East Asian Mathematical Journal
Vol. 25 (2009), No. 3, pp. 279-297

Comparison of NAEP, PISA, and TIMSS-R

Young-Ok Kim · Henry Wakhungu · In-Suk Ku

Abstract. This paper discusses the key differences and common themes among three major assessments of pre-college mathematics learning: Program for International Student Assessment (PISA), Third International Mathematics and Science Study Repeat (TIMSS-R), and National Assessment of Educational Progress (NAEP).

I. Introduction

In this paper, we discuss the key differences and common themes among three major assessments of pre-college mathematics learning: Program for International Student Assessment (PISA), Third International Mathematics and Science Study Repeat (TIMSS-R), and National Assessment of Educational Progress (NAEP). At the beginning of this paper, we elucidate the backgrounds of the three assessments, their respective goals, and assessment cycles. Next, we focus on the mathematics frameworks of the three assessment programs including mathematics knowledge, mathematics abilities, and content areas. Then, we discuss the different methodologies used in the three assessments, and their common themes. Finally, we discuss the extent to which each the NAEP program is of value to teachers, policy makers, and researchers in the U.S. vis-à-vis the PISA and TIMSS assessments that have stakeholders nationally and internationally across the participating countries. Thus whereas beneficiaries of NAEP are restricted to the U.S., PISA and TIMSS will be of interest to individuals worldwide.

2009년 8월 투고, 2009년 9월 심사 완료.
This work was supported by Kyungnam University Foundation Grant, 2009.
2000 Mathematics Subject Classification: 97D60
Key words: TIMSS, PISA, NAEP, Assessments

II. Background of the NAEP, TIMSS-R and PISA Assessments

In this section, the backgrounds, goals and assessment cycles NAEP, PISA and TIMSS-R are discussed. First, NAEP which is also known as "the nation's report card" restricted to the U.S., whereas PISA and TIMSS are international. NAEP is an on going congressionally mandated survey designed to measure *what* students in the U.S. *know* and *can do*. The NAEP program is sponsored by the U.S. Department of Education and administered by the National Center of Education Statistics (NCES) whose policy has been determined by the non partisan independent National Assessment Governing Board (NAGB). The congressional legislation that established the state NAEP program also mandated standard-based reporting of NAEP results; stating that NAEP results should be presented both as overall scores and in terms of percentages of students in grades who meet established standards for performance. NAEP consists of two distinct assessment programs referred to as *trend NAEP* and *main NAEP* with different instrumentation, sampling, administration, and reporting practices. In this paper, we will focus on the main NAEP which has two components, *national* NAEP and *state* NAEP. National NAEP typically tests nationally representative samples of students in grades 4, 8, and 12. The object is to measure achievement in NAEP subject areas in relation to current thinking about curriculum and instruction. In most but not all subjects, NAEP is administered two, three, or four times during a 12-year period, which makes it possible to examine changes over a decade. State NAEP assessments are administered to state-representative samples of grade 4 and 8 students in states that elect to participate in the state assessment programs via the same large-scale assessment materials that are used in *national* NAEP. In the last decade, most NAEP assessments have reported summary scores and percentages of students performing at or above basic, proficient, and advanced levels of performance by subgroups (i.e., gender, race/ethnicity, region, school (public/nonpublic, etc). NAEP conducts directly comparable state assessments using collected mathematics and science data from nationally represented samples of students in public and private schools.

NAEP results have been increasingly used by policy makers, researchers, teachers and the public as indicators of the nation's educational health. Subsequently, there has been a departure from the original design of NAEP through the 1990s as a result of the pressures on NAEP to do more than and more beyond its established purposes. Various educators and policy makers have suggested that NAEP be used as a lever for education reform, as an anchor for other assessments, as an accountability tool, and asan international assessment tool. It is against this backdrop of change in design and pressure on NAEP that warranted the comparison with other assessments. Although the NAEP assessment cycle is yearly, the last mathematics, science, and reading assessments were in 2000 and the next assessments in these domains will take place in 2004.

Second, the original TIMSS which was administered in 1995 was the largest, most comprehensive, and most rigorous international study of the schools and student achievement that involved the testing of more than a half-million students in mathematics and science at different grade levels in 41 countries. In addition to testing students' achievement in mathematics and science, TIMSS-R collected a range of information about the contexts for learning these subjects. In particular, TIMSS-R examined the curricular goals of the education system and how the system is organized to attain those goals; the educational resources and facilities provided; the teaching force and how it is educated, equipped, and supported; classroom activities and characteristics; home support and involvement; and the knowledge, attitudes, and predisposition that students and teachers themselves bring to the educational enterprise. The contextual framework adapted by TIMSS-R identifies the major characteristics of the educational and social contexts that will be studied with a view to improving student learning opportunities.

The TIMSS-R curriculum model has three parts: the *intended* curriculum, the *achieved* curriculum, and the *achieved* curriculum. Working from this model, TIMSS-R uses mathematics achievement tests to describe student learning in the participating countries, together with questionnaires which ask about the structure and content of the intended curriculum in mathematics, the preparation, experience, and attitudes of teachers, the mathematics content actually taught, the instructional approaches used, the organization and resources of schools and classrooms, and the experiences and attitudes of students in the schools. The original TIMSS-R had three student populations and three assessments: Population I, students in the two grades enrolling the largest number of 9-year-old students (third and fourth grade in most countries); Population II, students in the two grades enrolling the largest number of 13-year olds (seventh and eighth grade in most countries); and Population III, students in the final year of secondary education. TIMSS-R, administered in 1999 to students in 38 countries, was essentially a repeat of the population II assessment.

In brief, TIMSS-R is based on the same framework as TIMSS, and approximately one third of the assessment items are the identical to those administered to the Population II. Like NAEP, TIMSS-R seeks to assess students' mastery of basic knowledge, concepts, and subject-specific thinking skills tied to extensive frameworks of curriculum topics. As a result, both NAEP and TIMSS-R consist of large numbers of items covering a broad range of topics, with items generally focused on a single, identifiable piece of knowledge, concept, or skill. Some items draw on a combination of topic areas or are more focused on students' mathematical thinking abilities than on content topic.

Finally, the PISA (Program for International Student Assessment) assessments first administered in the year 2000 aim to assess the extent to which 15-year old students in 32 countries have acquired the broader knowledge and skills in the mathematical, scientific and reading literacy domains corresponding to school subjects that they will need in adult life. Thus PISA features separate assessments in these domains of literacy with one of the three domains being designated the "major" domain, with approximately two thirds of assessment time devoted to it in each three year administration cycle of PISA.

Nohara (2001) states that the goal of the PISA program is to measure the "cumulative yield" of education systems, that is, students' knowledge and abilities near the end of their primary-secondary educational careers through its focus on students' ability to function in situations common in adult life in a mathematically literate society. In the second cycle of the year 2003, mathematical literacy will be the major domain and the third cycle year 2006; the major domain will be science. In cases where a domain is not the major domain, given that less time is available for it, the assessments do not attempt to encompass the full range of all aspects identified in the assessments frameworks. Therefore as we compare the three assessments for the year 2000, we note that mathematical literacy and scientific literacy were minor domains and thus fewer items were developed for PISA in these areas than for either NAEP or TIMSS-R. PISA also differs from NAEP and TIMSS-R in that most items are clustered, in groups of two to four, around a common situation described partly by text, graph, or chart, with the sequence of questions increasing in complexity or difficulty.

In short, there are clear differences in the purposes and philosophical underpinnings of each assessment. Most salient is the fact that while both NAEP and TIMSS-R seek to find out how well students have mastered curriculum-based scientific and mathematical knowledge and skills, the purpose of PISA is to assess students' scientific and mathematical "literacy", that is their ability to apply scientific and mathematical concepts and thinking skills to everyday, non-school situations. In addition, the assessment cycles for the three assessments differ significantly.

III. NAEP-TIMSS-PISA Mathematics Frameworks

Content areas

This section describes each of the mathematics content strands in each assessment. First, NAEP has gathered data on students' understanding and performance and captures a range of thinking skills needed by students in order to deal with the complex issues they encounter inside and outside their classrooms. NAEP assessment focuses essentially on the following five broad strands:

a) Number sense, properties, and operations; they include whole numbers, fractions, decimals, integers, operations and their application to mathematical and real-world situation.

b) Measurement; it includes concepts of time, money, temperature, length, area, capacity and angles measure.

c) Geometry and spatial sense; it includes properties of angles and polygons and the application of reasoning skills to geometric situations.

d) Data analysis, statistics, and probability; it includes problem solving involving data and graphical representations of data, informal measure of central tendency.

e) Algebra and functions; it includes algebra notation, the meaning of variable, and the solution of the simple equations and inequalities (Silver &Kenney, 2000).

These divisions are not intended to separate mathematical into discrete elements. Rather, they are intended to provide a helpful classification scheme that describes a wide range of mathematical content assessed by NAEP. Classifying items into one primary content area is not always clear-cut, but doing so helps to meet the goal of ensuring that all important mathematical concepts and skills are assessed (Silver & Kenney, 2000).

Second, TIMSS-R tends to measure students' mastery of knowledge, skills, and concepts. As a result the description of TIMSS mathematics content area is highly detailed and serves as the primary consideration in item development. The following content area categories are general areas covered at eighth grade in TIMSS:

a) Fractions and number sense include: whole numbers, fractions and decimals, integers, exponents, estimation and approximation, and proportionality.

b) Measurement includes: standard and non-standard units, common measures, perimeter, area, volume, and estimation of measures.

c) Data representation, analysis, and probability include: representing and interpreting tables, charts, and graphs; range, and mean; informal likelihood, and simple numerical probability.

d) Geometry includes: points, lines, planes, angles, visualization, triangles, polygons, circles, transformations, symmetry, congruence, similarity, and constructions.

e) Algebra includes: number patterns, representation of numerical situations, solving simple linear equations, operations with expressions, representations of relations and functions (Nohara, 2001; TIMSS, 1999).

Finally, the content area of PISA mathematics literacy includes two major aspects: mathematical competency and mathematical big idea, as well as two minor aspects: mathematical curriculum strands and situations. These aspects are used to organize the mathematics domain.

Major aspects

(1) Mathematical competencies are the processes of working on mathematics task. These include the use of mathematical language, modeling, and problem-solving skills. The questions are organized in terms of three competency classes and mathematisation.

Competency classes

- Class 1: reproduction, definitions, and computations.
- Class 2: connections and integration for problem solving.
- Class 3: mathematical thinking, generalization and insight (PISA, 2000).

Mathematisation refers to the organization of perceived reality through the use of mathematical ideas and concepts. The process of mathematisation occurs in two different phases: horizontal mathematisation, which is the process of translating the real world into the mathematical world, and vertical mathematisation; that is, working on a problem within the mathematical world and using mathematical tools in order to solve the problem (PISA, 2000; the PISA, 2000).

(2) Mathematical big ideas refer to content of mathematics. They represent clusters of relevant, connected mathematical concepts that appear in real situation and contexts. Mathematical big ideas contain the following: (a) chance, (b) change and growth, (c) space and shape, (d) qualitative reasoning, (e) uncertainty, and (f) dependency and relationships (PISA, 2000).

Minor aspects

(1) Mathematical curricular strands

Mathematical curricular strands constitute a minor domain. This represents the content of school mathematics as implemented in many school curricula. It contains: (a) number, (b) measurement, (c) estimation, (d) algebra, (e) functions, (f) geometry, (g) probability, (h) statistics, and (j) discrete mathematics.

(2) Situations

An important part of the mathematical literacy is that students use mathematics in a variety of situations. Situations contain: (a) personal, (b) educational, (c) occupational, (e) public, and (f) scientific (PISA, 2000)

Content area and the three assessments

NAEP and TIMSS have to some extent similarities in focusing on their content area of mathematics. They both focus on specific areas and present their items to measure students' understanding of these topics. For example, number sense, in both NAEP and TIMSS mainly focus on students' understanding the operations, integers, and rational numbers, as well as fractions, decimals, and estimations. In contrast, PISA focuses mainly on the process and content of a broad mathematical concept such as mathematical thinking, generalization, change and growth, and qualitative reasoning. On the other hand, the content of school mathematics in PISA assessment is less important although it is somewhat similar to the content area of NAEP and TIMSS. Furthermore, PISA includes situations domain which concentrate on a variety of mathematics situations such as educational, public, and scientific.

Item complexity

This section focuses on the mathematical knowledge and ability that students should use and have in each assessment. That is, knowledge and ability imply the level of complexity and amount of thinking necessary to solve problems. Although the three assessments used fairly different terminologies, they all pay attention to the cognitive demand of various type of items. In NAEP, the mathematical ability, power, and achievement are the main concerns. TIMSS-R focuses on performance expectations, and PISA focuses on general mathematical skills.

First, NAEP includes: (a) Mathematical abilities to describe the level of mathematical complexity that an item demands of students. Each level describes the degree of procedural knowledge, conceptual understanding, and problem solving. The notion of conceptual understanding, procedural knowledge, and problem solving send a strong message about the depth and breadth of engaging in a mathematical activity. (b) The mathematical power gives further emphasis to the idea that certain activities cut across content areas. The abilities and power dimensions are not intended for reporting, but rather to provide for a wide range of mathematical activity in the item which includes: reasoning, connections, and communication (Silver & Kenney, 2000). (c) The achievement levels are: (1) Basic performance which denotes partial mastery of the knowledge, content, and skills for desired work at each grade level, but not a level of work that is deemed satisfactory. (2) Proficient performance which represents solid academic performance at each grade levels, and (3) Advanced performance that signifies a superior level of work at each of the grades (Silver & Kenney, 2000).

Next, TIMSS-R specifies the following levels of items complexity:

(a) Knowing: knowing facts encompass factual knowledge that provides the basic language of mathematics, and the essential mathematical facts and properties that form the foundation for mathematical thought.

(b) Using routine procedures: procedures for a bridge between more basic knowledge and the use of mathematics for solving routine problems, especially those encountered by many people in their daily lives.

(c) Investigating and problem solving: problem solving is a desired outcome of mathematics instruction linked with many mathematics topics.

- Mathematical reasoning: it involves the capacity for logical, systematic thinking. It includes intuitive and inductive reasoning based on patterns and regularities that can be used to arrive at solutions to non-routine problems.
- Communicating (Nohara, 2001; TIMSS, 1999).

Finally, PISA focuses on general mathematical skills which fall under the major aspect of assessing students' knowledge and ability as follows:

(a) Mathematical thinking and skills: This includes posing question characteristics of mathematics, distinguishing between different kinds of statements, and understanding and handling the extent and limits of given mathematical concepts.
(b) Mathematical argumentation skill: This includes knowing what mathematical proofs are and how they differ from other kinds of mathematical reasoning.
(c) Modeling skill: This includes structuring the field or situation to be modeled.
(d) Problem posing and solving skill: This includes posing, formulating, and defining different kinds of mathematical problem.
(e) Representation skill: This includes decoding, interpreting and distinguishing between different forms of representation.
(f) Symbolic, formal and technical skills: These include decoding and interpreting symbols, and understanding their relationship to natural language.
(g) Communication skill: This includes expressing oneself in variety ways on matters with a mathematical content.
(h) Aids and tool skill: These include knowing about and being able to make use of various aids and tools. (PISA, 2000)

Item complexity: common themes

NAEP contains mathematical ability, power, and achievement which focus on the level of complexity of the question being asked, and give an idea of the process and the knowledge that are required from students to deal with items. Unlike NAEP, TIMSS includes knowing facts, bridge between basic knowledge and the use of mathematics for solving routine problem and how students would achieve these expectations. On the contrary, PISA deals with general mathematics skills that aim at more broad goals than other assessments. For example, mathematical argumentation skills include knowing what mathematical proofs are and how they differ from other kind of mathematical reasoning.

Question type

NAEP and TIMSS-R had similar number of items with 160 and 164, respectively. Since mathematics was a minor domain in PISA, only 32 items were given to students. Questions from the three assessments can be categorized into three formats: (1) Multiple choices, (2) Short answers; and (3) Extended answers.

Each assessment, however, labels its questions differently. For example, PISA uses the labels: multiple-choice, closed-constructed response, and open-constructed response. The PISA development team believes that multiple-choice format has limitation in measuring the high-order goals so they prefer to devote more items to the open questions (PISA, 2000). TIMSS-R and NAEP devoted 77 and 60 percent of all items to multiple-choice format, respectively. However, the percentage of the multiple-choice items in PISA was just 34 percent. Unlike NAEP and TIMSS-R where the multiple-choice dominated, the short answer items dominated the question type in PISA. The percentage of the extended answer items did not exceed 16 percent in any one of the three assessments (Nohara, 2001). The percentages of the extended answers were the less on the three assessments because such answers take students longer to write than the multiple-choice or short answer items. However, the extended answers items on the three assessments account for a significant portion of the overall assessments.

A Portion of Items situated in real-world contexts

How much items situated in real-world contexts are included in the assessment is an important comparison because connecting mathematics to the world outside of school is a major goal of many mathematics education reform initiatives. It is also significant because it means that students have to choose for themselves the operations and solutions most appropriate for the problem and figure out how they relate to the information provided, thereby adding to the difficulty of an item. Gravemeijer and Doorman (1999) emphasized a reflexive relation between the use of context problems and the development of the student's experiential reality. On the one hand, the context problems are rooted in this reality on the other hand, solving these context problems helps the students to expand their reality. Notwithstanding this dynamic character of reality that defines context problems, starting points for instructional sequences will often link up with everyday life experience of students. Three assessments contained many items situated in real-world contexts, 48 percent of items on NAEP, 44 percent of items on TIMSS-R, and all but one item on PISA, 97 percent.

Figure 1. Percent and number of items that present students with real-life situations or scenarios as settings for the problem

	Percent	Number of items
NAEP	48	79
TIMSS-R	44	72
PISA	97	31

Resource: A comparison of the national assessment of educational progress(NAEP), the Third International Mathematics and Science Study Repeat(TIMSS-R), and the Programme for International Student Assessment(PISA) by David Nohara (2001).

PISA has several items that represent real-life situations, defined as items not presented strictly in the language of mathematics. These contexts are highly unique, that is, not typically encountered in mathematics instruction or textbooks, and requires significantly more thought regarding how the nature of the context affects the mathematics involved in the problem.

NAEP and TIMSS-R have context items on subsets of items, but the number of items is only a few. However, it is not surprising because the PISA's goal and focuses are very different to NAEP and TIMSS-R. The stated goal of the PISA is to measure the "cumulative yield" of education systems, that is, to measure students' knowledge and abilities near the end of their primary-secondary educational careers. It focuses on students' ability to function in situations common in adult life in a mathematically literate society, as opposed to their mastery of detailed sets of curriculum topics. Otherwise, NAEP and TIMSS-R serve as the primary source of information on students' knowledge and skills in the mathematics. They address knowledge and skills commonly found in school curricula and national curriculum documents, including both specific content topics and broader thinking skills.

Aids and Tools

NAEP permitted students to use tools such as rulers, protractors, calculators, and manipulatives such as geometric shapes. However, the use of these tools was restricted to particular blocks of items; students did not have access to these tools at all times during the assessment. Approximately two-thirds of the assessment measures students' mathematical knowledge and skills without access to a calculator; the other third of the assessment allows the use of a calculator. NAEP assessment contains blocks for which calculators are not allowed, and calculator block, which contain some items that would be difficult to solve without a calculator (Silver & Kenney, 2000). PISA, on the other hand, allowed students to use their own calculators and tools whenever they wanted. PISA adopted this policy because it represents the most reliable assessment of what students can achieve. The PISA development team claimed that the test items were chosen so that the use of calculators is not likely to enhance a student’s performance in the assessment because the test items were chosen to meet this criterion (PISA, 2000). PISA questions, nevertheless, were more likely to need calculators to answer comparing to the NAEP blocks where the use of calculators was not allowed. TIMSS-R was different than the two other assessments in that it did not allow the use of calculators.

Questionnaires

This section presents differences and similarities among the NAEP, TMISS-R, and PISA questionnaires. More specifically, we investigate a series of questionnaires and specific questions in teacher questionnaire, student questionnaire and school questionnaire for the three assessments.

Interpretations of any observed students’ achievement similarities or differences must be made with caution because of the complex influences that affect mathematics learning, such as school factors that include tracking policies that affect or instructional practices and availability of appropriate resources; students’ attitude and beliefs, their self-perceptions and expectations regarding their mathematics ability, and their beliefs about mathematics; and family influences that include parental involvement and expectations, socioeconomic status, and cultural customs (Silver & Kenney, 2000).

Although the main assessments of NAEP, TIMSS-R and PISA assessments provide important data on student achievement, the three assessments are adapting the questionnaires to obtain additional information to provide a context for the achievement data and to help the interpretation of result in achievement. NAEP developed four kinds of questionnaires, such as student questionnaires, teacher questionnaires, school questionnaires, and SD/LEP (Students with Disabilities or Limited English Proficiency) questionnaire. TMISS-R has three kinds of questionnaires for student, teacher, and school. PISA gathered information through student and school questionnaires.

TIMSS-R developed separately mathematics teacher questionnaire and science teacher questionnaire, but one student questionnaire consists of both questions of mathematics and science. In PISA, one student questionnaire is including three domain questions: reading literacy, mathematical literacy, and scientific literacy.

In addition to reporting a series of questionnaires of each assessment, we compared specific questions of teacher questionnaire and student questionnaire for three assessments. Firstly, teacher questionnaires of NAEP and TIMSS-R provide various and detailed information about teachers so that they include background of teachers, belief and attitude about teaching and learning, instructional practice, and so forth. Particularly, TIMSS-R teacher questionnaire is very detailed because one of purposes of TIMSS-R is to examine the curricular goals of the educational system and how the system is organized to attain those goals.

We also compared student questionnaires from the three assessments in terms of perceptions of and attitudes toward their school experience. NAEP student questionnaire includes specific questions to survey students' responses about course enrollment history, but TIMSS-R and PISA provide few such questions. The three assessment questionnaires include questions related to time spent studying mathematics and attitudes toward and perceptions about mathematics. However, the questions for PISA are not useful because the questions themselves are superficial.

Therefore, when we compare the three assessment questionnaires as math educators, the NAEP and TIMSS-R questionnaires provide more abundant background information than PISA for the interpretation of the student achievement. However, we cannot say exactly which assessment questionnaire is better than the others because the goal and focus of each assessment are different.

IV. Common Themes of Three Assessments

Assessment frameworks

All three assessments are based on multi-dimensional frameworks that outline the important facts, concepts, and competencies to be covered on the assessments and other desirable characteristics for items. In all three frameworks, there is one dimension consisting of content topics and sub-topics (e.g., "algebra") and at least one describing non-topic-based cognitive processes (e.g., "reasoning"). Although these various sub-dimensions in one dimension may make each framework as a whole appear somewhat complex, they reflect the idea that the importance of any subject comes not just from its body of facts and concepts, but also from processes and skills related to it, not associated with any one topic or sub-topic.

	Questionnaire	Completed by	Questions focus on
NAEP	Student Questionnaire	students attending in NAEP assessment	background characteristics, subject area experience, and motivation on the assessment
	Teacher Questionnaire	mathematics teacher in which students are being assessed	background and training and classroom-by-classroom information
	School Questionnaire	principal or another official of each school in NAEP	information about school polices
	SD/LEP Questionnaires	teachers of those students who were selected to participate in NAEP and who were classified as students with disability and students with limited English proficiency, or individual education plans	information about students within the sample who have disabilities or limited English proficiency.
TIMSS-R	Student Questionnaire	each student who takes the TIMSS-R mathematics and science	daily activities, family attributes, educational resources in the home, attitudes and beliefs about learning, instructional processes in the classroom, and study habits and homework
	Teacher Questionnaire	mathematics teachers of each student in TIMSS-R	attitudes and beliefs about teaching and learning, teaching assignments, class size and organization, topics covered, the use of various teaching tools, instructional practices, and participation in professional development
	School Questionnaire	principal of each school in TIMSS-R	community attributes, personnel, teaching assignments, policy and budget responsibilities, curriculum, enrollment, behavioral problems, instructional organization, and mathematics and science courses offered
PISA	Student Questionnaire	Students attending in PISA	family, experience of students' school, and what students plan to do in the future
	School Questionnaire	principal or designate	characteristics of the student body, relationship the school has with the students, some of the administrative structures within the school and some of the pedagogical practices of the school

Table 1. Series and factors of questionnaires for NAEP, TIMSS-R, and PISA assessment questionnaire

Another common feature of framework categories and topics is that they are not mutually exclusive: all three frameworks recognize that a single item may address more than one content topic or involve more than one type of cognitive skill.

Content

When assessment items were placed in the NAEP mathematics Content Strands, there were clear differences in the content emphases of the three assessments (see Figure 2.). While approximately one fifth of the items on all three assessments dealt with Algebra and Functions, the degrees of emphases on the other four categories differed considerably (Nohara, 2001).

Figure 2: Percent and number of items that address NAEP mathematics Content Strands

	NAEP (165 items)		TIMSS-R (164 items)		PISA (32 items)	
	Percent	Number of items	Percent	Number of items	Percent	Number of items
Number sense, properties, and operations	32	52	46	76	9	3
Measurement	15	24	15	24	25	8
Geometry and spatial sense	20	33	12	20	22	7
Data analysis, statistics, and probability	14	23	11	18	31	10
Algebra and functions	20	33	19	31	19	6

Resource : A comparison of the national assessment of educational progress(NAEP), the Third International Mathematics and Science Study Repeat(TIMSS-R), and the Programme for International Student Assessment(PISA) by David Nohara (2001).

Response type

Over 75 percent of items on all three assessments were either multiple-choice or short answer. On TIMSS-R, these types of items accounted for all but four percent of items, with 77 percent of all items being multiple-choice and 20 percent being short-answer. On NAEP, 60 percent of items were multiple-choice and 16 percent were short answer. PISA differed from the other two assessments in that there were more short answer items, 50 percent of all items, than multiple-choice, 34 percent (Nohara, 2001).

Beneficiaries of three assessments

NAEP is of greatest value to policy makers given their dominant role in its initiation and design as highlighted in the background section. It has been suggested that policy makers can use NAEP as a lever for education reform and as an accountability tool (Pellegrion, Jones, & Mitchell, 1999). However, beside policy makers and comparing to the other two assessments, NAEP could be of great value to teachers if more items are released. NAEP has this potential because it is extensively tied to curricula that used in most American schools. Moreover, NAEP compares the performance of students in individual states with each other and with the U.S as a whole (Johnson, Siegendorf, & Philips, 1998). Such comparison is more reasonable and makes more sense to American teachers than the comparisons that are made by TIMSS-R or PISA. Nonetheless, with NAEP being more restricted in releasing items- it is expected that no items will be released from NAEP 2000- teachers will be no longer able to examine the items that students could not answer in order to improve their teaching methods. Researchers, likewise, have encountered difficulty in doing secondary analysis of NAEP results. Although NAEP data have the potential of being a valuable resource for researchers, the restricted policy about items released prevents many researchers from conducting secondary analysis of NAEP data.

According to OECD, the main reason for developing PISA is to provide empirically grounded information that will inform policy decisions. However, other individuals such as researchers, teachers, and laypersons expected to benefit from the PISA result (PISA, 2000).

PISA is of great value to policy makers because it provides policy-oriented indicators. It produces international comparative analyses of student achievements in the participant countries. It provides direction for schools instructional efforts and insights into curriculum strengths and weaknesses. It allows policy makers to compare the performance of their education systems with those of other countries. Therefore, PISA provides tools to allow policy makers to monitor the education system in their countries. Even in the U.S where is the education system is not central; the government is gradually getting involved in steering the education system.

Since PISA produces contextual indicators relating results to students and school characteristics, researchers can benefit from it results especially when designing and evaluating curricula, pedagogies, and school practices. The high performances of students in the top countries correlate with certain type of curricula or pedagogies, so PISA results could be used to investigate those factors.

Although PISA is not of great value to teachers as is it to policy makers and researchers, its results, nevertheless, could be self-promoting for teachers who want to acquire professional development. Such teachers usually take the initiative to implement different promising techniques in their classroom without waiting for the high authorities to make such decisions. For example, since Japan was on the top in the mathematical literacy assessment in PISA and TIMSS. Teachers can use the TIMSS videotape classroom study to see how Japanese teachers teach mathematics concepts differently.

Like PISA for teachers, TIMSS may help teachers to see another kind of pedagogy comparing different teaching styles. Since the result of TIMSS is international to every participating country, many researches may be conducted in many different sites and different population which might lead to help researchers to answer their questions research. Conducting many researches is richened the nature of research and strengthened the outcomes of the studies. TIMSS has a great deal of released items that may enable more meaningful reports. On the other hand, policy maker may take the students' results compared to other countries as a reason to change the curricula and the system in order to help improve their students' achievement. They have the ultimate decision to support developing their country education whether it is very high and to preserve it or low and try to elevate it.

V. Conclusion

From the discussion in the background section, clear differences among the three assessments in their purposes and philosophical underpinnings emerged coupled with different assessments cycles. Based on the mathematical frameworks we found that NAEP and TIMSS are somewhat similar in the form of their respective curricula content areas. They both focus on specific content areas and present their question items to measure students' understanding of the relevant domains. PISA, on the other hand, concentrates on assessing the extent to which students have acquired the wider knowledge and general skills needed in every day life. Although the three assessments label their questions differently, the questions can be categorized into three formats: multiple-choice, short answer, and extended answer.

Both NAEP and PISA allow the use of aids and tools. However, TIMSS does not permit the use of calculators. The scope of the questionnaires of the three assessments differ remarkably, albeit the common themes with respect to assessment frameworks. The beneficiaries of the three assessments oscillate on a broad spectrum of policy-makers, researchers, teachers, educators, and the general public. Subsequently, the degree of value to be derived from each assessment is a function of specific questions or issues that may need to be addressed or answered by the respective beneficiaries. Evidently, policy-makers in political arenas are likely to exploit any of these assessment programs to address their political agendas as exemplified by the envisaged use of NAEP in the No Child Left Behind (NCLB) initiatives.

References

[1] Braswell, J. S., Lutkus, A. D., Grigg, W. S., Santapau, S. L., Tay-Lim, B., & Johnson, M (2001). *The Nation's Report Card: Mathematics 2000*. Washington, DC: U.S. Department of Education. Office of Educational Research and Improvement. National Center for Education Statistics.

[2] Gravemeijer, K., & Doorman, M. (1999). Context problems in realistic mathematics education: a calculus course as an example, *Educational Studies in Mathematics*. 39, pp. 111-129.

[3] Johnson, E. G., Siegendorf, A., & Philips, G. W. (1998). *Linking the National Assessment of Educational Progress and the Third International Mathematics and Science Study: Eighth-Grade Result.* Washington, DC: U.S Department of Education. Mathematics Items, ISC, Boston College. http://isc.bc.edu/TIMSS-R1999i/study.html

[4] Nohara, D.(2001). *A comparison of the National Assessment of Educational Progress(NAEP), the Third International Mathematics and Science Study-Repeat (TIMSS-R), and the Program for International Student Assessment (PISA).* U.S. Department of Education, National Center for Education Statistics, Washington, DC.

[5] PISA 2000 measuring student knowledge and skills: A new framework for assessment.
http://www.pisa.oecd.org/Docs/Download/PISAFrameworkEng.pdf

[6] PISA 2000 assessment of reading, mathematical, and scientific literacy.
http://www.pisa.oecd.org/Docs/Download/PISAsampleitemsEng.pdf

[7] Silver & Kenney, (2000). *Results from the seventh mathematics assessment of the National Assessment of Educational Progress*. National Council of Teachers of Mathematics (NCTM).
[8] Strutchens, M. E., & Silver, E. A. (2000). *NAEP findings regarding race/ethnicity: students' performance, school experiences, and attitudes and beliefs*. In E. A. Silver and P. A. Kenney (Ed.), Results from the Seventh NAEP Mathematics Assessment, pp. 45-72: Reston, VA National Council of Teachers of Mathematics.

Kim, Young Ok
Department of Mathematics Education
Kyungnam University
Masan, 631-701, Korea
E-mail address: youokim@kyungnam.ac.kr

Henry Kerre Wakhungu
School of Public and Environmental Affairs
Indiana University Bloomington
IN 47408, USA
E-mail address: hewakhun@indiana.edu

In-Suk Ku
Graduate School of Kyungnam University
Masan, 631-701, Korea
E-mail address: insuk1230@hanmail.net