiding the learning process (Weinstein & Mayer, 1986). These strategies have also been acknowledged as important predictors of academic achievement (Hong et al., 2006). The Program for International Student Assessment (PISA), an international large-scale assessment developed by the Organization for Economic Co-operation and Development (OECD), measures 15-year-old students' use of learning strategies on an ordinal scale (OECD, 2012). The PISA uses three learning strategies—memorization, elaboration, and metacognitive strategies—and defines them as follows: memorization involves learning key terms and repeated learning of materials; elaboration includes making connections to related areas and thinking about alternative solutions; and metacognition involves planning, monitoring, and regulation (OECD, 2005; Zimmerman, 2001). Since PISA 2012, they have turned the ordinal scale of learning strategy items into a nominal scale.

By analyzing learning strategy items in PISA, research has demonstrated that these mathematics learning strategies are associated with students' mathematics achievement (Areepattamannil & Caleon, 2013; Kiliç et al., 2012; Lin & Tai, 2015; Wu et al., 2020). For instance, Areepattamannil and Caleon (2013) found that in East Asian education systems, including Shanghai-China, Hong Kong-China, Korea, and Singapore, memorization strategies were negatively associated with mathematics achievement, and the magnitude of the negative correlation differed among the countries. Some studies have also explored the use of learning strategies without using PISA data. Elaboration strategies and metacognitive strategies have been found to be positively correlated with learning achievement across 34 countries (Chiu et al., 2007), including Germany (Glogger et al., 2012; Murayama et al., 2013), Hong Kong (McInerney et al., 2012), and Sweden (Rosander & Bäckström, 2012). Memorization is generally considered less effective than other learning strategies (e.g., elaboration and metacognitive strategies; McInerney, 2012).

Among studies that have investigated learning strategies for mathematics in the East Asian educational system (Lin & Tai, 2015; Liu et al., 2019; Wu et al., 2020), only Wu et al. (2020) included South Korea in the East Asian data of PISA 2012. In fact, very limited research has examined the relationship between learning strategies and mathematics achievement in the South Korean education system (e.g., with CSAT and exam-driven culture). Furthermore, considering that the magnitude of correlation coefficients between learning strategy use and mathematics achievement is different across countries (Areepattamannil & Caleon, 2013; Wu et al., 2020), research is necessary to interpret and discuss a specific education system to better understand learning strategy use and mathematics achievement.

Thus, this study focuses on the South Korean education system for two reasons. First, Korean students are encouraged to use memorization strategies. The College Scholastic Aptitude Test (CSAT), a university entrance exam in South Korea, is considered the sole determinant of which university a student can attend (Blazer, 2012). The CSAT has led students to rely on memorization strategies to learn test-taking skills and improve their ability to solve multiple-choice questions in a limited amount of time (Kim, 2004). Second, South Korea consistently shows high levels of academic achievement in international large-scale assessments in the domain of mathematics, indicating that South Korea a top-tier country (Choi et al., 2019; Park, 2004). For instance, South Koreans ranked 4th in PISA 2009, 5th in PISA 2012, and 7th in PISA 2015 on mathematics performance, and 5th in science, reading, and mathematics in PISA 2018.

In the present study, we use the nominal response model (NRM) to score Korean students' learning strategies in PISA 2012 and examine the association between learning strategy use and mathematics achievement. The NRM is an item response theory (IRT) model for modeling the probability of responses to items with nominal categories (i.e., unordered responses; Zu & Kyllonen, 2020) as a type of learning strategy item in PISA 2012. To explore the extent to which the NRM fits the data, we compare it to the generalized partial credit model (GPCM), which assumes the order of categories (i.e., the ordinal relationship between learning strategies) in IRT modeling. The advantage of applying IRT in scoring compared to the sum of scores is that a parametric model can estimate the uncertainty of point estimates (i.e., standard errors), which can be taken into account in the subsequent analysis of the relationship between learning strategy and mathematics achievement using the plausible values approach.

This study aims to examine and explore the relationship between learning strategy use and mathematics achievement in the South Korean education system. We explore (1) the extent to which the NRM fits the data and (2) the linear and nonlinear relationship between learning strategy use and mathematics achievement in South Korea. To achieve the first goal, we compared the NRM to the GPCM and expected that the NRM would fit the data better than the GPCM because of the nominal nature of learning strategies. The second purpose was achieved by conducting a correlation analysis between the Korean students' learning strategy scores from the NRM and mathematics scores, as well as a correlation analysis between raw scores of single strategies and mathematics scores in PISA 2012. In addition, to examine the nonlinear relationships between these variables, we used quadratic regression analysis.

The article is organized as follows. We review the literature on learning strategies using self-regulated learning theory, the relationships between learning strategy use and mathematics achievement, and learning strategies used in the East Asian and South Korean contexts. We briefly introduce two IRT models, the NRM and GPCM, then generate the two research questions according to the research gaps. The methodology section describes the South Korean sample and the measures of learning strategies and mathematics achievement in PISA 2012. In the statistical analysis, model comparison between NRM and GPCM is performed to answer the first research question. Then, correlation analyses and nonlinear regression analysis between learning strategy use and mathematics achievement are conducted to answer the second research question. Finally, we discuss our findings and elaborate on the reasons for them.

Literature Review

Learning Strategies in Self-regulated Learning Theory

The self-regulated learning (SRL) process was introduced by Zimmerman (2001) to describe how students regulate their own learning processes, including learning strategies, motivation, and behavior. According to SRL, a self-oriented feedback loop occurs during learning (Carver & Scheier, 1981; Zimmerman, 1990). In this cyclical loop process, students monitor the effectiveness of their learning strategies and respond to this feedback in several ways, such as replacing one learning strategy with another to achieve more desirable results (Zimmerman, 2001).

In the SRL process, students are regarded as self-regulated learners to the degree that they are metacognitively, motivationally, and behaviorally active participants in their own learning processes (Zimmerman, 1986). These students self-generate thoughts, feelings, and actions as their learning goals. SRL includes students' metacognitive strategies for planning, monitoring, and modifying their cognition (Campione et al., 1984; Corno, 1986; Zimmerman & Pons, 1986, 1988), and the actual cognitive strategies that students use to learn, remember, and understand the material (Corno & Mandinach, 2009; Zimmerman & Pons, 1986, 1988).

These different cognitive strategies, such as rehearsal, elaboration, and organizational strategies, have been found to foster active cognitive engagement in learning and to result in higher levels of achievement (Weinstein & Mayer, 1986).

Learning Strategies and Achievement

SRL theory has prompted many empirical studies to define different types of learning strategies and demonstrate their efficiency (Dent & Koenka, 2016; Pintrich & Groot, 1990; Zimmerman & Pons, 1986). Although various classifications of learning strategies have been suggested (Kember et al., 2004; Lee & Shute, 2010; Marton & Säljö, 1976; Weinstein & Mayer, 1986; Zimmerman & Pons, 1986), many studies have followed the concept of Weinstein and Mayer’s (1986) framework to define cognitive and metacognitive strategies.

Cognitive strategies (e.g., memorization and elaboration) refer to mental procedures related to learning, storing, organizing, summarizing, and understanding information by relating it to new and prior knowledge (Weinstein & Mayer, 1986; Zimmerman & Pons, 1986). While learning mathematics, students may recall formulas, summarize a mathematical concept that they have absorbed, or connect a mathematics concept to their actual experiences (Wu et al., 2020). Metacognitive strategies (i.e., control strategies in PISA) refer to supervising, controlling, and regulating cognitive activities (Weinstein & Mayer, 1986; Zimmerman & Pons, 1986). During mathematics learning, metacognitively aware students may devise plans to solve the next mathematics tasks, review their own understanding of the concepts learned, ask for help, and assess their own learning strategies to improve performance (OECD, 2013).

PISA uses self-reported learning strategy items in the mathematics domain (OECD, 2005), following Weinstein and Mayer’s (1986) concept of learning strategies. We discuss memorization, elaboration, and metacognitive strategies in the following subsections.

Memorization strategies and mathematics achievement

Memorizing factual knowledge might be useful in the introductory stage of acquiring mathematics knowledge (Dinsmore & Alexander, 2016), but exclusive use of memorization as a strategy does not generally lead to improvements in complex problem solving or advanced logical skills (Biggs, 1993; Liu et al., 2019; Marton & Säljö, 1976; McInerney et al., 2012). For example, Liu et al. (2019) suggested that Chinese students who use the memorization strategy with the other learning strategies (e.g., elaboration and metacognition) perform better than those who use only the memorization strategy in mathematics.

Educational studies have investigated the impact of memorization strategies on mathematics achievement (Areepattamannil & Caleon, 2013; Kiliç et al.,2012; Pintrich & Groot, 1990). In general, the exclusive use of memorization is negatively correlated with mathematics achievement. Pintrich and Groot (1990) found that the use of memorization without metacognitive strategies was not conducive to academic performance. Similarly, Kiliç et al. (2012) found that memorization had a negative effect on learning in Turkey and its neighboring countries, and Areepattamannil and Caleon (2013) concluded that memorization strategies were negatively associated with mathematics achievement in four East Asian education systems: Shanghai-China, Korea, Hong Kong-China, and Singapore. While studies have suggested that mixed use of learning strategies including memorization may lead to better academic performance than the use of a single strategy (Dent & Koenka, 2016; Wu et al., 2020), educational researchers tend to hold negative views of using memorization only.

Elaboration strategies and mathematics achievement

Elaboration is defined as mental processes and behaviors that involve integrating information from different sources to create meaningful interpretations, relate new concepts to prior knowledge, and summarize material into one's own words (Pintrich & Groot, 1990; Trigwell & Prosser, 1991; Walker et al., 2006; Wolters, 2004). Elaboration can occur during self-study, discussions, note-taking, or answering questions (Pires et al., 2020).

Elaboration strategies that deepen understanding of knowledge and skills lead to high-quality learning outcomes, whereas students who use a surface approach (e.g., rehearsal or memorization; Ramsden, 1988) are more likely to achieve lower-quality outcomes (Marton & Säljö, 1976; Prosser & Millar, 1989). Studies have found that elaboration strategies lead to a positive effect on student learning, including mathematics (Donker et al., 2014; Murayama et al., 2013). One meta-analysis study (Donker et al., 2014) found that elaboration was the only sub-strategy that demonstrated a significantly positive relationship with mathematics achievement among a variety of sub-strategies. The longitudinal study (Murayama et al., 2013) suggested that growth in students' mathematics achievement was positively predicted by deep learning strategies from Grades 5 through 10, and was negatively predicted by surface learning strategies (Ramsden, 1988).

In contrast, the relationship between elaboration strategy and mathematics achievement did not demonstrate a consistent pattern of results across different educational systems in a study by Chiu et al. (2007), who found that these strategies were not linked to achievement in any domain or culture. Liu et al. (2009) indicated that elaboration strategy use by Chinese eighth-grade students showed either a positive or negative relationship with mathematics achievement, depending on unique Chinese demographic variables (e.g., only-child families and residential locations). Thus, the effect of elaboration strategies varied between countries.

Metacognitive strategies and mathematics achievement

According to SRL theory, self-regulated learners are able to monitor the efficiency of their learning strategies and change one learning strategy to another to achieve their goals. This is referred to as a metacognitive strategy (Zimmerman, 2001). Several researchers have shown that metacognition plays an important role in mathematics success (Borkowski & Thorpe, 1994; De Clercq et al., 2000; Schoenfeld, 2016). Artz and Armour-Thomas (2009) found that the main reason for students' failures in mathematical problem solving was that they were not able to monitor their own mental procedures.

Many empirical studies have demonstrated the effectiveness of metacognitive strategies for improving students' mathematics performance (Areepattamannil & Caleon, 2013; Desoete et al., 2001; Dignath & Büttner, 2008; Perels et al., 2009). Desoete et al. (2001) indicated that metacognitive knowledge and skills accounted for 37% of achievement in mathematical problem solving. Dignath and Büttner (2008) demonstrated a stronger relationship between metacognitive strategies and mathematics than with other subjects. Perels et al. (2009) investigated the effects of training metacognitive strategies (i.e., self-regulative strategies) on mathematical achievement for Grade 6 students in Germany. The students of the experimental group whose teacher taught mathematics topics combined with metacognitive strategies (i.e., self-regulative strategies) showed more improvement in their mathematics skills in a pre/posttest comparison than the control group whose teacher taught merely mathematical topics. Areepattamannil and Caleon (2013) found that metacognitive strategies were positively associated with mathematics achievement in four East Asian education systems: Shanghai-China, Korea, Hong Kong-China, and Singapore. In addition, Wu et al.'s (2020) findings showed that the combined use of metacognitive and elaboration strategies was the most effective way for mathematics achievement in most East Asian countries, followed by the mixed use of metacognitive and memorization strategies.

East Asian Students' Learning Strategy Use

In recent decades, Western educators have explored the reasons for East Asian students' high mathematics performance. They believed that East Asian students relied on memorization, but these students performed better on international large-scale assessment (ILSA) than Western students (Biggs, 1998; Leung, 2014). However, several studies have found that East Asian students do not depend on a single strategy like memorization, but instead use mixed learning strategies (Lin & Tai, 2015; Liu et al., 2019; Wu et al., 2020). According to Wu et al. (2020), most East Asian students use multiple learning strategies for learning mathematics, and students who use both metacognitive and elaboration strategies achieve the highest scores on the mathematics exam, followed by those who use metacognitive and memorization strategies. Several studies have also shown that memorization does not necessarily imply rote learning without understanding (Biggs, 1998; Kember, 2016; Leung, 2014). For instance, as an application of memorization strategy, continuous practice with increasing variation could help learners understand new material (Hess & Azuma, 1991; Marton & Booth, 1997). Thus, the use of memorization strategy does not always mean rote learning and East Asian students does not rely entirely on memorization strategy.

Learning Strategies in the South Korean Context

The South Korean education system introduced the College Scholastic Aptitude Test (CSAT) in 1994 to encourage students to develop high-level thinking abilities rather than fragmented short-term memorization. However, CSAT was criticized for triggering a different kind of memorization because of its multiple-choice formats and for causing repetition of problem-solving exercises in test subjects, including mathematics (Kim, 2004). Students were intent on learning test-taking skills that would ensure their ability to solve these multiple-choice questions in a limited amount of time (Kim, 2004). As CSAT has become an essential determinant of which university a student can attend, South Koreans have expressed concern about whether students rely on rote learning only to get high scores on the exam (Blazer, 2012; Li, 2011).

A previous study (Wu et al., 2020) explored the relationship between learning strategies and mathematics achievement for South Korean students as well as the other education systems in East Asia (e.g., Hong Kong, Japan, Korea, Shanghai, Singapore, Taiwan, and Macau) using latent class analysis. According to Wu et al. (2020), the largest percentage (65%) of South Korean students primarily used metacognitive strategies with memorization strategies (Class 2). Only 14.2% of South Korean students primarily used metacognitive strategies with elaboration (Class 4). Class 4 was found to have the best performance among the classes, followed by Class 2. Although Wu et al. (2020) investigated South Korean students' learning strategy use, the study lacked discussion about the Korean education system.

Nominal Response Model

The nominal response model (Darrel Bock, 1972) is designed for items with nominal categories (Thissen et al., 2010). Nominal categories imply there is no assumption that Category 2 indicates higher ability than Category 1 (Zu & Kyllonen, 2020). In other words, the NRM does not assume that using a metacognitive strategy is better than using a memorization strategy in response to items in PISA 2012. It enables partial credit for different option selections and allows for differential item weights and varying category discriminations. The NRM is expressed as:

$$P(X_i = k|\theta) = \frac{e^{a_{ik}\theta + c_{ik}}}{\sum_{j=1}^{m_i} e^{a_{ij}\theta + c_{ij}}} \quad (1)$$

where $a_{ik}$and c_{ik} are the category slope and category intercept parameters, respectively, for the k^{th} category of item i. In this equation, the expression on the right gives the probability that a person with trait-level θ selects response category k ($k = 1, 2, 3, \dots m_i$) on item i.

Within an item, the order for the response categories with respect to latent ability is determined by the value of the a_{ik}s. Within item i, response k indicates higher θ than response q if and only if $a_{ik} > a_{iq}$ (Thissen et al., 2010). The category intercept parameters (c_{ik}) reflect the relative frequency of choosing that category, where the larger c_{ik} (intercept parameter) represents a greater relative frequency for option k (Zu & Kyllonen, 2020).

Generalized Partial Credit Model

The GPCM can be seen as a generalization of the dichotomous 2PL model for handling polytomous data and a constrained version of the NRM. The GPCM requires responses to be ordered from best to worst with respect to latent ability, which could be accomplished through prior knowledge, expert ratings, in- or out-of-sample response popularity, or other means (e.g., the a_{ik} values from the NRM analysis; see Equation 1; Zu & Kyllonen, 2020). In other words, in the GPCM, the options are coded [Memorization = 1, Elaboration = 2, Metacognitive = 3] based on prior knowledge (Biggs, 1987; OECD, 2014; Weinstein & Mayer, 1986; Zimmerman & Pons, 1986). This implies that students who have high learning strategy scores tend to use metacognitive strategies, and students who have low learning strategy scores tend to use memorization strategies. With the prior ordering of the response categories, the GPCM is the NRM with the constraint that the degree of discrimination between adjacent categories is the same for all adjacent categories in an item. Due to these constraints, category slopes within an item can be represented by one item slope parameter. An expression of the GPCM is the NRM shown in Equation 2, with constraints:

$$a_{ik} = a_i(k-1) \tag{2}$$

where a_i is the slope parameter for item i (Zu & Kyllonen, 2020). The number of parameters for item i under the GPCM is the number of response categories, m_i. In other words, within item i, category slopes are all the same ($a_{i1} = a_{i2} = a_{i3} = a_i$).

Aims of the Present Study

We summarize two research gaps: (1) There are few published studies applying NRM to examine learning strategy use and mathematics achievement in PISA 2012; and (2) There is little research exploring the relationship between these variables in the South Korean education system. This study will address these gaps by first exploring to what extent NRM fits learning strategy data in PISA 2012 compared to GPCM, and second, by investigating how South Korean students' learning strategies are correlated to mathematics achievement.

The present study seeks to answer the following two research questions:

1. To what extent does the NRM fit the response data in learning strategies in PISA 2012 South Korean data compared to the GPCM?

2. To what extent is the learning strategy use of South Korean students correlated to mathematics achievement linearly and nonlinearly?

Method

PISA 2012 Sampling Design

PISA is an OECD study of the achievement of 15-year-olds in mathematics, reading, and science. PISA 2012, the fifth PISA survey, covered reading, mathematics, science, problem solving, and financial literacy, with a primary focus on mathematics. In 2012, 65 countries and economies (all 34 OECD countries and 31 partner countries and economies) and around half a million students, representing 28 million 15-year-old students, participated in the PISA assessment. PISA 2012 adopted a two-stage complex survey design to select a representative sample of 15-year-old students in each educational system. In the first stage, around 150 schools were sampled, and then at least 35 students were selected in each sampled school (OECD, 2014). To acquire sufficiently high response rates, PISA required the school to have a minimum participation rate of 50%.

Sample

In the present study, the South Korean educational system was examined. PISA 2012 collected data from 5,033 15-year-old South Korean students (female = 47%) who participated (Dong et al., 2012). In PISA 2012, the total sample was 5,201: 6.1% middle school students, 73.7% general high school students, and 20.2% vocational high school students.

Mathematics Learning Strategies

PISA 2012 adopted a rotation design for the student questionnaire (OECD, 2014). The questionnaire included a common part and two of three rotating parts: set1, set2, and set3. Each student randomly received one of three questionnaire booklets. Therefore, each item for learning strategies had 33% of the data missing by design. Of 5,033 participating students, 3,310 students provided complete responses. Thus, the present study examined 3,310 South Korean students (female = 46%) in the analysis. Listwise deletion was employed, which involves deleting all persons with missing data, before proceeding with the analysis (Newman, 2014).

In PISA 2012, three types of learning strategies—memorization, elaboration, and metacognition—were measured using nominal scales. Four items were used to determine students' use of learning strategies in mathematics. Thus, students chose only one learning strategy from the three options (see Table 1).

Table 1

Mathematics Learning Strategy

Strategy	Statement
Item 1	
Metacognitive	When I study for a mathematics test, I try to work out what the most important parts to learn are.
Elaboration	When I study for a mathematics test, I try to understand new concepts by relating them to things I already know.
Memorization	When I study for a mathematics test, I learn as much as I can off by heart.
Item 2	
Metacognitive	When I study mathematics, I try to figure out which concepts I still have not understood properly.
Elaboration	When I study mathematics, I think of new ways to get the answer.
Memorization	When I study mathematics, I make myself check to see if I remember the work I have already done.
Item 3	
Metacognitive	When I study mathematics, I start by working out exactly what I need to learn.
Elaboration	When I study mathematics, I try to relate the work to things I have learnt in other subjects.
Memorization	When I study mathematics, I go over some problems so often that I feel as if I could solve them in my sleep.
Item 4	
Metacognitive	When I cannot understand something in mathematics, I always search for more information to clarify the problem.
Elaboration	I think about how the mathematics I have learnt can be used in everyday life.
Memorization	In order to remember the method for solving a mathematics problem, I go through examples again and again.

Mathematics Achievement

Each student was randomly assigned one of 13 booklets, which means that they tested a portion of the items from the entire item pool. PISA 2012 used the item response theory (IRT) framework to estimate a latent posterior distribution for each student. As the students did not answer all booklets, missing data must be inferred from the observed item responses. As one of several alternative approaches for making this inference, PISA used the imputation methodology called plausible values. Five plausible values were drawn from the posterior distribution with a mean of 500 and a standard deviation of 100 to represent students' mathematics scores (OECD, 2014). In this study, we used all five plausible values when conducting correlation analysis, as well as quadratic regression analysis, to take measurement errors into account.

Statistical Analysis

To answer the first research question, we compared the NRM to the GPCM in three domains: model fit index, empirical reliability, and item characteristic curve. As learning strategy items in PISA 2012 were measured on a nominal scale, the NRM was expected to be a better fit than models for ordinal data (e.g., the GPCM or graded response model). The GPCM assumes that the degree of discrimination between adjacent categories is the same for all adjacent categories in an item, whereas NRM releases these assumptions. If the NRM fits better than the GPCM, we can conclude that the nominal relationship among the three strategies is maintained. Then, we are able to create learning strategy scores based on their latent ability, considering the posterior distribution of estimates, and calculate plausible values of learning strategy scores in NRM.

To answer the second research question, we conducted two correlation analyses and a quadratic regression analysis with learning strategy use and mathematics achievement as three relationships (see Table 2). First, the correlation between the South Korean students' learning strategy scores created by the NRM and mathematic scores was obtained with the plausible values approach for taking the measurement error into account (OECD, 2009). Second, we examined the correlation between the observed raw score of each learning strategy and mathematics achievement to suggest the baseline value of each learning strategy. Third, we performed a quadratic regression analysis, adding a quadratic component of learning strategy score to a linear model to investigate the nonlinear relationship between the learning strategy score created by NRM and mathematics achievement. Finally, we tested the hypothesis for the coefficient of the quadratic term to understand the curvilinear relationship between mathematics achievement and learning strategy use.

Table 2

Analysis for Research Question 2

Purpose	Model or Statistics
Linear relationship	Correlation (Learning strategy score, Math)
Baseline of the linear relationship	Correlation (Memorization strategy, Math)
	Correlation (Elaboration strategy, Math)
	Correlation (Metacognitive strategy, Math)
Nonlinear relationship	Quadratic regression (Math ~ Learning strategy score + (Learning strategy score)^2)

Note. Learning strategy score is created by NRM. Math stands for mathematics achievement score.

Model comparison between NRM and GPCM

We used the *mirt* package in R (Chalmers et al., 2022) by the function of *mirt* with the argument of itemtype = “nominal” or itemtype = “gpcm” to introduce the NRM and GPCM models. In this study, under the GPCM, the options were recoded (Memorization = 1, Elaboration = 2, Metacognitive = 3) based on prior literature (Biggs, 1987; OECD, 2014; Weinstein & Mayer, 1986; Zimmerman & Pons, 1986). We compared the model data fit index of the NRM to that of the GPCM with the Akaike information criterion (AIC; Bozdogan, 1987), and Bayesian information criterion (BIC; Schwarz, 1978). In addition, empirical IRT reliability via sampling variances and empirical variances estimated from the expected a posteriori (EAP) method was used to indicate the features of these four items and how precise they were. To understand the meaning of the scores, we compared the item characteristic curves (ICCs) of the NRM to those of the GPCM, which also indicated that the equal discrimination constraint in the GPCM was inadequate.

Learning strategy score

We computed the five plausible values of ability estimates in NRM via the *fscores* function in the *mirt* package, which randomly sample five scores from the posterior distribution of θ in Equation 1. When a model contains latent regression predictors, the plausible values approach accounts for latent regression predictor effects and measurement error simultaneously (Chalmers et al., 2022). The coefficients of predictor effects with the plausible values approach are unbiased compared to other estimators, such as weighted likelihood estimates that underestimate coefficients and EAP, which overestimates coefficients (OECD, 2009). The plausible values of the learning strategy score were thus used for further correlation analysis and quadratic regression analysis.

We also computed raw scores of the learning strategies to suggest the baseline value of the association between single learning strategies and mathematics achievement. We created each raw score for memorization, elaboration, and metacognitive strategies equal to the frequency of choosing the corresponding strategy in the four items. For instance, if a student chose memorization once, elaboration twice, and metacognition once among four items, then the raw score for the student would be 1 for memorization, 2 for elaboration, and 1 for metacognition. If a student chose elaboration twice and metacognition twice, then the raw score would be 0 for memorization, 2 for elaboration, and 2 for metacognition. Compared to the NRM scores, which represented multiple learning strategies mixed together, the raw scores indicated the frequency of using single learning strategies.

Linear relationship (1): Correlation between learning strategy score and mathematics scores

The average of plausible value statistics was used for the point estimates of the population statistics. Thus, to obtain a correlation coefficient between the learning strategy scores and a mathematics achievement scores, the five correlation coefficients in the diagonal were computed and then averaged (OECD, 2009). Mathematically, secondary analyses with plausible values can be described as follows: The population coefficient ρ is the formulation of ρ_i, which is the coefficient computed on one plausible value, then:

$$\rho = \frac{1}{M}\sum_{i=1}^{M}\rho_i \quad (3)$$

where *M* is the number of plausible values.

To compute the uncertainty in the averaged correlation coefficient, the measurement variance, usually denoted as imputation variance, is equal to:

$$B_M = \frac{1}{M-1}\sum_{i=1}^{M}(\rho_i - \rho)^2. \quad (4)$$

This corresponds to the variance of the five plausible value statistics of interest. Finally, the sampling variance and the imputation variance should be combined as follows:

$$V = U + \left(1 + \frac{1}{M}\right)B_{\mathrm{M}} \quad (5)$$

where U is the sampling variance, and V is the squared standard error of the correlation coefficient between learning strategies and mathematics achievement.

Linear relationship (2): Correlation between learning strategy raw score and mathematics score

Besides the plausible values approach, we conducted correlation analysis between the learning strategy raw score and mathematics score to understand the correlation between the use of a single strategy and mathematics achievement. Different from the NRM scores representing the Korean students' multiple strategies used in learning mathematics, the raw scores of the learning strategies represented the use of a single strategy (i.e., memorization, elaboration, or metacognition) regardless of the Korean context. The correlation based on raw scores is the baseline value to understand to what extent the learning strategy score was related to mathematics achievement. For instance, if the correlation coefficient based on the NRM score is larger than that based on each single strategy raw score, we can conclude that the Korean student's learning strategy is more efficient than using a single strategy.

Nonlinear relationship: Quadratic relationship between learning strategy score and mathematics score

We performed a quadratic regression analysis on the basis of the previously examined linear model. The quadratic regression model contains a linear term and a quadratic term to capture the linear and quadratic relationships between learning strategies and mathematics achievement (Cohen et al., 2002). More specifically, we first specified a model assuming a linear relation. In the second step, we added a quadratic component to examine whether a curvilinear relationship described the data better. A curvilinear (second-order) predictor, such as X^2, is added to the linear regression equation ($Y = B_1X + B_0 + \varepsilon$) as follows:

$$Y = B_1X + B_2X^2 + B_0 + \varepsilon \tag{6}$$

where X is the learning strategy score from the NRM as a predictor, X^2 is squared value of learning strategy score as a curvilinear (second-order) predictor, Y is the mathematics score as an outcome variable, and ε is an error term with mean to zero and variance to the residual variance. In addition, B_0 is the intercept of the equation, implying the mean of the mathematics score when the learning strategy score is zero. B_1 is a regression coefficient of X (learning strategy score) and B_2 is a quadratic coefficient of X^2(the squared value of learning strategy score). When conducting hypothesis testing, the significance of the quadratic coefficient (B_2) was explored. If a quadratic term B_2 is significant ($p < .001$), this implies that mathematic achievement scores did not monotonically increase as learning strategy scores increased.

In the quadratic regression analysis, the plausible values of regression coefficients were calculated using the same method in the linear relationship (1), and the five plausible values of the regression coefficients (e.g., LS1-MATH1, LS2-MATH2, … LS5-MATH5) were averaged (OECD, 2009) to take measurement errors into account.

Results

Frequencies of Learning Strategy Used Among South Korean Students

The frequencies and percentages of learning strategy use by South Korean students are presented in Table 3. The primary learning strategy varied across items, and metacognitive strategy was the most frequent learning strategy overall. For instance, more than 80% of South Korean students reported using elaboration and metacognition (43.3% and 40.0% each), with only 20% using memorization in Item 1. The percentage of metacognitive strategy use was the most noticeable in Items 2 and 3. In Item 2, South Korean students reported using metacognition, memorization, and elaboration (51.1%, 29.6%, and 19.3%, respectively). In Item 3, they chose metacognition, elaboration, and memorization at 62.5%, 22.9%, and 14.6%, respectively. In Item 4, more than half of the students reported the use of memorization (54.9%), followed by metacognition (30.9%) and elaboration (14.2%).

Table 3

Frequencies and Percentages of Students' Responses to Learning Strategy Items

	Item1		Item2		Item3		Item4	
	n	%	n	%	n	%	n	%
Metacognitive	1,323	40.0	**1,693**	**51.1**	**2,068**	**62.5**	1,023	30.9
Elaboration	**1,432**	**43.3**	638	19.3	759	22.9	470	14.2
Memorization	555	16.8	979	29.6	483	14.6	**1,817**	**54.9**

Note. Numbers in bold indicate the highest response probability within an item.

Comparison Between NRM and GPCM

Model fit

We used the likelihood ratio test to compare the model data fit between the NRM and the GPCM. Lower AIC and BIC values represent a better fit. Table 4 shows that the NRM had lower value on AIC, and BIC than the GPCM. Thus, the NRM was found to be a better fit for the data. BIC and AIC penalize the number of parameters, so the superiority of fit for the NRM is not merely due to the number of parameters.

Table 4

Comparison of Fit Indices in Models

Model	logLik	AIC	BIC
NRM	**-12,873.10**	**25,778.21**	**25,875.88**
GPCM	-12,980.61	25,985.22	26,085.47

Note. The values in bold represent the smaller value in log likelihood, AIC, and BIC.

Reliability

To estimate the reliability of learning strategy items, we reported empirical reliability. The empirical reliability for the NRM was 0.365, and 0.214 for the GPCM. Considering that there were only four items, it is probable that both models' empirical reliabilities were low. In general, the more information (i.e., more items in the test) we have, the more reliably we can measure the underlying trait (Cheng et al., 2012). Still, the reliability of the NRM was higher than that of the GPCM.

Item characteristic response curve

The NRM redefined each learning strategy regardless of recoded numbers. As c_{ik} indicates the relative frequency of choosing that option (compare Tables 3 and 5), the P1 curve showed memorization, the P2 curve showed elaboration, and the P3 curve showed metacognition (see Figure 1). Under GPCM, within item *i*, category slopes are all the same (a_{i1} = a_{i2} = a_{i3}= a_i). These constraints led Item 2, Item 3, and Item 4 response curves to be shaped differently from the curves of the NRM.

Table 5

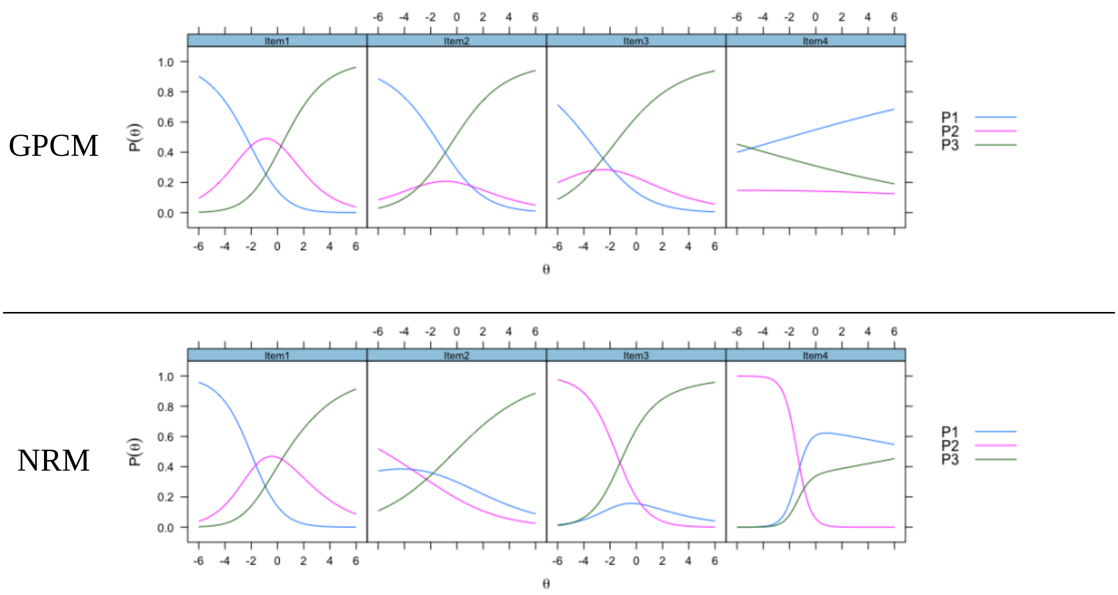
Item Parameters for GPCM and NRM

Model	Parameter	Item			
		1	2	3	4
GPCM					
	a	0.567	0.333	0.303	-0.059
	b1	- 2.034	1.066	- 1.775	-23.078
	b2	0.272	-2.847	- 3.300	13.240
NRM					
	a1	- 0.63	-0.055	0.125	0.628
	a2	0.106	-0.186	-0.54	-1.322
	a3	0.524	0.241	0.415	0.695
	c1	- 0.762	-0.03	-0.555	0.989
	c2	0.451	-0.482	-0.315	-1.39
	c3	0.312	0.512	0.87	0.401

Note. a) An item discrimination parameter of GPCM and the slope of each category in NRM; b) A threshold of GPCM; c) An intercept of each category in NRM

Figure 1

Item Characteristic Response Curve of GPCM and NRM



Note. X-axis = θ (learning strategy score), Y-axis = Probability of selecting the strategy, P1= Memorization strategy, P2 = Elaboration strategy, and P3 = Metacognitive strategy

Overall, the response curve patterns of the GPCM and those of the NRM were different in Items 2, 3, and 4, while Item 1 had a similar pattern in both models. In Item 2, under the GPCM, as learning strategy score (θ) increased, the probability of choosing memorization (P1/ blue) decreased with a steeper slope than under the NRM, and the probability of choosing elaboration (P2/pink) increased for learning strategy scores lower than 0 and decreased for learning strategy scores higher than 0. Under the NRM, the probability of choosing elaboration (P2/pink) decreased monotonically as learning strategy score (θ) increased. In Item 3, under the GPCM, as learning strategy score (θ) increased, the probability of choosing memorization (P1/blue) decreased with a steep slope, and the probability of choosing elaboration (P2/pink) increased for learning strategy scores lower than about -2 but decreased for learning strategy scores higher than -2. Under the NRM, the probability of choosing memorization (P1/blue) increased for learning strategy scores lower than 0 but decreased for learning strategy scores higher than 0, and the probability of choosing elaboration (P2/pink) decreased steeply as learning strategy score (θ) increased. Item 4 curves of the GPCM showed very different shapes from those of NRM. Under GPCM, as learning strategy scores (θ) increased, the probability of choosing memorization (P1/blue) increased monotonically, the probability of choosing elaboration (P2/pink) did not seem to change, and the probability of choosing metacognition (P3/green) decreased monotonically. Under the NRM, as learning strategy score (θ) increased, the probability of choosing memorization (P1/blue) increased for learning strategy scores lower than 0 but decreased slightly for learning strategy scores higher than 0, the probability of choosing elaboration (P2/pink) decreased with a steep slope, and the probability of choosing metacognition (P3/green) increased gradually with a slope change around 0.

As comparing the NRM to the GPCM, the NRM showed the better model. Therefore, we used NRM to create the South Korean students' learning strategy scores. To understand the implication of the learning strategy scores, item characteristic response curves and parameter estimates of the NRM (a_{ij}, c_{ik}) were considered (see Table 5 and Figure 1). In general, a_3 values were larger than the other two a_1 and a_2 values in all four items. However, a_1 and a_3 had similar values in Item 4. The NRM curves in Figure 1 showed that memorization strategy (P1/blue) was slightly higher than metacognitive strategy (P3/green) in Item 4 when θ (learning strategy score) was around 6.0. Thus, in general, the learning strategy score in the NRM might suggest the use of metacognitive strategies with memorization strategies. The higher θ (higher learning strategy score) implied more frequent use of metacognitive strategies with memorization strategies, depending on the context. For example, the students who have a high learning strategy score tended to use memorization in Item4 but metacognition in Items 2 and 3.

The Relationship Between Learning Strategy and Mathematics Achievement

To summarize briefly, the result of the first correlation analysis is that the learning strategy score from the NRM was positively and linearly associated with mathematics achievement. As the result of the raw score correlation, we found that mixed use of learning strategies was more effective than use of a single strategy for South Korean students. Finally, we explored a curvilinear relationship by adding a quadratic term of the learning strategy score from the NRM and found a significant negative association with mathematics achievement. More detailed results about the findings are presented below.

Linear relationship (1): Correlation between learning strategy score and mathematics score

All correlation coefficients between learning strategy scores from the NRM and mathematics scores were significantly larger than zero ($p < .05$). In other words, the confidence intervals of the correlations between the variables did not include zero. The mean of the correlation coefficients was 0.18 (SE = 0.00075, Range = 0.17–0.22). The results indicate that there was a tendency that the higher the mathematics score, the higher the learning strategy score, and the reverse also applied. Thus, the South Korean students who primarily used the metacognitive strategy with memorization, depending on the context, obtained high scores on mathematics exams.

Linear relationship (2): Correlation between learning strategy raw score and mathematics score

The second correlation analysis between the raw score of learning strategy and mathematics score was conducted. The correlation based on raw scores was the baseline value for understanding to what extent the students' learning strategy scores were related to mathematics achievement. The correlation coefficients between the raw score of the single learning strategy and mathematics score were all significantly different from zero ($p < .05$). The mean of correlation coefficients between raw score of metacognitive strategy and mathematics score was 0.12. In contrast, the mean correlation coefficients between elaboration strategy and mathematics score were negative, as they were for memorization strategy. The means of the correlation coefficients for elaboration and memorization were -0.04 and -0.10, respectively. These results support the evidence that those students who used metacognition exclusively tended to achieve higher mathematics scores than those who used elaboration or memorization exclusively. More specifically, the sole use of memorization or elaboration learning strategies had a negative impact on mathematics scores.

The mean of correlation coefficients between a raw score of metacognitive strategy and mathematics score was 0.12, which was less than 0.18 (i.e., the mean of correlation coefficients between the learning strategy score from the NRM and mathematics score). This indicates that students using mixed learning strategies had higher mathematics achievement scores than those who used only metacognitive strategies, in line with previous research (Wu et al., 2020).

Nonlinear relationship: Quadratic relationship between learning strategy score and mathematics score

The quadratic regression coefficients were significantly different from zero ($p < .001$). The average R-squared difference between the quadratic regression model and the linear regression model was 0.00828. In fact, the quadratic regression model showed a better fit than the linear model; the mean value of BIC for quadratic regression was smaller than that of the linear model (39.760 and 39.771, respectively). This finding confirmed our expectation that a linear relationship could not fully capture the students' learning strategies.

We presented a significance test for each of the individual regression coefficients with 95% confidence intervals on the mean of the five regression coefficients (see Table 6). We also summarized both linear and quadratic regression models in the scatter plot (see Figure 2).

Table 6

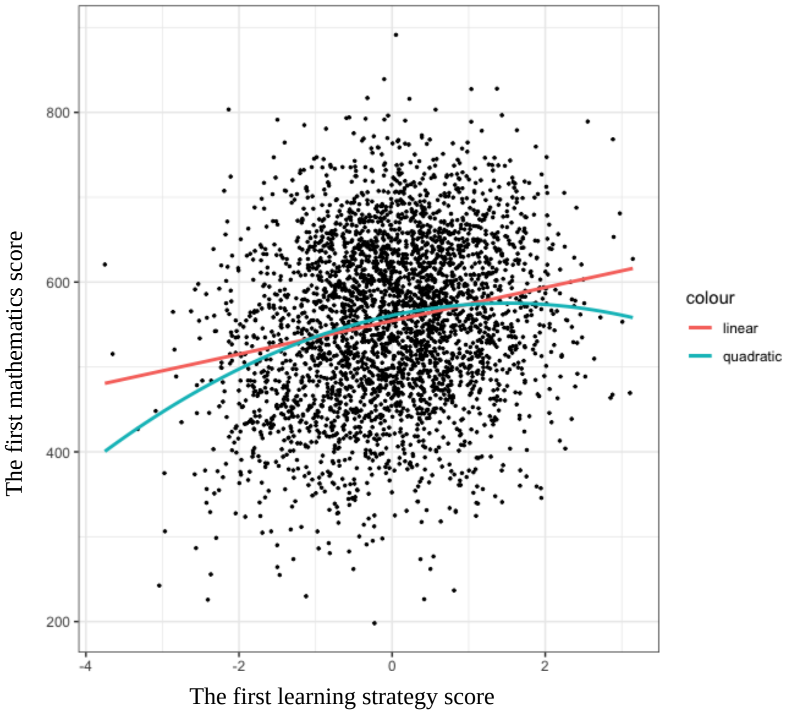
Quadratic Regression Coefficients of the Learning Strategy Score on the Mathematics Score

Variable	B	SE	95% CI [Lower, Upper]	β	SE	95% CI [Lower, Upper]
Intercept	559.531***					
LS	18.010***	3.468	[11.212, 24.809]	0.1802***	0.00069	[0.179, 0.182]
LS^2	-5.411***	0.519	[-6.428, -4.396]	-0.0754***	0.00045	[-0.0763, -0.0745]
R^2	0.0424					

Note. N = 3,310. B = coefficient, β = standardized coefficient, CI = confidence interval. LS is the learning strategy score. The mathematics score from PISA 2012 is an outcome variable.
***$p < .001$.

Figure 2

Association between Students' Learning Strategy Score and Mathematics Score



Note. Each dot indicates an individual participant's first plausible value of mathematics scores and learning strategy scores. The red line characterizes the best fit linear regression of mathematics score on the first learning strategy score in the NRM; the blue line represents the best fit quadratic regression of the first mathematics score on the first learning strategy score in the NRM.

Table 6 shows the mean of the standardized quadratic regression coefficients of learning strategy scores (i.e., LS^2) was significantly negative at -0.0754 (p < .001). The negative average coefficient of quadratic regression estimates implies that the initially positive association between the learning strategy score and mathematics achievement diminished slightly and became negative as the value of learning strategy score increased.

The scatter plot in Figure 2 shows both the positive linear relationship and negative quadratic relationship between learning strategy score in NRM (x-axis) and mathematics score. To simplify the scatter plot, we used the first plausible value of learning strategy scores and the first plausible value of mathematics scores as x-variable and y-variable, respectively, rather than five plausible values. The red line and blue curve in Figure 2 are the fitted linear regression line and the fitted quadratic regression curve, respectively. The linear regression line (red line) was fitted to $\hat{Y} = 19.640\ LS1 + 554.386$, where $\hat{Y}$ is the predicted mathematics achievement score and $LS1$ is the first plausible value of learning strategy scores in the NRM (i.e., x-variable in Figure 2). The coefficient of $LS1$ in the linear equation was 19.640 (CI = [16.38, 22.900]), which was significantly larger than zero, with *t*(3307) = 1.663 and *p* < .001. The standardized coefficient of $LS1$ was 0.195. The blue curve in Figure 2 indicates the fitted quadratic regression equation with predictors of both $LS1$ and $LS1^2$, which was $\hat{Y} = 19.019\ LS1 + (-6.319)LS1^2 + 560.841$, where the coefficient of $LS1^2$ was -6.319 with the 95% confidence interval did not include zero (CI = [-8.659 -3.978]), and the hypothesis testing for the coefficient of $LS1^2$ showed a significant negative value with *t*(3307) = -5.293 and *p* < .001. Likewise, as shown in Table 6, the negative quadratic relationship between the variables took the shape of an inverted U-curve rather than a straight line. The implications of the inverted U-curve are presented in the discussion section.

Discussion

This study aimed to explore the link between South Korean students' learning strategy use and their achievement in mathematics exams using the NRM. We found that the Korean students who primarily used the metacognitive strategy with memorization, depending on the context, achieved high scores on mathematics exams with the limited effect. Our investigation extended previous research in two ways. First, it created scores for learning strategy use with the NRM. Second, it addressed the existence of a curvilinear relationship between learning strategy scores and mathematics achievement, as well as the linear relationship between the variables, focusing on one of top-performing East Asian education system (i.e., South Korea). A more detailed discussion of the findings is presented below.

The curvilinear relationship between learning strategy score and mathematics achievement

Learning strategy score had a positive linear relationship with mathematics achievement. Even so, Table 6 shows that a linear relationship may not accurately reflect the nature of the association, and Figure 2 also indicates a curvilinear pattern. The negative curvilinear coefficient indicates the presence of a curvilinear association between learning strategy score and mathematics achievement. The increasing use of metacognitive and memorization strategies was correlated with higher achievement in mathematics until it reached an optimum value; then, this association decreased slightly as the use of both strategies increased. This nonlinear pattern indicates that excessive use of metacognition and memorization may have diminishing returns for increasing student achievement, and that more use of metacognition and memorization does not necessarily lead to better performance. In other words, the learning strategy combination of metacognition and memorization might not be the best strategy combination for every high-performing Korean student. Our finding is also in line with the previous study (Wu et al., 2020). Wu et al., (2020) suggested the 14.2% of students who primarily used metacognition with elaboration performed slightly better on the mathematics exam than the 65% of students who primarily used metacognition with memorization. Still, the effect size of the curvilinear model is small; Cohen (1992) suggested that 0.02 is reflective of a small effect size.

In addition, two possible responses to Item 4, memorization and metacognition, would lead to the high learning strategy score. It could be a probable reason for the negative nonlinear relationship between learning strategy score and mathematic achievement. In other words, both the students who used metacognition for all items and those who used metacognition for Items 1, 2, and 3 but memorization for Item 4 would get a high learning strategy score. The former students could get lower mathematics scores than the latter because memorization strategy of Item 4 is more effective than metacognitive strategy of Item 4 according to Item contents (see Table 1). A more detailed explanation of the high frequent use of memorization strategy in Item 4 is suggested in another subsection.

Use of metacognitive strategies with other learning strategies

We found that high-performing students in South Korea reported heavy use of metacognitive strategies with memorization strategies. This implies that the students did not use metacognitive strategies or memorization strategies alone, which is in line with previous studies (Nathan, 2021; Quigley et al., 2018). Nathan (2021) suggested that metacognition could only be developed in within-subject or content-based lessons, and with other learning strategies. Thus, metacognitive strategies rely upon the use of other cognitive strategies (e.g., memorization and elaboration) and content that learners can use to plan, monitor, and evaluate. For example, if students who are self-regulated learners (Zimmerman, 1986) were asked to solve a math question with regard to mathematical formulas, they would start with some knowledge of the task and strategies. They could utilize one of the formulas that they already knew (i.e., elaboration strategy). In the process of recalling possible formulas, understanding the formula and practicing its use repeatedly in advance are necessary (i.e., memorization). They could then evaluate their overall success and check whether they were correct. If their answers were wrong, they could try other strategies (Quigley et al., 2018). Therefore, the finding that high-achieving students in South Korea use mixed learning strategies makes sense.

Variation in the use of memorization learning strategies

More than half of the students reported using the memorization strategy in Item 4 (i.e., In order to remember the method for solving a mathematics problem, I go through examples again and again). Although memorization is generally regarded as a relatively inefficient strategy (e.g., rote learning), the memorization strategy of Item 4 (In order to remember the method for solving a mathematics problem, I go through examples again and again) is no closer to the rote learning concept than the other memorization strategies in Items 1 and 3 (Wu et al., 2020). In fact, "go through examples" represents a common practice method in mathematics learning, especially in the introductory stages (Dinsmore & Alexander, 2016).

In South Korea, the most common way to learn mathematics in class is by doing different examples repeatedly, regardless of the students' level. The variation of examples is associated with the students' mathematics level or step of the mathematics learning process. In the beginning stages of learning, most students do examples with minor variations (e.g., numbers or $\pm,\times,\div$). It is common for low-achieving students to do less varied examples and even the same examples from the textbook repeatedly, which can lead to rote learning. When relatively high-achieving students do the examples with increasing variation, this can be considered a "route to understanding" (Marton & Booth, 1997; Hess & Azuma, 1991). High-performing students even create and solve examples of their own.

In South Korea, students are usually encouraged to make their own review notes for wrong answers, called *Odabnote* (i.e., incorrect answer note or incorrect note), particularly in mathematics exams (Moon, 2019). After exams, they take notes to review the wrong answers. They report what they did wrong, why it was wrong, and even what a new question could be based on the concepts they got wrong. Then, they review their notes by going through not only the same questions, but also their own examples before exams. Thus, frequent use of the memorization strategy, like Item 4, does not necessarily mean rote learning. Still, to examine whether the use of memorization strategies causes rote learning needs further research with different methods, such as cognitive lab or think-aloud.

Use of learning strategies in the South Korean education system

The majority of PISA test-taking students in South Korea (79.8%) are from general secondary schools, which are academically oriented and sometimes called college preparatory schools, where most Korean secondary school students are enrolled (Kim & Byun, 2014). Most students consider university entrance exams to be very important (Lee, 2010; Ripley, 2013), prompting them to study mathematics, which is one of the core subjects that determines their future college options (Hwang, 2001; Yoon et al., 2021). According to the OECD, South Korean high school students study mathematics for 10.4 hours per week on average, which is 3 hours more than the OECD country average (7.6h; Lee, 2014). In addition, 50.2% of South Korean students engage in private tutoring (at a *hagwon* or through informal private instruction by a university student) to study mathematics, which is more than in other subjects (e.g., English, Korean, and science).

Considering how much time they spent studying mathematics and their reasons for doing so, we can understand why the fewest South Korean students reported using elaboration strategies (19.3%) in Item 2, while more than half (51.1%) reported using metacognitive strategies. The elaboration strategy in Item 2 (When I study mathematics, I think of new ways to get the answer) might not be an efficient way for them to learn mathematics, especially for 15-year-olds who learn mathematics in a highly stressful and competitive environment. If they already know how to solve a problem, they do not need to find another way. They are more likely to spend time figuring out what they do not understand (i.e., metacognitive learning strategy) to get more answers correct on their mathematics exams. This may explain why more than half of the students chose the metacognitive strategy in Item 2 (When I study mathematics, I try to figure out which concepts I still have not understood properly). Likewise, the fewest students reported using the elaboration strategy in Items 3 and 4. These items asked if students thought about and related their knowledge to other subjects (Item 3) or their lives (Item 4), both of which are unnecessary for finding an answer in a mathematics exam. This finding is related to why the raw score of elaboration strategies was negatively correlated to mathematics achievement. Regardless of whether using elaboration strategies deepens leaners' understanding of knowledge and leads to high-quality learning outcomes (Marton & Säljö, 1976; Prosser & Millar, 1989), it does not necessarily mean achieving high scores on mathematic exams.

In contrast, 62.5% of South Korean students reported using metacognitive strategies in Item 3 (When I study mathematics, I start by working out exactly what I need to learn). As the importance of metacognition is emphasized in education, education stakeholders in South Korea, including private cram schools, are very interested in metacognitive learning strategies (Ji, 2021). Not only has the school curriculum focused on how to teach these strategies, but more cram schools are also advertising themselves using the slogan; "The secret to getting 100% on a mathematics exam: metacognitive learning strategies." Although metacognition should not be misunderstood as a process of verifying true and false, right and wrong, or good and bad (Park, 2021), South Korean education stakeholders could misuse metacognitive strategies to verify which mathematical knowledge is helpful for performing well on exams. Thus, it is probable that South Korean PISA test-taking students think of Item 3 as "When I study mathematics, I start by working out exactly what I need to learn *for the mathematics exam*." To accept the widely held assumption that metacognition is beneficial, it could, at least in part, be understood as a result of its close relationship to self-regulation (Efklides, 2011; Norman, 2020; Zimmerman, 2008). Therefore, future studies should investigate how South Korean students use metacognitive strategies to illustrate that they can produce positive effects.

Limitations

There are some limitations of this study that are worth noting. First, as we focused on the Korean context, the degree to which the findings generalize to other populations is uncertain. Thus, to generalize the relationship between learning strategy and mathematics achievement in other countries, other factors, such as cultural context, should be considered. Second, this study is based on self-reported learning strategy data, which may not mirror students' actual learning strategy use. A follow-up study in which learning strategy is assessed using different methodologies (e.g., observational data, think-aloud, and retrospective think-aloud) would add to the weight of these findings (Wu et al., 2020). Third, the present study focuses on learning strategies and mathematics achievement without accounting for other psychological variables (e.g., motivation and behavior), which SRL theoretical framework suggests (Wu et al., 2020), and other test-taking strategies. To fully understand South Korean students' high achievement in mathematics, further studies need to consider psychological characteristics and other practical strategies that students might use for exams.

Conclusion

This research explored the relationship between learning strategy use and mathematics achievement in the South Korean education system using the NRM. The findings show that frequent use of metacognitive strategy with memorization is positively related to South Korean students' mathematics achievement until it reaches an optimum value. We extended earlier research by creating learning strategy scores via the NRM. Our results also provide insight into the multifaceted nature of the association between learning strategy use and mathematics achievement by examining the existence of a curvilinear relationship. This study also discussed the relationship between the variables based on the South Korean education system. The necessity of further studies on how students use each learning strategy based on a specific education system was highlighted. Overall, these results are useful for understanding South Korean students' learning strategy use for mathematics achievement.

References

Areepattamannil, S., & Caleon, I. S. (2013). Relationships of cognitive and metacognitive learning strategies to mathematics achievement in four high-performing East Asian education systems. *The Journal of Genetic Psychology*, *174*(6), 696–702. https://doi.org/10.1080/00221325.2013.799057

Artz, A. F., & Armour-Thomas, E. (2009). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*. Advance online publication https://doi.org/10.1207/s1532690xci0902_3

Biggs, J. B. (1987). Student approaches to learning and studying. Learning process questionnaire manual. *Australian Council for Educational Research*. https://eric.ed.gov/?id=ED308199

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, *63*(1), 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x

Biggs, J. (1998). Learning from the Confucian heritage: So size doesn't matter? *International Journal of Educational Research*, *29*(8), 723–738. https://doi.org/10.1016/S0883-0355(98)00060-3

Blazer, C. (2012). Is South Korea a case of high-stakes testing gone too far? Information capsule. Volume 1107. *Research services, Miami-Dade County Public Schools*. https://eric.ed.gov/?id=ED536521

Borkowski, J. G., & Thorpe, P. K. (1994). Self-regulation and motivation: A life-span perspective on underachievement. In *Self-regulation of learning and performance: Issues and educational applications* (pp. 45–73). Lawrence Erlbaum Associates, Inc.

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*(3), 345–370. https://doi.org/10.1007/BF02294361

Campione, J. C., Brown, A. L., Ferrara, R. A., & Bryant, N. R. (1984). The zone of proximal development: Implications for individual differences and learning. *New Directions for Child and Adolescent Development*, *1984*(23), 77–91. https://doi.org/10.1002/cd.23219842308

Carver, C. S., & Scheier, M. F. (1981). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology*, *17*(6), 545–568. https://doi.org/10.1016/0022-1031(81)90039-1

Chalmers, P., Pritikin, J., Robitzsch, A., Zoltak, M., Kim, K., Falk, C. F., Meade, A., Schneider, L., King, D., Liu, C.-W., & Oguzhan, O. (2022). MIRT: Multidimensional item response theory (Version 1.31) [Computer software]. https://cran.r-project.org/web/packages/mirt/mirt.pdf

Cheng, Y., Yuan, K.-H., & Liu, C. (2012). Comparison of reliability measures under factor analysis and item response theory. *Educational and Psychological Measurement*, *72*(1), 52–67. https://doi.org/10.1177/0013164411407315

Chiu, M. M., Chow, B. W.-Y., & Mcbride-Chang, C. (2007). Universals and specifics in learning strategies: Explaining adolescent mathematics, science, and reading achievement across 34 countries. *Learning and Individual Differences*, *17*(4), 344–365. https://doi.org/10.1016/j.lindif.2007.03.007

Choi, Y., Kim, S. won, & Hong, W.-P. (2019). Is the role of cultural capital in student achievement in South Korea different? A systematic review. *British Journal of Sociology of Education*, *40*(6), 776–794. https://doi.org/10.1080/01425692.2019.1592662

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge. https://doi.org/10.4324/9780203774441

Corno, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology*, *11*(4), 333–346. https://doi.org/10.1016/0361-476X(86)90029-9

Corno, L., & Mandinach, E. B. (2009). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*. Advance online publication. https://doi.org/10.1080/00461528309529266

Darrell Bock, R. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, *37*(1), 29–51. https://doi.org/10.1007/BF02291411

De Clercq, A., Desoete, A., & Roeyers, H. (2000). Epa2000: A multilingual, programmable computer assessment of off-line metacognition in children with mathematical-learning disabilities. *Behavior Research Methods, Instruments, & Computers*, *32*(2), 304–311. https://doi.org/10.3758/BF03207799

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Desoete, A., Roeyers, H., & Buysse, A. (2001). Metacognition and mathematical problem solving in Grade 3. *Journal of Learning Disabilities*, *34*, 435–449. https://doi.org/10.1177/002221940103400505

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Dinsmore, D. L., & Alexander, P. A. (2016). A multidimensional investigation of deep-level and surface-level processing. *The Journal of Experimental Education*, *84*(2), 213–244. https://doi.org/10.1080/00220973.2014.979126

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, *11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

Dong, H., Ok, H., Lim, H., Jeong, H., Son, S., & Seong, B. (2012). OECD international academic achievement assessment study: PISA 2012 main test implementation report (RRE 2012-3-1), 178–179. (in Korean)

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, *104*(2), 452. https://doi.org/10.1037/a0026683

Hess, R. D., & Azuma, H. (1991). Cultural support for schooling: Contrasts between Japan and the United States. *Educational Researcher*, *20*(9), 2–9. https://doi.org/10.3102/0013189X020009002

Hong, E., Sas, M., & Sas, J. C. (2006). Test-taking strategies of high and low mathematics achievers. *The Journal of Educational Research*, *99*(3), 144–155. https://doi.org/10.3200/JOER.99.3.144-155

Hwang, Y. (2001). Why do South Korean students study hard? Reflections on Paik's study. *International Journal of Educational Research*, *35*(6), 609–618. https://doi.org/10.1016/S0883-0355(02)00014-9

Ji, S. (2021, November 11). What is the professionalism of teachers required by the 2022 revised curriculum and high school credit system? Senior teachers' association emphasizes 'metacognition and Q&R learning methods'. *Education Plus*. http://www.edpl.co.kr/news/articleView.html?idxno=3199

Kember, D., Biggs, J., & Leung, D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, *74*(2), 261–279. https://doi.org/10.1348/000709904773839879

Kember, D. (2016). Why do Chinese students out-perform those from the West? Do approaches to learning contribute to the explanation? *Cogent Education*, *3*(1), 1248187. https://doi.org/10.1080/2331186X.2016.1248187

Kim, H. (2004). Analyzing the effects of the high school equalization policy and the college entrance system on private tutoring expenditure in Korea. *KEDI Journal of Educational Policy*, *1*(1). https://www.proquest.com/docview/1013966600/abstract/8C5C215702634B8APQ/1

Kim, K., & Byun, S. (2014). Determinants of Academic Achievement in Republic of Korea. In H. Park & K. Kim (Eds.), *Korean Education in Changing Economic and Demographic Contexts* (pp. 13–37). Springer. https://doi.org/10.1007/978-981-4451-27-7_2

Kilic, S., Cene, E., & Demir, I. (2012). Comparison of learning strategies for mathematics achievement in Turkey with eight countries. *Educational Sciences: Theory and Practice*, *12*(4), 2594–2598.

Lee, J., & Shute, V. J. (2010). Personal and social-contextual factors in K–12 academic performance: An integrative perspective on student learning. *Educational Psychologist*, *45*(3), 185–202. https://doi.org/10.1080/00461520.2010.493471

Lee, Y. (2010). Views on education and achievement: Finland's story of success and South Korea's story of decline. *KEDI Journal of Educational Policy*, *7*, 379–401.

Lee, S. (2014, July 14). Mathematics scores are not proportional to study time. *Science Times*. https://bit.ly/3spXR4K

Leung, F. K. S. (2014). What can and should we learn from international studies of mathematics achievement? *Mathematics Education Research Journal*, *26*(3), 579–605. https://doi.org/10.1007/s13394-013-0109-0

Li, A. (2011, October 4). Korea cracks down on clandestine study groups. *The Toronto Star*. http://thestar.com/news/article/1064697

Lin, S.-W., & Tai, W.-C. (2015). Latent class analysis of students' mathematics learning strategies and the relationship between learning strategy and mathematical literacy. *Universal Journal of Educational Research*, *3*(6), 390–395. https://doi.org/10.13189/ujer.2015.030606

Liu, Q., Du, X., Zhao, S., Liu, J., & Cai, J. (2019). The role of memorization in students' self-reported mathematics learning: A large-scale study of Chinese eighth-grade students. *Asia Pacific Education Review*, *20*(3), 361–374. https://doi.org/10.1007/s12564-019-09576-2

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process*. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Marton, F., & Booth, S. A. (1997). *Learning and awareness*. Psychology Press. https://doi.org/10.4324/9780203053690

McInerney, D. M., Cheng, R. W., Mok, M. M. C., & Lam, A. K. H. (2012). Academic self-concept and learning strategies: Direction of effect on student academic achievement. *Journal of Advanced Academics*, *23*(3), 249–269. https://doi.org/10.1177/1932202X12451020

Moon, Y. (2019, April 19). How to write an incorrect answer note to be top in the mathematics class. *Edugene Internet Education Newspaper*. https://www.edujin.co.kr/news/articleView.html?idxno=30692

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

Nathan, B. (2021, January 30). 5 myths about metacognition that we need to banish. *Tes Magazine*. https://www.tes.com/magazine/archived/5-myths-about-metacognition-we-need-banish

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, *17*(4), 372–411. https://doi.org/10.1177/1094428114548590

Norman, E. (2020). Why metacognition is not always helpful. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/article/10.3389/fpsyg.2020.01537

OECD. (2005). *PISA 2003 technical report*. Paris, France: OECD Publishing.

OECD. (2009). *PISA data analysis manual: SPSS, second edition*. OECD. https://doi.org/10.1787/9789264056275-en

OECD. (2012). *PISA 2009 technical report*. Paris, France: OECD Publishing. https://doi.org/10.1787/9789264167872-en

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.

Park, K. (2004). Factors contributing to Korean students' high achievement in mathematics. *Korea*, *547*, 84. http://matrix.skku.ac.kr/For-ICME-11/ICME/Chap5(kPark).htm

Park, H. (2021, July 23). The beginning of empathy is 'metacognition'. *Shin-A Ilbo*. http://www.shinailbo.co.kr/news/articleView.html?idxno=1439847

Perels, F., Dignath, C., & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education*, *24*(1), 17–31. https://doi.org/10.1007/BF03173472

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33. https://doi.org/10.1037/0022-0663.82.1.33

Pires, E. M. S. G., Daniel-Filho, D. A., de Nooijer, J., & Dolmans, D. H. J. M. (2020). Collaborative learning: Elements encouraging and hindering deep approach to learning and use of elaboration strategies. *Medical Teacher*, *42*(11), 1261–1269. https://doi.org/10.1080/0142159X.2020.1801996

Prosser, M., & Millar, R. (1989). The "How" and "What" of learning physics. *European Journal of Psychology of Education*, *4*(4), 513–528. https://doi.org/10.1007/BF03172714

Quigley, A., Muijs, D., & Stringer, E. (2018). Metacognition and self-regulated learning: Guidance report.

Ramsden, P. (1988). Context and strategy. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 159–184). Springer. https://doi.org/10.1007/978-1-4899-2118-5_7

R Core Team. (2020). R*: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ripley, A. (2013). *The smartest kids in the world: And how they got that way*. Simon and Schuster.

Rosander, P., & Bäckström, M. (2012). The unique contribution of learning approaches to academic performance, after controlling for IQ and personality: Are there gender differences? *Learning and Individual Differences*, *22*(6), 820–826. https://doi.org/10.1016/j.lindif.2012.05.011

Schoenfeld, A. H. (2016). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics (Reprint). *Journal of Education*, *196*(2), 1–38. https://doi.org/10.1177/002205741619600202

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Thissen, D., Cai, L., & Bock, R. D. (2010). The nominal categories item response model. In *Handbook of polytomous item response theory models*. Routledge.

Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: The influence of learning context and student approaches to learning on learning outcomes. *Higher Education*, *22*(3), 251–266. https://doi.org/10.1007/BF00132290

Walker, C. O., Greene, B. A., & Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences*, *16*(1), 1–12. https://doi.org/10.1016/j.lindif.2005.06.004

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). Macmillan.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to Predict Students' Motivation, Cognition, and Achievement. *Journal of Educational Psychology*, *96*(2), 236–250. https://doi.org/10.1037/0022-0663.96.2.236

Wu, Y.-J., Carstensen, C. H., & Lee, J. (2020). A new perspective on memorization practices among East Asian students based on PISA 2012. *Educational Psychology*, *40*(5), 643–662. https://doi.org/10.1080/01443410.2019.1648766

Yoon, H., Bae, Y., Lim, W., & Kwon, O. N. (2021). A story of the national calculus curriculum: How culture, research, and policy compete and compromise in shaping the calculus curriculum in South Korea. *ZDM – Mathematics Education*, *53*(3), 663–677. https://doi.org/10.1007/s11858-020-01219-w

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5

Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, *23*(4), 614–628. https://doi.org/10.3102/00028312023004614

Zimmerman, B. J., & Pons, M. M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, *80*(3), 284–290. https://doi.org/10.1037/0022-0663.80.3.284

Zimmerman, B. J. (1990). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329. https://doi.org/10.1037/0022-0663.81.3.329

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 1–39). Lawrence Erlbaum Ass. Publishers.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*, 166–183. https://doi.org/10.3102/0002831207312909

Zu, J., & Kyllonen, P. C. (2020). Nominal response model is useful for scoring multiple-choice situational judgment tests. *Organizational Research Methods*, *23*(2), 342–366. https://doi.org/10.1177/10944281188

Appendix

Appendix I: GDPR Documentation and Ethical Approval

The current research was not subject to GDPR (General Data Protection Regulation) because we used only anonymous data in PISA 2012 Database. Thus, we attached the mock registration NSD (Norwegian Centre for Research Data) application form.

Which personal data will be processed?

What are personal data?

What is processing?

Name (also with signature/written consent)

Yes No

National ID number or other personal identification number

Yes No

Date of birth

Yes No

Address or telephone number

Yes No

Email address, IP address or other online identifier

Yes No

Photographs or video recordings of people

Yes No

Sound recordings of people

Yes No

Chat with us on weekdays from 12-14

GPS data or other geolocation data (electronic communications)

Yes No

Background data that can identify a person

Yes No

Genetic data

Yes No

Biometric data

Yes No

Other data that can identify a person

Yes No

You have indicated that no personal data will be processed in the project.
If you will only be processing anonymous data you should not notify your project. Anonymous data are data where individual persons are not/no longer identifiable; not directly, indirectly or via email/IP address or scrambling key.

Note that this is not a formal assessment but is guidance based on the answers you have given above.

Continue

Appendix II: Data Management & Analysis Code

The following syntax code is presented for reproducibility of the findings. All data analysis were conducted using R version 4.0.3. (R Core Team, 2020).

```
setwd( "/Users/jiyounkim/Desktop/Thesis/Data")

# Load and attach add-on packages
library(apaTables)
library(mirt)
library(ggplot2)
library(dplyr)
library(broom)

#Import Data for South Korean in PISA 2012
example1<-read.csv("/Users/jiyounkim/Desktop/Thesis/Data/2012_Korea.csv")

# Extract necessary variables
# 8 ; StIDStd
# 134:137 ; ST53Q01, ST53Q02, ST53Q03, ST53Q04 ; Learning Strategy Items
# 501:505 ; PV1MATH, PV2MATH, PV3MATH, PV4MATH, PV5MATH ; Plausible
    values of Mathematics Achievement
example<-example1[c(8,134:137,501:505)]
# Rename Item names
example<-rename(example, Item1 = ST53Q01,Item2 = ST53Q02,Item3 = ST53Q03,Item4 =
    ST53Q04)

# Recode each learning strategy statement as numbers
# 3:metacognitive strategy 2:elaboration strategy 1:memorization strategy 9:NA
example[,c(2)] <-
  ifelse(example[,c(2)] == "Most important", 3,
        ifelse(example[,c(2)] == "relating to known", 2,
              ifelse(example[,c(2)] == "by heart",1, 9)))
```

```
example[,c(3)] <-
  ifelse(example[,c(3)] == "Improve understanding", 3,
         ifelse(example[,c(3)] == "new ways", 2,
                ifelse(example[,c(3)] == "check memory", 1, 9)))

example[,c(4)] <-
  ifelse(example[,c(4)] == "learning goals", 3,
         ifelse(example[,c(4)] == "Relating to other subjects", 2,
                ifelse(example[,c(4)] == "in my sleep", 1, 9)))

example[,c(5)] <-
  ifelse(example[,c(5)] == "more information", 3,
         ifelse(example[,c(5)] == "everyday life", 2,
                ifelse(example[,c(5)] == "Repeat examples", 1, 9)))

# Cleaning the data as deleting the response which has a NA or more.
example=na.omit(example)

# Create Learning strategy column
thedata<-example[2:5]
for(i in 1:ncol(thedata)){
  thedata[,i] <- as.factor(thedata[,i])
}

for (i in 1:nrow(thedata)) {
  thedata[i,"Memorization"] <- table(thedata[i,1:4]==1)["TRUE"]
}

for (i in 1:nrow(thedata)) {
  thedata[i,"Elaboration"] <- table(thedata[i,1:4]==2)["TRUE"]
}
```

```
for (i in 1:nrow(thedata)) {
  thedata[i,"Metacognitive"] <- table(thedata[i,1:4]==3)["TRUE"]
}

thedata[is.na(thedata)]<-0



# Add Learning strategy columns to original data(i.e., example)
example<-cbind(example,thedata)
example<-example[,c(1:10,15:17)]




################# RESEARCH QUESTION 1 #################
# Model Comparison
# 1. NOMINAL RESPONSE MODEL (NRM)
myNOMINAL<- mirt(example[2:5],model=1, itemtype="nominal",SE = TRUE, verbose =F)
coef(myNOMINAL,IRT=TRUE) # return a and b
coef(myNOMINAL, IRT=FALSE)  #return a and d (Note that d = -a*b)

# 2. Generalized Partial Credit Model (GPCM)
myGPCM<- mirt(example[2:5],model=1, itemtype="gpcm",SE = TRUE, verbose =F)
coef(myGPCM,IRT=TRUE) # return a and b
coef(myGPCM, IRT=FALSE) #return a and d (Note that d = -a*b)

# Compare TWO MODELS
## Criteria 1. Fit Index

### Fit Index of NRM
extract.mirt(myNOMINAL, "AIC")
extract.mirt(myNOMINAL, "BIC")
```

```
### Fit index of GPCM
extract.mirt(myGPCM, "AIC")
extract.mirt(myGPCM, "BIC")

anova(myNOMINAL,myGPCM)

## Criteria 2. Empirical reliability
### NRM
EAP.est <- fscores(object = myNOMINAL, method = "EAP", full.scores.SE = T)
empirical_rxx(EAP.est)
### GPCM
EAP.est.g <- fscores(object = myGPCM, method = "EAP", full.scores.SE = T)
empirical_rxx(EAP.est.g)

## Criteria 3. Item Characteristic Curve
### NRM
plot(myNOMINAL, type = "trace")

### GPCM
plot(myGPCM, type = "trace")

################# RESEARCH QUESTION 2 #################

# Data Preparation
## Create Learning Strategy score from NRM ; Use Ability Estimation
fscores<-fscores(myNOMINAL, plausible.draws=5,se=TRUE)
example<-cbind(example,fscores)
names(example)[14:18]<-c("LS1","LS2","LS3","LS4","LS5")

# 1. Linear Relation between Learning Strategy (LS) Scores and Mathematics Scores
## (1) Correlation between LS scores from NRM and Mathematics Scores
```

```
round(cor(example[,c(6:10,14:18)],use = "pairwise.complete.obs"),2)
apa.cor.table(example[,c(6:10,14:18)],filename =
    "cor_LS(NRM)_MATH.doc",table.number=1)



## (2) Correlation between Raw scores of LS and Mathematics Scores
round(cor(example[,c(6:13)],use = "pairwise.complete.obs"),2)
apa.cor.table(example[,c(6:13)],filename = "cor_LS(raw)_MATH.doc",table.number=1)

# 2. Nonlinear Relation between LS scores from NRM and Mathematics Scores

#Preparation
#regCoef
regCoef<-function(m,data,q=.95,digits=3){ #q:confidence interval
  sm = summary(m)
  ci = confint(m, level=q)
  D = cbind(m$model[,1],model.matrix(m,data)[,-1])
  S = cov(D,use="complete.obs")
  Sd = sqrt(diag(S))
  BetaZ = coef(m)*c(0,Sd[-1]/Sd[1])
  R = cov2cor(S)
  Tol.x = c(NA,1/diag(qr.solve(R[-1,-1])))
  Ry.x = c(NA,R[-1,1])
  P = qr.solve(R)
  S.P = diag(1/sqrt(diag(P)))
  Ryx.x = c(NA,-(S.P%*%P%*%S.P)[-1,1])
  Ry.x.x = BetaZ*sqrt(Tol.x)
  tab = cbind(sm$coefficients,ci,BetaZ,Tol.x,Ry.x,
            Ryx.x,Ry.x.x)
  colnames(tab)=c("b","SE","t","p","[b.l,","b.u]","
  b_Z","Tol","Ryx","Ryx.x","Ry(x.x)")
  return(round(tab,digits))
```

```
}


## Quadratic regression models
m1=lm(PV1MATH~LS1+I(LS1^2), example3)
summary(m1)
regCoef(m1,example3)

m2=lm(PV2MATH~LS2+I(LS2^2), example3)
summary(m2)
regCoef(m2,example3)

m3=lm(PV3MATH~LS3+I(LS3^2), example3)
summary(m3)
regCoef(m3,example3)

m4=lm(PV4MATH~LS4+I(LS4^2), example3)
summary(m4)
regCoef(m4,example3)

m5=lm(PV5MATH~LS5+I(LS5^2), example3)
summary(m5)
regCoef(m5,example3)


## Linear Models
m01=lm(PV1MATH~LS1, example3)
summary(m01)
regCoef(m01,example3)

m02=lm(PV2MATH~LS2, example3)
summary(m02)
```

```
regCoef(m02,example3)

m03=lm(PV3MATH~LS3, example3)
summary(m03)
regCoef(m03,example3)

m04=lm(PV4MATH~LS4, example3)
summary(m04)
regCoef(m04,example3)

m05=lm(PV5MATH~LS5, example3)
summary(m05)
regCoef(m05,example3)


### To show quadratic regression model is better (BIC)

# LS1 - PV1MATH
glance(m01)
glance(m1)

# LS2 - PV2MATH
glance(m02)
glance(m2)

# LS3 - PV3MATH
glance(m03)
glance(m3)

# LS4 - PV4MATH
glance(m04)
glance(m4)
```

```
# LS5 - PV5MATH
glance(m05)
glance(m5)

### Scatter Plot for both linear and quadratic relation
ggplot(example,
       aes(x = LS1, y = PV1MATH)) +
  geom_point(size=0.5) +
  geom_smooth(method = "lm",
              aes(color = "linear"),
              se = FALSE) +
  geom_smooth(method = "lm",
              formula = y~x+I(x^2),
              aes(color = "quadratic"),
              se = FALSE) +
  theme_bw()
```