What TIMSS Tells Us about Instructional Practice in K-12 Mathematics Education

Dabae Lee & Yeol Huh
Indiana University, United States

Abstract

Effort to determine teachers' effects on student has been continuously made with national data. However, paucity of research has been conducted on how teachers' instructional strategies impact on student learning with national data, although instructional theories suggest a direct relationship between instructional strategies and learning outcomes. Therefore, the relationship between teachers' use of instructional strategies and learning outcomes should be examined with national data. This study investigates how much teacher's instructional strategies explain student learning in mathematics and what instructional strategies are positively related to student learning outcomes. Revised Bloom's taxonomy was used to define instructional strategies that support different levels of cognitive processes. The U.S. 8th grade mathematics data from the 2007 Trends in International Mathematics and Science Study was analyzed using multilevel modeling. As results, teachers' instructional strategies explained approximately 12% at the individual level and 17% at the teacher level of the learning outcome. Also, asking student to write equations and functions to represent relationships and to decide on their own procedures for solving complex problems were positively and significantly related to student learning outcomes.

Keywords: *Instructional strategies; TIMSS; Mathematics education; Academic achievement; Multilevel modeling*

Introduction

Mathematical thinking, literacy, and skills have long received a heavy emphasis in K-12 education in the U.S. The Act of No Child Left Behind mandates all states to periodically assess and report students' academic performance in mathematics as well as in reading and writing. However, the academic performance of the U.S. students has not shown much improvement over time. Especially, academic performance of 8th grade students has been around the international average since 1995, in international assessments including The Trends in International Mathematics and Science Study (TIMSS) (National Center for Education Statistics, n.d.).

Policymakers' interest in reforming teaching has created a demand for data on K-12 teachers' instructional practices and their impacts on student learning. Effort to determine teachers' effects on student has been continuously made with national data (e.g., Chetty, Friedman, & Rockoff,

2011; D. P. Mayer, 1999; National Mathematics Advisory Panel, 2008; Rowan, Correnti, & Miller, 2002). However, such effort has been focused on teachers' content knowledge, academic background, professional experience, and professional development activities. There is little research conducted on how teachers' instructional strategies impact on student learning with national level data, although instructional theories suggest that there is a direct relationship between instructional strategies and learning outcomes (Merrill & Boutwell, 1973; Merrill, Olsen, & Coldeway, 1976; Merrill, Tennyson, & Posey, 1992; Merrill & Wood, 1974; Reigeluth, 1999; Reigeluth & Merrill, 1979). Therefore, it is important to examine how teachers' instructional strategies in their classrooms are related to student learning outcomes using national level data.

TIMSS provides data on academic achievement in mathematics and science of 4th and 8th grade students with teachers' instructional practices that are internationally comparable. In this study, we examined how much teachers' instructional practices explain learning outcomes in mathematics of the U.S. 8th grade students and how each instructional strategy is related to academic achievement.

Literature Review and Theoretical Framework

Numerous scholarly efforts were made in examining teachers' effects on student learning outcomes. For example, Borich (1996) examined how teachers' high expectations of their pupils impact student learning and found that teachers' high expectations resulted in improved academic achievement. Several scholars examined how effective teachers managed their classrooms with minimized disruptions (Brophy & Good, 1986; Evertson, Anderson, Anderson, & Brophy, 1980; Griffin and Barnes, 1986; Lampert, 1988; Secada, 1992).

More importantly, a number of studies examined how effective teachers used their class time. The results were consistent that effective teachers committed their class time more to teacher-led discussion rather than to students' individual work time on their own (Borich, 1996; Brophy, 1986; Evertson et al., 1980; Good, Grouws, DeWayne, Slavings, & Cramer, 1990; Good, Grouws, & Ebmeier, 1983; Walberg; 1986; Mason & Good, 1993;). Furthermore, these teachers structured the learning material in an effective way and presented the material in an active way by asking many questions to students and eliciting a great deal of participation and discussion from students (Brophy, 1986; Brophy & Good, 1986; Borich, 1996; Good, Grouws, & Ebmeier, 1983). Given that teacher-led discussions are a feature of effective teaching, the question is what specific instructional strategies should be used, what types of learning teachers need to promote in mathematics, and how they can allocate their class time to each strategy to maximize learning outcomes.

Discussion of what instructional strategies should be used is inherently linked with what type of learning teachers need to promote. R. E. Mayer (2002) has distinguished meaningful learning from rote learning. Rote learning occurs when learners remember important information but are unable to use the information. Meaningful learning occurs when learners possess relevant knowledge and are able to use that knowledge to solve problems (R. E. Mayer, 2002). That is, rote learning concerns retention; and meaningful learning concerns transfer. Transfer requires learner to

achieve a higher level of understanding on the revised Bloom's taxonomy (Krathwohl, 2002). R. E. Mayer (2002) has illustrated cognitive processes for retention and transfer based on the revised Blooms' taxonomy, which was used as a framework in this study. This was selected because its close resemblance to the survey items about instructional strategies used in TIMSS.

In the revised Bloom's taxonomy, there are six levels of cognitive processes: remember, understand, apply, analyze, evaluate, and create, each of which is a prerequisite to the next level (Krathwohl, 2002). For example, *remembering* is necessary to *understanding*. According to R. E. Mayer (2002), *remembering* is closely related to retention and becomes a means to the other five cognitive processes, and the other five are related to transfer in an increasing order. The six cognitive processes can be further broken down into 19 sub-processes as in Table 1. For example, remembering involves recognizing and recalling relevant information.

In line with rote learning versus meaningful learning, Muijs and Reynolds (2010) have claimed that both rote learning and meaningful learning are important in learning mathematics. Students need to "master the conventional systems of math and gain automaticity in the use of skills such as multiplication facts and times tables" in order to work efficiently for more meaningful work (p. 260). However, they have warned that an overemphasis on rote learning may lead to difficulty in transferring students' knowledge's to other situations.

The six cognitive processes of the revised Bloom's taxonomy resemble how the TIMSS measures student academic achievement in mathematics. Therefore, the revised Bloom's taxonomy has been adopted as a theoretical framework for this study. Table 1 presents the revised Bloom's taxonomy in relation to the TIMSS assessment framework. The TIMSS mathematics test measured content knowledge on number, algebra, geometry, data and chance for three cognitive domains: knowing, applying, and reasoning (Mullis et al., 2005). *Knowing* is equivalent to remembering and classifying under understanding in the revised Bloom's taxonomy, because knowing "covers the facts, procedures, and concepts students need to know" (Mullis et al., 2005, p. 33), and it covers behaviors such as recalling, recognizing, computing, retrieving, measuring, and classifying. *Applying* corresponds to understanding and applying in the revised Bloom's taxonomy, because "applying focuses on the ability of students apply knowledge and conceptual understanding to solve problems" (Mullis et al., 2005, p. 33), and it covers behaviors such as selecting, representing, modeling, and implementing and solving routine problems. *Reasoning* corresponds to implementing an unfamiliar task under applying and analyzing, and evaluating, because reasoning "encompasses unfamiliar situations, complex contexts, and multi-step problems" (Mullis et al., 2005, p. 33), and it covers behaviors such as analyzing, generalizing, synthesizing, justifying, and solving non-routine problems in real life contexts. The TIMSS does not measure the student' ability at the Create level, because it is not feasible to measure the students' ability to *create* something using standardized tests.

Table 1. Cognitive Processes based on the Revised Bloom's Taxonomy & TIMSS Assessment Framework

Revised Bloom's taxonomy					
Rote learning	Meaningful learning				
Remember	Understand	Apply	Analyze	Evaluate	Create
Recognizing Recalling	Interpreting Exemplifying Classifying Summarizing Inferring Comparing Explaining	Executing Implementing	Differentiating Organizing Attributing	Checking Critiquing	Generating Planning Producing
TIMSS assessment framework					
Knowing	Applying		Reasoning		
Recalling Recognizing Computing Retrieving Measuring Classifying	Selecting Representing Modeling Implementing Solving routine problems		Analyzing Generalizing Synthesizing Justifying Solving non-routine problems		N/A

As discussed above, types of learning are closely and inherently related to instructional strategies. In TIMSS, teachers were surveyed what instructional strategies were used in promoting the three types of learning in their classrooms: Knowing, Applying, and Reasoning. Table 2 presents the instructional strategy items and corresponding types of learning or cognitive process. These instructional strategies are micro-level strategies to support a particular cognitive process that are specific to mathematics education. For example, in supporting *Knowing,* teachers may have students memorize formulas and procedures, practice the four fundamental arithmetic operations without a calculator, or work on fractions and decimals.

Table 2. The TIMSS Cognitive Domains and Corresponding Instructional Strategy Items in TIMSS

Cognitive processes	Instructional strategy items in TIMSS
Knowing	• Memorize formulas and procedures. • Practice adding, subtracting, multiplying, and dividing without using a calculator. • Computing: Work on fractions and decimals.
Applying	• Use knowledge of the properties of shapes, lines and angles to solve problems. • Interpret data in tables, charts, or graphs. • Write equations and functions to represent relationships. • Apply facts, concepts and procedures to solve routine problems.

Reasoning	• Decide on their own procedures for solving complex problems.
	• Work on problems for which there is no immediately obvious method of solution.
	• Relate what they are learning in mathematics to their daily lives.

In this study, we investigated how instructional strategies that support students' cognitive processes are related to student academic achievement in mathematics using the TIMSS 2007 8th grade mathematics dataset. The research questions are:

(1) How much teachers' such instructional strategies explain student learning in mathematics?

(2) Which instructional strategies that support the cognitive processes are positively related to student learning in mathematics?

Methods

Dataset

TIMSS 2007 U.S. 8th grade mathematics dataset was used. In TIMSS, two-stage probability proportional-to-size sampling was used; and the sample was weighted to be representative of the nation (Joncas, 2008). The coverage of the U.S. target population was 100%, and there were 0% of school-level exclusions, 7.9% of within-sample exclusions, and 7.9% of overall exclusions. 239 private and public schools and 7,377 students participated at grade eight (Joncas, 2008). 532 teachers participated.

Variables

In the TIMSS 2007, 12 items were dedicated to investigate teachers' instructional practices as shown in Table 2. Based on the description of the TIMSS assessment framework (Mullis et al., 2005), the items were categorized into the three cognitive domains. Three of them were related to knowing. Four of them were related to applying. Three of them were related to reasoning. The remaining two items, "explain their answers" and "work together in small groups" were not included because they were not specific enough to determine the cognitive domain that they are concerned with.

Table 3 presents the variable names and the corresponding items. Student mathematics scores were used as the outcome variable. According to the description of TMSS international mathematics benchmarks (as cited in Gonzales et al., 2008), the higher the scores, the higher levels of understanding were required. For example, students in the advanced category (625 and above) were able to apply their knowledge and solve highly complex, non-routine problems. Those in the low category (400 and below) demonstrated some mathematical knowledge but were not

able to apply their knowledge. In the TIMSS, a complex rotated booklet design was used (Ruddok, O'Sullivan, Arora, & Erberer, 2008). In the design, only a portion of the total test was administered and treated the other portion as missing data. Multiple imputations were employed to generate five plausible scores for individual student. The average values of the five scores were used.

The ten items above from the teacher questionnaire were used as predictor variables: three items for knowing, four items for applying, and three items for reasoning. The items asked how often teachers asked students to do the corresponding instructional strategies in the TIMSS class. The response categories were: 1) every or almost every lesson, 2) about half the lessons, 3) some lessons, and 4) never. The responses were reverse coded, so the higher number, the more frequent instructional strategies were implemented. Because there were more than one mathmatics teachers for individual student, we used average values of teachers' responses for each predictor as suggested by Snijders and Bosker (2012a) and assigned classification identification number to student groups who shared the same combination of teachers. There were 517 groups that share the same teachers.

Table 3. Variable Names and Description

Category	Var. Name	Description
	Math	Average values of plausible mathematics scores
Knowing	Memorize	Memorize formulas and procedures.
	Practice	Practice adding, subtracting, multiplying, and dividing without using a calculator.
	Fraction	Work on fractions and decimals.
Applying	Use	Use knowledge of the properties of shapes, lines and angles to solve problems.
	Interpret	Interpret data in tables, charts, or graphs.
	Write	Write equations and functions to represent relationships.
	Apply	Apply facts, concepts and procedures to solve routine problems.
Reasoning	Decide	Decide on their own procedures for solving complex problems.
	Workon	Work on problems for which there is no immediately obvious method of solution.
	Relate	Relate what they are learning in mathematics to their daily lives.

Data Analysis

Multilevel modeling (Snijders & Bosker, 2012b) was used in order to partition the outcome variance into the student and teacher levels.

Exploratory Analyses

We first performed exploratory analyses to investigate how the teacher level predictors were related to student academic achievement. As a result, we found that predictors in a cognitive domain have different slopes, which suggest different effects on student learning depending on teaching strategies.

Model Comparison

With the results from exploratory analyses in mind, we fitted two models: 1) Model 1, where variables within a domain were combined, accordingly, there were three predictors, knowing, applying, and reasoning, and 2) Model 2, variables were not combined, accordingly there were 10 predictors, and compared their AIC values. The AIC value of Model 1 was 71,962, and that of Model 2 was 71,115, which was smaller by 847. This suggests that Model 2 is a better fitting model. Therefore, we decided to use Model 2.

Final Model Specification

The following model was fitted. In level 1, there are a group varying intercept and level 1 residual. In level 2, there are a population intercept, ten predictors, and level 2 residual. The residuals are assumed to be normally distributed. The means of the level 1 and level 2 residuals are zero; and their variances, σ^2 and τ_0^2, respectively, are assumed to be constant.

Level 1: Student level
$$math_{ij} = \beta_{0j} + R_{ij}$$
Level 2: Teacher level
$$\beta_{0j} = \gamma_{00} + \gamma_{01}memorize + \gamma_{02}practice + \gamma_{03}fraction$$
$$+\gamma_{04}use + \gamma_{05}interpret + \gamma_{06}write + \gamma_{07}apply$$
$$+\gamma_{08}decide + \gamma_{09}workon + \gamma_{10}relate + U_{0j}$$

Checks for Assumptions

After fitting the final model, we inspected level 1 and level 2 residuals to check if there is any evidence that suggests violation of the normality, linearity, and homoscedasticity assumptions.

Normality

Inspecting QQ plot of level 1 residuals in Figure 1, although there are a few over and under dispersed cases in the tails, these residuals seem quite normal. However, looking at QQ plot of level 2 in Figure 2, there is some evidence of departure from normality in the intercepts.
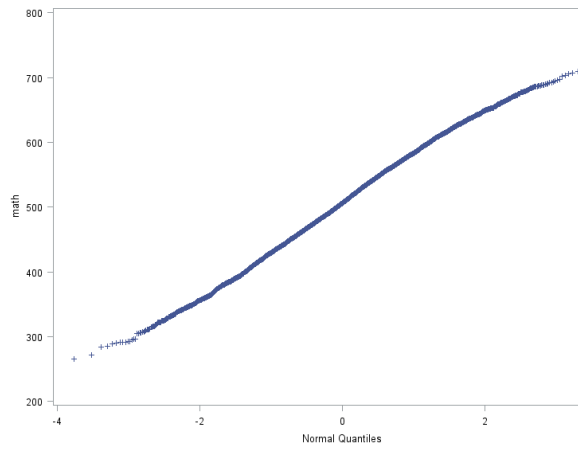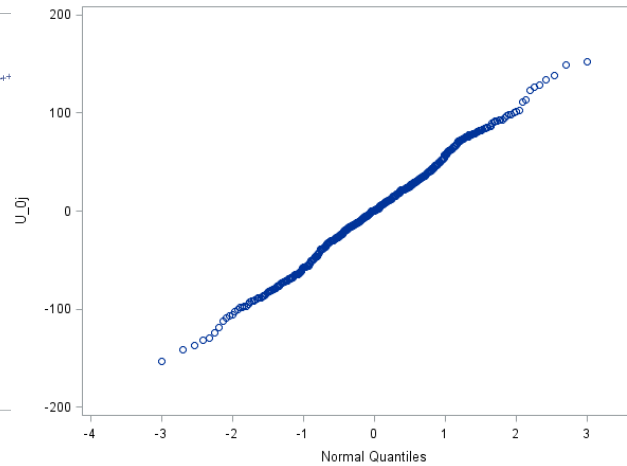
Figure 1. QQ Plot of Level1 Residuals

Figure 2. QQ Plot of Level 2 Residuals

Linearity and Homoscedasticity

We examined the level 2 residuals against each predictor by inspecting box plots. The level 2 residuals did not have discernable patterns or strong evidence of heterogeneity against most of the predictors. However, the box plots of the use and interpret predictors in Figure 3 and Figure 4 respectively showed some curve linear patterns, which suggests violation of linearity. Also, as shown in Figure 5, the box plot of the write predictor showed a corn-shape distribution of the level 2 residuals, which suggests violation of homoscedasticity.
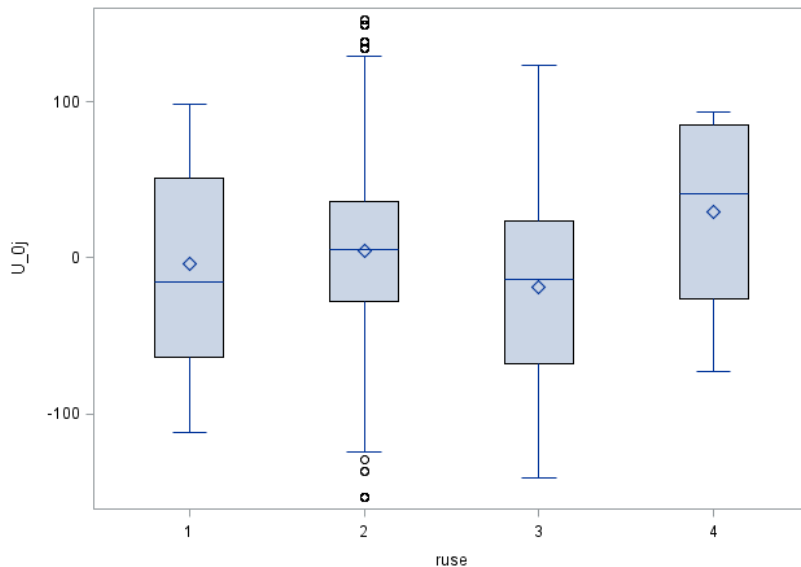

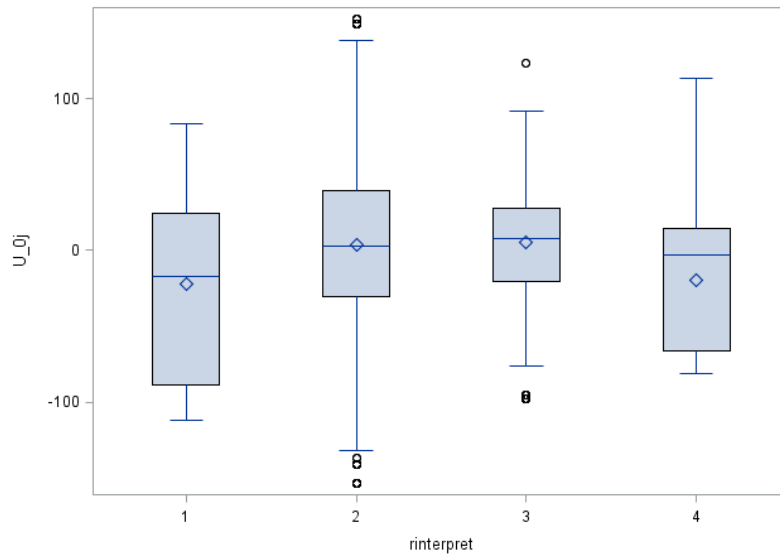Figure 3. Box Plot of Level 2 Residuals and the Use Variable

Figure 4. Box Plot of Level 2 Residuals and the Interpret Variable
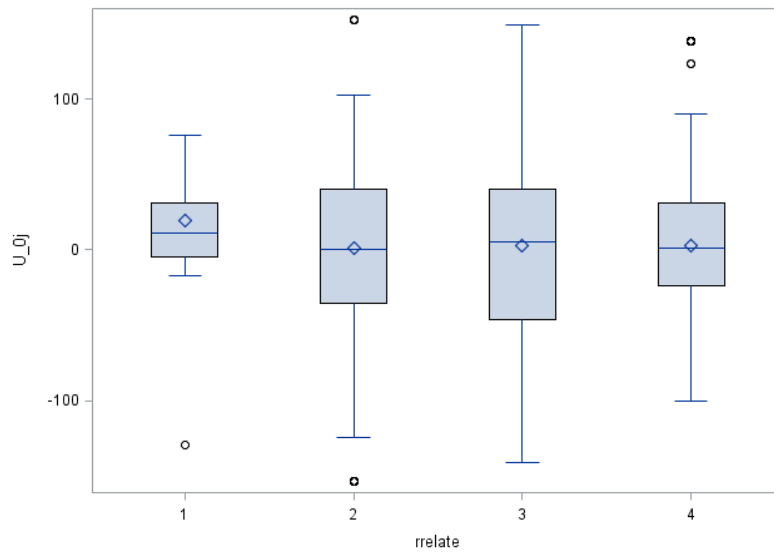

Figure 5. Box Plot of Level 2 Residuals and the Relate Variable

Results

Fixed Effects and Random Effects

Table 4 presents the results of the fixed and random effects of the final model. The fixed effects of practice, write, decide, and relate were statistically significant. The write and decide variables were positively related to student mathematics scores. One-unit changes in write and decide resulted in 26.82 and 9.04 score higher in mathematics scores respectively. The practice and relate effects were negatively related to student mathematics scores. One-unit changes in practice and relate resulted in 7.32 and 10.28 score lower in mathematics scores respectively. None of the other fixed effects were statistically significant. In teacher level, 3,209.98 of outcome variance remained unaccounted by the model. In student level, 2,007.04 of outcome variance remained unaccounted.

Table 4. Fixed and Random Effects

Fixed effects			
Parameter	Estimate	SE	Pr > |t|
γ_{00} = Intercept	436.58***	19.67	<.0001
γ_{01} = memorize	0.00	3.98	0.999
γ_{02} = practice	-7.32**	2.96	0.014
γ_{03} = fraction	1.37	4.48	0.760
γ_{04} = use	10.39	5.79	0.073
γ_{05} = interpret	-2.98	6.20	0.631
γ_{06} = write	26.82***	4.34	<.0001
γ_{07} = apply	0.88	3.90	0.822
γ_{08} = decide	9.04*	4.56	0.048
γ_{09} = workon	2.12	4.57	0.642
γ_{10} = relate	-10.28**	3.75	0.006
Random effects			
Parameter	Estimate	SE	Pr > Z
τ_0^2 = Level 2 residual	3209.98***	227.10	<.0001
σ^2 = Level 1 residual	2007.04***	36.05	<.0001

Note. ***p<0.001. **p<0.01. *p<0.05.

Pseudo R Square

In order to examine how much teachers' instructional strategies explain student learning in math, pseudo R square values in level 1 and level 2 were calculated based on the values in Table 5. For

group sample sizes, we used harmonic mean, 9.89. As a result, the model explained 11.54% of the outcome variance at the student level and 16.54% of the outcome variance at the teacher level.

Table 5. Random Effects of Null and Final Model

Null Model			Final Model (Random Intercept)		
Parameter	Estimate	SE	Parameter	Estimate	SE
Level 2 residual	3891.56***	257.84	Level 2 residual	3209.98***	227.10
Level 1 residual	2006.36***	34.30	Level 1 residual	2007.04***	36.05

Note. ***p<0.001.

Discussion

RQ1. How Much Teachers' Instructional Strategies Explain Student Learning in Math?

Teachers' instructional strategies measured by the ten items from TIMSS teacher background questionnaire explained approximately 12% and 17% of the student mathematics achievement. This means that the teachers' instructional strategies account for 29% of variance of academic achievement in mathematics. The proportion would be even larger, because the ten items may not represent all of the possible instructional strategies. Also, in the present study, we used static mathematics scores, however, Rowan et al. (2002) suggested that "if one really wants to assess the size of teacher effects on changes in student achievement, models of annual gains in achievement are preferable" (p. 1532), showed a substantial increase in the student achievement variance explained by teachers when annual gains in achievement were used as an outcome variable.

This result emphasizes the importance of choosing effective instructional strategies in that about a third of variance of student learning outcome can be attributed to teachers' instructional strategies. Given that the result comes from a nationally represented sample, this finding can be generalized to the 8th grade students in the U.S. Furthermore, the implications of this result go beyond educational practice. Especially, the result provides empirical evidence that supports how important it is to carefully choose instructional strategies when designing instructional systems.

RQ2. Which Instructional Strategies Are Positively Related to Student Learning?

Results from the second research question provide specific guidelines as to what instructional strategies are effective in teaching and learning mathematics in 8th grade classrooms. Among the three knowing variables, the practice variable was significant and negatively related to mathematics achievement, and the fraction and memorize variables were not significant. As R. E. Mayer (2002) claimed, memorizing itself did not result in increase in student academic achievement. Among the four applying variables, the write variable was significant and positively related to mathematics achievement. Therefore, asking student to write equations and functions

to represent relationships resulted in increase in student mathematics scores. Among the reasoning variables, the decide effect was significant and positive, and the relate effect was significant and negative. Therefore, asking students to decide on their own procedures for solving complex problems resulted in increase in student mathematics scores.

In sum, asking students to write equations and functions to represent mathematical relationships and letting them decide on their own procedures for solving complex problems were significantly effective than other strategies. These findings are consistent with classroom practice illustrated by several researchers. According to their studies, effective teachers have placed a great emphasis on interactivity by asking questions and eliciting students' participation (Borich, 1996; Brophy, 1986; Brophy & Good, 1986; Good, Grouws and Ebmeier, 1983). On the other hand, too much emphasis on rote learning such as practicing the four arithmetic functions can even have negative effects on student learning as Muijs and Reynolds (2010) have warned.

The present study adds to the knowledge base that asking students write equations and decide on their own procedures can help students better learn, whereas letting students blindly practice the four arithmetic functions may not be effective. These findings advise teachers to allocate enough time on the instructional activities of asking students write equations and decide on their own procedures, but not to spend too much time on practicing the four arithmetic functions. Furthermore, when designing instructional systems for mathematics, the two effective strategies should be incorporated to create effective learning environments.

Asking students to relate what they are learning in mathematics to their daily lives resulted in decrease in student mathematics scores. However, caution should be used when interpreting the negative impact of the relate variable. Brown, Collins, and Duguid (1989) claimed that when learning is situated in real life, students learn more effectively and easily apply their knowledge. However, in order to make learning situated, a set of complex, inter-related instructional strategies should be executed.

Educators have devised several instructional strategies to make learning more situated such as project-based learning (Bell, 2010) and problem-based learning (Hmelo-Silver, 2004). Substantial research has supported effectiveness of such instructional strategies (Cindy, Duncan, & Clark, 2007; Kirschner, Sweller, & Clark, 2006; Ravitz, 2009; Strobel & Van Barneveld, 2009; Walker & Leary, 2009).

The conflicting result of the relate effect can be explained by possible validity problems of the questionnaire item. Given that project-based learning or problem-based learning consists of several interrelated instructional sub-methods, it is hard to measure such practice with one item. D. P. Mayer (1999) claimed that individual indicator could misguide in measuring instructional practice, instead composite indicators should be used. The violation of homoscedasticity of the relate variable supports this claim.

Limitations and Future Research Directions

Cross-sectional observational data were utilized in the study. Unlike the case-control studies, causal inferences should not be made based on the results of the current study. Also, there existed some evidence of violation of normality, linearity, and homoscedasticity assumptions. Therefore, results should be interpreted with those issues in mind.

The current study suggests mainly two future research directions. First, effort to create a valid and reliable measure of instructional practice should be. As explained earlier, some of the TIMSS items on instructional practice were not sufficient to measure teachers' instructional practice. Therefore, more research should be conducted to create such measure to make an informed policy decision in reforming teaching.

In addition, it should be investigated how teachers' knowledge in the subject areas and teaching experience interplay with the instructional strategies that teachers choose. Shulman (1987) has emphasized teachers' pedagogical content knowledge, which is "special amalgam of content and pedagogy" identifies "the distinctive bodies of knowledge for teaching" (p. 8). In order for teachers to form pedagogical content knowledge and choose effective instructional strategies, teachers need to understand the content and to have practiced teaching to let content and pedagogical knowledge intertwine. Therefore, teachers' content knowledge and teaching experience should be examined with instructional practice.

References

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *Clearing House, 83*(2), 39-43. doi: 10.1080/00098650903505415

Borich, G. (1996) *Effective teaching methods* (3rd ed.). New York : Macmillan.

Brophy, J. (1986) Teaching and learning mathematics: Where research should be going. *Journal for Research in Mathematics Education*, 17(5), 323-346.

Brophy, J. & Good, T. L. (1986) Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching.* New York : MacMillan.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood (Working Paper No. 17699).* National Bureau of Economic Research website: http://www.nber.org/papers/w17699

Cindy, E., Duncan, R.G., & Clark, A.C. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107.

Evertson, C. M., Anderson, C. W., Anderson, L. M., & Brophy, J. E. (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and english classes. *American Educational Research Journal*, *17*(1), 43-60.

Good, T. L., Grouws, D. A., DeWayne, A. M., Slavings, R. J., & Cramer, C. (1990) An observational study of small-group mathematics instruction in elementary school. *American Educational Research Journal, 27*(4), 735-782.

Good, T. L., Grouws, D. A., & Ebmeier, H. (1983) *Active mathematics teaching.* New York: Longman.

Griffin, G. A. & Barnes, S. (1986). Using research findings to change school and classroom practice: Results of an experimental study. *America Educational Research Journal*, *23*(4), 572-586.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235-266.

Joncas, M. (2008). TIMSS 2007 sampling weights and participation rates. In J. Olson, M. Martin & I. Mullis (Eds.), *TIMSS 2007 Technical Report*. Chestnut Hills, MA: TIMSS & PIRL International Study Center, Boston Collge.

Kirschner, P.A., Sweller, J., & Clark, R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Lampert, M. (1988). What can research on teacher education tell us about improving quality in mathematics education? *Teaching and Teacher Education*, *4*(2), 157-170.

Mason, D. A. & Good, T. L. (1993) Effects of two-group and whole- class teaching on regrouped elementary students' mathematics achievement. *American Educational Research Journal, 30*(2), 328-360.

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*(1), 29-45.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice, 41*(4), 226-232.

Merrill, M. D. & Boutwell, R. C. (1973). Instructional development: Methodology and research. In F. N. Kerlinger (Ed.), *Review of research in education* (pp. 95-131). Itasca, IL: Peacock Publishers.

Merrill, M. D., Olsen, J. B, & Coldeway, N. A. (1976). *Research support for the instructional strategy diagnostic profile*. Provo, UT: Courseware Incorporated.

Merrill, M. D., Tennyson, R. D, & Posey, L. O. (1992). *Teaching concepts: An instructional design guide*. Englewood Cliffs, NJ: Educational Technology Publications.

Merrill, M. D. & Wood, N. D. (1974). *Instructional strategies: A preliminary taxonomy*. Columbus, OH: Ohio State University.

Muijs, D. & Reynolds, D. (2010). *Effective teaching: Evidence and practice* (3rd ed.). London, UK: Sage.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberer, E. (2005). *TIMSS 2007 assessment framework*. Chestnut Hill, MA: TIMSS& PIRLS International Study Center, Lynch School of Education, Boston College.

National Center for Education Statistics. (n.d.). *Trends in International Mathematics and Science Study (TIMSS)*. Retrieved from https://nces.ed.gov/TIMSS/results11_math11.asp.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Ravitz, J. (2009). Introduction: Summarizing findings and looking ahead to a new generation of PBL research. *Interdisciplinary Journal of Problem-based Learning, 3*(1), 2.

Reigeluth, C. M. (1999). What is instructional-design theory and how is it changing? . In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 5-29). Mahwah, NJ: Lawrence Erlbaum.

Reigeluth, C. M. & Merrill, M. D. (1979). Classes of instructional variables. *Educational Technology, 19*, 5-24.

Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. *Teachers College Record, 104*(8), 1525-1567.

Ruddok, G. J., O'Sullivan, C. Y., Arora, A., & Erberer, E. (2008). Developing the TIMSS 2007 mathematics and science assessments and scroing guides. In J. Olson, M. Martin & I. Mullis (Eds.), *TIMSS 2007 Technical Report*. Chestnut Hills, MA: TIMSS & PIRL International Study Center, Boston Collge.

Secada, W. G. (1992) Race, ethnicity, social class, language and achievement in mathematics, In D. A. Grouws, (Ed.), *Handbook of research on mathematics teaching and learning*. New York : MacMillan.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-23.

Snijders, T. A. B. & Bosker, R. J. (2012a). Imperfect hierarchies. In T.A.B. Snijders & R.J. Bosker (Eds.), *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2 ed., pp. 205-215). Los Angeles, CA: Sage.

Snijders, T. A. B. & Bosker, R. J. (2012b). *Multilevel analysis: An introduction to basic and advanced multilevle modeling* (2nd ed.). Los Angeles, CA: Sage.

Strobel, J. & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problembased Learning, 3*(1).

Walberg, H. J. (1986) Syntheses of research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching.* New York: MacMillan.

Walker, A. & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-based Learning, 3*(1), 6-28.

Correspondence: Dabae Lee, Doctoral Candidate in the Department of Instructional Systems Technology, School of Education, Indiana University, Bloomington, Indiana, USA