

Korean Perspectives on Assessment of Student Achievement

PARK, Hyun-Jeong*

In this paper, two nationwide assessments for elementary and secondary students' educational achievement in Korea were reviewed for their assessment frameworks and the overall results; the Diagnostic Test for Basic Skills (DTBS) of Grade 3 elementary students and the National Assessment of Educational Achievement (NAEA) for Grade 6, 9 and 10 students. Also, the results for Korea in two large-scale international comparison studies were reviewed; the Trends in International Mathematics & Science Study (TIMSS) and the Programme for International Student Assessment (PISA). Finally, the author makes some suggestions for future research and policies on the student achievement from the Korean perspective.

1 Introduction

The Korean government has taken initiatives on the nationwide assessments of student achievement since the late 1980s. Currently, there are two nationwide tests targeting elementary and secondary school students in Korea; one is the Diagnostic Test for Basic Skills (DTBS) of the Grade 3 elementary students and the other is the National Assessment of Educational Achievement (NAEA) for the Grade 6, 9 and 10 students. Along with these two nationwide achievement tests, Korea has participated two international surveys on student achievement since 1994; one is the Trends in International Mathematics & Science Study (TIMSS) and the other is the Programme for International Student Assessment (PISA). In this paper, I would like to review the assessment framework and the results from these assessments currently in operation, and then make suggestions for future research and policies on the student achievement from Korean perspectives.

2 Nationwide Assessments of Student Achievement in Korea

The Diagnostic Test for Basic Skills (DTBS) of the Grade 3 elementary students started in 2002. The legal foundation for this test is contained in Article 9.1 of the elementary and secondary education act, which states that the Ministry of Education can carry out tests to evaluate the achievement of students in school education. Besides this legal foundation, the DTBS is based

*Seoul National University
e-mail: hjp@snu.ac.kr

on the 7th national curriculum and the school education normalization policy (Chae et al., 2003). The 7th national curriculum in Korea states that the government can carry out evaluations of student achievement, schools and educational institutions, and planning and implementation of school curriculum periodically on the national level, in order to control the quality of national curriculum. Also, the school education normalization policy of March 2002 directly introduced the idea of testing Grade 3 elementary students on the 3R's (reading, writing, arithmetic) to diagnose their basic skills.

While this background of the introduction of DTBS can tell us much about the objectives of this test, the main objectives of the DTBS can be summarized as follows (Chae et al., 2003). First of all, the DTBS conducts a scientific evaluation of basic skills and provides necessary support to the nation, local educational authorities and schools based on the results of this test in order to maintain a high level of academic achievement. Secondly, the DTBS confirms whether students have reached the minimum proficiency level for reading, writing, and arithmetic at the Grade 3 level, and the test produces various policy-based indicators for lower-grade elementary school education. Thirdly, the DTBS develops and offers remedial education programs to support students under the minimum proficiency level based on the analysis of their characteristics.

To achieve these objectives, the Korea Institute for Curriculum and Evaluation (KICE), which is a management institution responsible for national achievement tests, has carried out the DTBS every October since 2002, and has distributed these results to students and schools the following December. The results for individual students include whether the student has reached the minimum proficiency level in each subject area (reading, writing, arithmetic), as well as diagnostic information on the sub-areas classified into abilities and content.

The other nationwide assessment utilized in Korea is the National Assessment of Educational Achievement (NAEA). The focus of this paper is on the NAEA because it has been recently in the center of educational arguments, while the DTBS has never been in the middle of controversy. The NAEA has been conducted with the nationally representative sample of Grade 6, 9, and 10 students every October since 2000. The NAEA was designed to benchmark the National Assessment of Educational Progress (NAEP) in USA. The objectives of the NAEA are as follows (Cho et al., 2007). The main objectives of the NAEA are to measure the educational achievement of elementary, middle, and high school students, and to analyze the trends of their achievement systematically and scientifically. The research design for the trend analysis of NAEA data was introduced in 2003, and the first report of trend analyses was published in 2006. In addition to these main objectives, the NAEA provides reference data to improve the national curriculum by analyzing students' achievement utilizing specific goals of the curriculum, and investigating the problems with curriculum implementation at school and classroom levels. In the process of analyzing test items and the relationship between students' achievement and their background variables, the NAEA also provides valuable information to improve teaching and learning methods, as well as necessary assistance to set up learning encouragement policies by the government.

To achieve these objectives, KICE has prepared systematically for the NAEA since 1998. KICE launched the initiating plan for the NAEA in 1998 and administered the field tests in social studies in 1999. The first nationwide administration of the NAEA was conducted with a 0.5% sample of the whole population for Grade 6, 9, and 10 students in October 2000. The sample sizes of the NAEA have increased since 2000; the samples were 1% of the 6th, 9th and 10th graders from 2001 to 2003, 1% of the 6th and 9th graders and 3% of 10th graders in 2004 and 2005, and 3% of

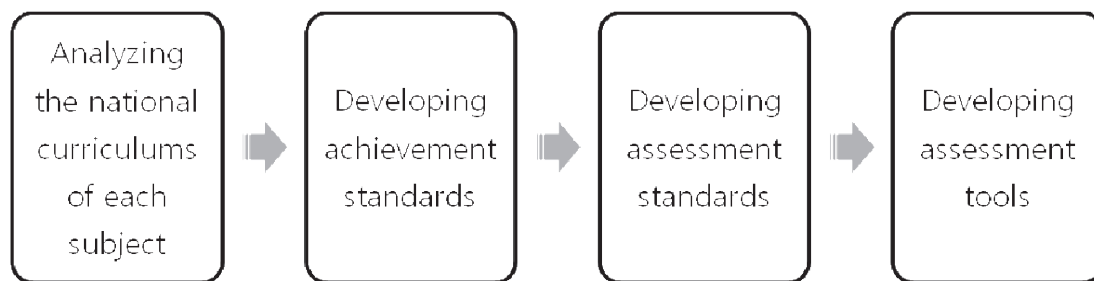


Figure 1 The process of developing the NAEA evaluation tools
 Reprinted from http://www.kice.re.kr/kice/eng/info/info_3.jsp

the 6th, 9th and 10th graders in 2006 and 2007.

For these three grades, students were tested on five subject matters, Korean language, social studies, mathematics, science, and English. Since the tests are administered in October, the tests for 6th graders include all the content covered from the 4th to 6th grades, the tests for 9th graders include all content covered from the 7th to 9th grades, and tests for 10th graders include all content covered in the 10th grade. The tests consist of constructed response items as well as multiple-choice items. The constructed response items comprise 20–40% of the total score depending on the subject matter. Also, listening comprehension tests are included in both Korean and English. Along with these achievement batteries, background questionnaires are administered to students, teachers and principals to investigate the relationship between the background variables and student academic achievement.

The process of developing NAEA evaluation tools can be summarized as follows (KICE, 2008). First, the national curriculum of each subject matter is analyzed to decide the assessment areas and to set up achievement standards for each area. The achievement standards are statements specifying the objectives and content of the national curriculum. Once the achievement standards are developed, assessment standards also need to be developed. The assessment standards are statements differentiating students' levels of achievement (four levels of Advanced, Proficient, Basic, and Below-Basic) to use as criteria in assessment activities for each subject. Finally, assessment tools are developed through a process of reviewing the achievement and assessment standards, setting a guideline for item development, appointing and training item writers, designing an item writing plan, developing items by the item writers, reviewing the items by reviewers, selecting the items for field test, conducting a field test, analyzing the results of the field test, revising and selecting items for a main test, and deciding the final assessment tools for NAEA.

The results from trend analyses of the NAEA can be summarized as follows (Cho et al., 2007). First of all, the variations of average standard scores from 2003 to 2006 were largest for the 6th graders. Especially for the English and science, the average scores showed a sharp increase until 2005 but changed directions and decreased in 2006. For the 9th graders, the average scores of English and social sciences increased in 2004, but steadily decreased afterwards, while the average scores of other subjects continued to increase until 2005 but showed a decline in 2006. Among the 10th graders, the average scores for Korean, mathematics and science fluctuated from time to time, while the average scores for English showed a slow increase from 2003 to 2005 before decreasing in 2006, and the average scores for social studies continued to slowly decline ever since 2003. In general, the average performance of students showed improvement until 2005, but decreased in 2006 for all grades. Comparing the years 2003 and 2006 the average achievement of 6th graders improved in English, science and mathematics, while 9th graders showed improvement

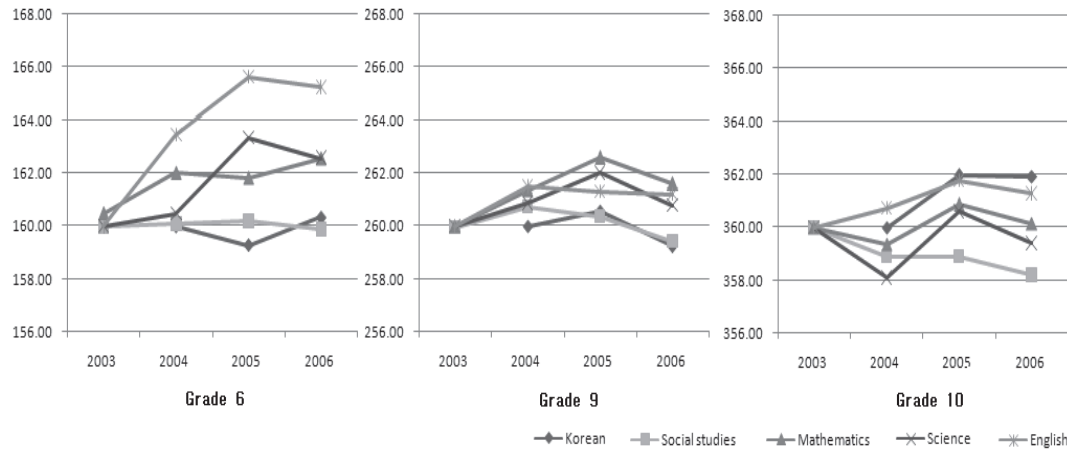


Figure 2 Changes in the average standard scores by grades and subject matters from 2003 to 2006

Note: For the subject of Korean, the referential year of the trend analysis was 2004 and not 2003. The standard deviations ranged from 7 to 11 depending on the subjects and years. The original figures from Cho et al. (2007) were re-organized by author.

in mathematics and English, and 10th graders in Korean and English. However, the average scores for social studies decreased from 2003 to 2006 for the 10th graders.

Secondly, in terms of proficiency levels, the proportion of students at the advanced level had generally increased for Grade 6 science, Grade 9 mathematics and Grade 10 Korean, while they had decreased for Grade 9 Korean, and Grade 10 Science. However, the proportion of students below basic level generally declined for Grade 6 mathematics, Grade 6 science, Grade 9 mathematics, Grade 9 science, and Grade 10 Korean. In short, there tended to be a higher proportion of 6 grade students at the advanced level, there was a higher proportion of students below basic levels in the 9th and 10th grades. For the 9th and 10th graders, the proportion of students below basic level was much higher in mathematics and science. The proportion of students below basic level was 19.8% for 10th graders in 2004.

Thirdly, in terms of gender differences, girls had higher average scores than boys in Korean, social studies, science, and English, while the mean difference between girls and boys fluctuated in mathematics for 6th graders. For 9th graders, girls had higher average scores in general across all the subjects except for mathematics. 10th grade girls had higher average scores in Korean, social studies and English, and 10 grade boys had higher average scores in mathematics and science. Moreover, there was a higher proportion of girls at the advanced level across all 6 grade subjects. For both 9th and 10th graders, there were more girls at the advanced level in Korean and English, and more boys at the advanced level in social studies, mathematics and science. However, there was a higher proportion of boys below the basic level across all the subjects for all grades.

In addition to these general descriptions of student achievement, there was found to be a relationship between student achievement and background characteristics of the NAEA (Cho et al., 2007). Private schools tended to have higher average scores in grades 6 and 10, while there was no significant difference in Grade 9. Schools with large class sizes had higher proportions of students in advanced levels across all grades, while schools with small class sizes tended to have higher proportions of students below basic level. This can be explained by the fact that many schools with large class sizes are located in the large cities, while ones with small class sizes are mostly in rural areas. Additionally, students whose teachers have high self-efficacy as teachers, high

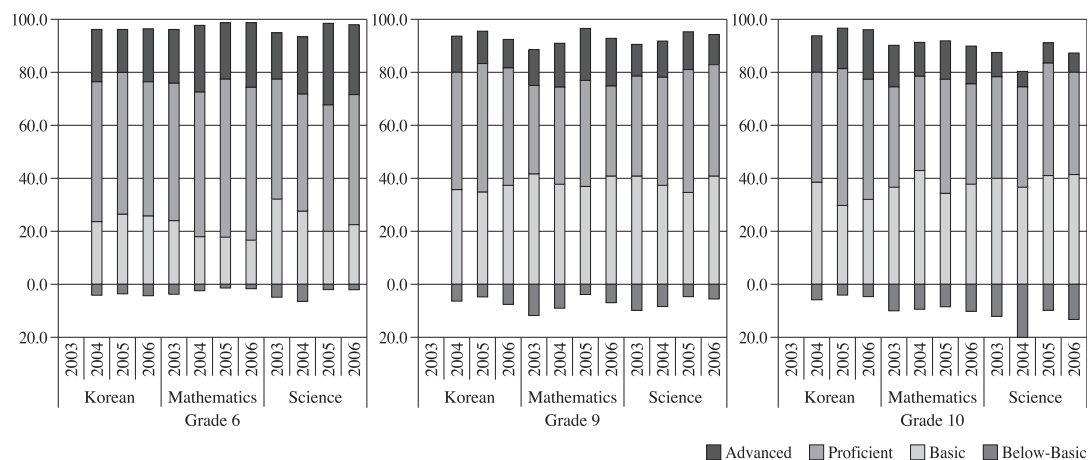


Figure 3 Changes in the student achievement levels by grades and subject matters from 2003 to 2006

Note: For the subject of Korean, the reference year of the trend analysis was 2004 instead of 2003. The proportion of students below basic level is presented below the x axis. The original figures from Cho et al. (2007) were re-organized by the author.

aspiration to teach and high expectations for student achievement tended to have higher achievement scores. Also, students who have parents with higher educational experiences and who spend more time talking with parents tended to have higher achievement scores. Finally, significant positive correlations with student achievement were found for self-regulated learning, school adaptation, student-teacher relations, positive self-concept and attitudes towards learning.

3 International Assessment of Student Assessment

In addition to these domestic assessments of educational achievement, Korea has also participated in two international comparison studies, the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). As a project of the International Association for the Evaluation of Educational Achievement (IEA), TIMSS provides information to improve teaching and learning in mathematics and science. TIMSS assesses achievement in mathematics and science at Grades 4 and 8 and collects a rich array of background information to address concerns about school resources and the quality of school curriculum and instruction. Conducted every four years on a regular cycle from 1995, TIMSS provides countries with an unprecedented opportunity to measure progress in educational achievement in mathematics and science.

Korea has been participating in TIMSS since the first cycle of 1995. TIMSS was designed for 4th and 8th graders, but only 8th graders in Korea have continuously participated in TIMSS from 1995 to 2007, and only 4th graders participated in 1995. The results of TIMSS can be summarized by the following (Martin et al, 2004; Mullis et al., 2004). The results of TIMSS from 1995 to 2003 showed that Korean students performed well in mathematics and science when compared to other participating countries. In 2003, Korea ranked 2nd with an average score of 589 in mathematics and ranked the 3rd with an average score of 558 in science among 46 participating countries. Furthermore, Korean students have showed an improvement with significant change over the 8-year period in both mathematics and science (average score of 581 to 587 and 589 score points for mathematics and 546 to 549 and 558 score points for sciences). However, the achievement of girls

was significantly lower than boys both in mathematics and science compared to other participating countries, even though these gender differences have continuously decreased since 1995. Despite the relatively strong performance of Korean students, they reported very low self-confidence in mathematics and science, as did Japanese students who took this test.

Korea has also participated in another large-scale international comparison study, the Programme for International Student Assessment (PISA). PISA was designed and developed by the OECD in the late 1990s as an ongoing periodic international comparative study in order to collect policy-oriented indicators on the educational systems. PISA assesses 15-year-olds in school. Unlike TIMSS, it is an age-based survey, rather than a grade-based one. The choice of this population means that the assessment is targeted to measure the extent to which students are prepared for the daily challenges of adulthood in modern society since compulsory education ends at this age in most countries. In this regard, PISA measures competencies, which is termed as “literacy,” rather than what is taught directly in schools. The definition of literacy is concerned with the capacity of students to extrapolate from what they have learned and to analyze and reason as they pose, solve and interpret problems in a variety of situations (OECD, 2007a). PISA surveys have taken place every three years since 2000. Although each cycle assesses all three assessment domains (reading, mathematics and science), the focus of the survey shifts from domain to domain in rotation, so that detailed analyses are periodically available for each domain, and in-depth comparisons are possible every nine years.

Korea has participated in PISA since the 1st cycle of 2000. The results of PISA have pro-

Table 1 Summary results of academic performance for Korean 15-year-olds

Domain	Country	Year	Average scale score	Percentage of students at level 5/6	Variance in student performance(SP)			Between-school variance in SP explained by	
					Total	Between schools	Within schools	SES of students	SES of students and schools
Reading	Korea	2006	556	21.7	80.2	33.0	48.7	2.1	14.6
		2003	534	12.2	75.2	27.6	48.9	4.6	16.7
		2000	525	5.7	54.5	20.8	34.6	2.3	10.4
	OECD average	2006	492	8.6	100.0	38.4	63.4	5.6	21.5
		2003	494	8.3	100.0	31.1	69.3	7.6	21.1
		2000	500	9.5	100.0	34.3	67.4	7.5	21.6
Math	Korea	2006	547	27.1	102.9	41.9	61.9	6.0	24.0
		2003	542	24.8	99.3	42.0	58.1	7.7	27.8
		2000	547	-	84.2	34.1	50.7	4.7	20.8
	OECD average	2006	498	13.3	100.0	36.8	64.6	7.3	21.9
		2003	500	14.6	100.0	33.0	67.4	8.3	22.6
		2000	500	-	100.0	32.4	68.6	8.2	20.5
Science	Korea	2006	522	10.3	90.2	31.8	59.3	3.8	16.9
		2003	538	-	101.4	38.9	63.2	5.8	23.2
		2000	552	-	74.2	29.4	45.9	2.9	15.7
	OECD average	2006	500	9.0	100.0	33.0	68.1	7.2	20.5
		2003	500	-	100.0	29.0	71.1	8.0	20.1
		2000	500	-	100.0	30.2	70.5	7.9	19.3

Note: The scale scores are standardized to have a mean of 500 and a standard deviation of 100 across OECD countries. They were then vertically scaled to allow for trend analysis. The information by level was reported from the year when the corresponding domain was the main focus of study. The highest level was 5 for reading and 6 for math and science. Total variance in student performance (SP) was expressed as a percentage of the average variance in SP across OECD countries. The author created this table based on data found in OECD (2001), OECD (2004) and OECD (2007b).

vided many insights on the strengths and weaknesses of educational systems in Korea and have facilitated much research (Yun & Lee, 2006; Lee, 2007; Shin et al., 2007; Park, 2008). The results of PISA have suggested that Korean students have performed relatively well compared to 15-year-olds in other participating countries. In PISA 2006, Korea ranked 1st in reading literacy, 3rd in mathematics literacy, and 10th in science literacy among 57 participating countries. As can be seen in Table 1, the performance in reading literacy was improved and mathematics performance remained stable, while science performance declined in PISA 2006 compared to PISA 2000 and PISA 2003 (OECD, 2007a). Also, the percentage of students at the highest levels, level 5/6, was much higher than the OECD average, except for science, which was near the OECD average.

Also, to examine the degree of educational inequality, we can look at the amount of total variance in student performance, the percentage of between-school variance among total variance, and the percentage of variance explained by SES of students and schools. The results of Table 1 suggest that there exist some amount of educational inequality in Korean education depending on the subject domain. Reading seems to be the domain which has least amount of educational inequality while mathematics seems to be the subject with most amount of educational inequality. Especially for mathematics, the percentage of between-school variance explained by SES of students and schools has been consistently above the OECD average, unlike other domains. This can be explained by *hakkun* effect in Korea which means that schools in wealthy areas tend to have higher academic performance.

One final, but very important result of PISA was the negative attitudes toward mathematics and science held by Korean students. As can be seen in Table 2, 15-year-olds in Korea had quite negative attitudes toward mathematics and science, even though they showed relatively high academic performances. Considering that these attributions can be regarded as foundations of lifelong learning, and that the relationship between these attributions and academic performance is quite positive, this phenomenon was taken quite serious and led to a variety of studies and policy changes to enhance student attitudes towards school learning.

4 New Challenges for the Assessment of Student Achievement in Korea

In Korea, the pursuit of higher educational attainment and entrance into prestigious schools has always been a priority of parents and students. To enter prestigious universities, Korean stu-

Table 2 Students' attitudes toward mathematics and science in relation to their academic performance

Year	Index	Mean index	Average performance of achievement scale		Percentage of explained variance in student performance	
			Bottom quarter	Top quarter	Korea	OECD average
2003	interest and enjoyment of math	-0.12	500	593	15.5	1.5
	self-concept in math	-0.35	493	604	21.4	10.8
	self-efficacy in math	-0.42	469	617	33.2	22.7
2006	general interest in science	-0.24	482	567	13.0	7.2
	enjoyment of science	-0.17	452	574	16.7	10.2
	self-concept in science	-0.71	487	569	12.6	8.8
	self-efficacy in science	-0.21	477	563	14.2	15.9

Note: The indexes were standardized to have a mean of 0 and standard deviation of 1 across OECD countries. This table was re-organized by author what were listed in OECD (2004) and OECD (2007b).

dents spend quite a lot of time for study at schools as well as at private tutoring industries from the early stage of school education. Considering these efforts of students and parents, the Korean government has made little attempt to monitor the educational progress of students and to diagnose and support students who are below the basic proficiency level. International studies such as TIMSS and PISA have introduced the need for the systematic assessment of student achievement, and are specifically designed for trend analysis, along with a variety of contextual questionnaires. Stimulated by the design and results of TIMSS and PISA, the Korean government introduced the National Assessment of Educational Achievement (NAEA) and the Diagnostic Test for Basic Skills (DTBS) in the early 2000s.

The results of these tests and those of the international studies have aroused people's attention to some concerns about Korean education, such as the relatively low performance in science, relatively large educational inequalities in mathematics, and students' negative attitudes towards school learning (Yun & Lee, 2006; Cho et al., 2007; Shin et al., 2007; Park, 2008). In spite of many existing studies, more studies are still needed to explain what makes schools perform well in Korean context, how students' attitudes toward school learning actually function to influence academic performances, and how to enhance students' attitudes toward school learning. Furthermore, more methodological research is needed to determine the cut-off scores for proficiency levels, vertical scaling method, and for identification of items with bias.

Recently, due to a newly introduced law, great changes are expected in the systems for the nationwide assessment of student achievement in Korea. In 2008, the Information Announcement Act on Educational Institutions was passed. According to this new law, all elementary and secondary schools are required to participate in the national assessment of student achievement such as the NAEA and to announce the results in public (Kim et al., 2007).

To satisfy this requirement, the Korean government changed the NAEA from a sample survey to a population survey. From 2008, all 6th, 9th and 10th graders in all elementary and secondary schools are required to participate in the NAEA. This is expected to change the nature of the test. Until now, the NAEA was not a high-stake test. The students and schools did not respond to this test sensitively because the results were only reported to the individual students and schools and were never used as a tool for school evaluation. Because of this nature of test, there was little worry about item exposure and it was easy to introduce common items for vertical scaling. However, it is now expected that schools will take initiatives to systematically prepare for this test in an attempt to make their schools look better. It will also make the items quite vulnerable for exposure and introduce more methodological issues for field testing and vertical scaling. There also has been a considerable amount of disagreement over how to develop a computerized scoring system for constructed response items, the need for reducing testing areas to fewer than five subjects, and changing the respondents' grades from the 6th, 9th and 10th to other grades (Kim, 2008; Cheong, 2008).

It is quite certain that changing from a sample survey to a population survey and becoming a high-risk test will introduce some new issues and conflicts in the NAEA. However, we always need to be cognizant of the reasons why we are doing these assessments. Essentially, it is important to continuously monitor the educational progress of the nation as well as schools, and we must find more in-depth answers for the conclusions made in international studies by bridging the results of these studies with domestic ones.

References

- Chae, Sunhee, Choi, Sukjin, Yang, Myonghee, Lee, Bongju, Lee, Jaegi, Kim, Donam, Kim Yunhee, & Kim, Jiyeon (2003). *The study on the national diagnostic test for basic skills of the grade 3 elementary school students in 2002*. Research paper published in Korean by the Korea Institute for Curriculum and Evaluation, RRE 2003-1. (in Korean)
- Cheong, Eunyoung (2008). *How to improve the systems of the national assessment of educational achievement (NAEA)*. Paper presented in Seminar on the ways to improve the national assessment of educational achievement hosted by the Korea Institute for Curriculum and Evaluation on Sep. 10, 2008. (in Korean)
- Cho, Jimin, Kim, Myonghwa, Choi, Inbong, Song, Miyoung, & Kim, Sujin (2007). *The study on the national assessment of educational achievement: Trend analysis from 2003 to 2006*. Research paper published in Korean by the Korea Institute for Curriculum and Evaluation, RRE 2007-3-1. (in Korean)
- KICE (2008). Internet homepage on the National Assessment of Educational Achievement. http://www.kice.re.kr/kice/eng/info/info_3.jsp (October, 3, 2008).
- Kim, Changhwan, Kang, Sungkuk, Kim, Byungju, Kim, Youngcheol, Kim, Heekyu, Park, Hyun-Jeong, Lee, Sunho, Yim, Hunam, Choi, Jungyun (2007). *Study on what to include in the MOE ordinance for the information announcement act on educational institutions*. Research paper published by the Korean Educational Development Institute. OR 2007-4. (in Korean)
- Kim, Shinyoung (2008). *Issues and Prospect of the systems of the national assessment of educational achievement (NAEA)*. Paper presented in Seminar on the ways to improve the national assessment of educational achievement hosted by the Korea Institute for Curriculum and Evaluation on Sep. 10, 2008. (in Korean)
- Lee, Kwanghyun (2007). Analysis on the effect of school choice on student achievement based on PISA 2003 database. *Journal of Educational Sociology* (published in Korea), Vol. 17, No. 4, 87–107. (in Korean)
- Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., & Chrostowski, S.J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., & Chrostowski, S.J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- OECD (2001). *Knowledge and skills for life: First results from PISA 2000*.
- OECD (2004). *Learning for tomorrow's world: First results from PISA 2003*.
- OECD (2007a). *PISA 2006 science competencies for tomorrow's world. Volumn 1: Analysis*.
- OECD (2007b). *PISA 2006 science competencies for tomorrow's world. Volumn 2: Data*.
- Park, Hyun-Jeong (2008). Competitiveness of secondary school education in Korea and the future directions, *The HRD Review*, 34–41. Periodicals published in Korean by Korea Research Institute for Vocational Education and Training, Korea. (in Korean)
- Shin, Jongho, Lee, Hyunjoo, Kim, Yongnam, & Yim Sunyoung (2007). *Effects of individual student and school on mathematics achievement: Comparisons between Korea and Japan*. Paper presented in 2007 International Conference: Competency-based Education Reform in Seoul on March 21, 2007. (in Korean)
- Yun, Chung-il & Lee, Bum-soo (2006). International comparative study on academic achievement determinants of PISA 2003: Focusing on Korea, Japan, the United States, Germany and Finland. *Journal of Educational Administration* (published in Korea), Vol. 24, No. 2, 523–537. (in Korean)