

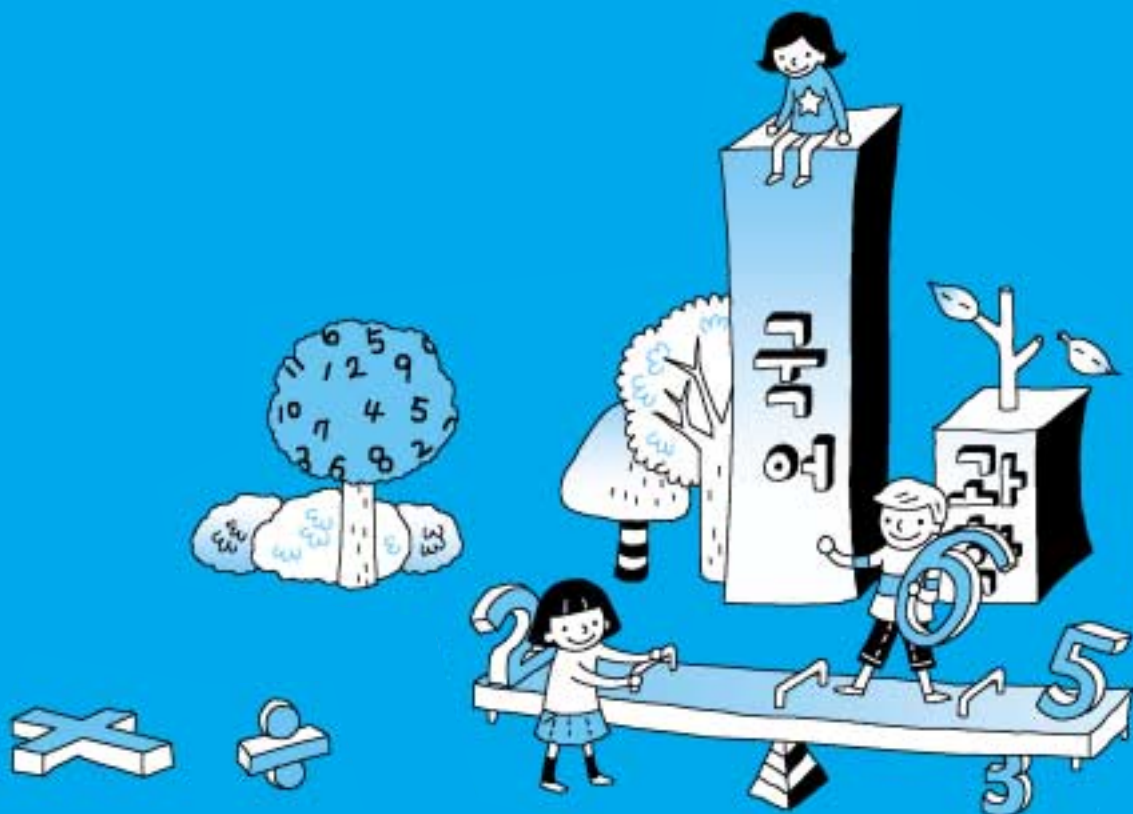


A Further Study on improving the National ASSESSment of Educational Achievement

- Scaling, equating and utilizing the results -

Research director_ **KIM Sung-Sook**

Collaborator_ SANG Kyoung-Ah, LEE Sang-Ha



I . Introduction

Need for and Goals of the Research

- A need was raised to design a new evaluation framework as the National Assessment of Educational Achievement (NAEA) expands to include all students in order to provide information helpful to increase the academic abilities of individual students.
- There should be a plan to implement equating and optimize data gathering and designing in the overall evaluation by analyzing the factors that damage the stability of test equating in the current academic achievement evaluation system.
- There is a need to propose ideas about the development of scale points and the scoring process by reviewing the scaling process in the current academic achievement evaluation system.
- There should be a plan to utilize the results of academic achievement on each level including individuals, schools, educational agencies, government and educational research institutions as well as enhance the methods of individual score reporting.
- It aims to suggest ideas and models for utilizing the results as well as the ways to improve the current scoring system including scaling and equating in the design of a new framework for academic achievement evaluation according to the full-scale implementation of the NAEA.

Content

■ Plans for Test Equating

- Analyzing the problems with test equating and data gathering and designing in the current academic achievement evaluation system according to its full-scale implementation.
- Examining the methods of test equating theoretically.
- Reviewing the testing agencies of foreign countries in charge of large-scale evaluations in terms of the scoring systems of each major test and equating methods: focusing on the NCA of the UK and the agencies in charge of the full-scale implementation in each state in the US (CTB, Pearson, ETS, and ACT).
- Collecting and comparing the overseas cases of designing data gathering and designing for equating in high-stake evaluations.

■ Plans for Resetting Scoring and Achievement Levels

- Comparing the scoring methods based on classical test theory and item response theory.
- Reviewing the scoring and scaling
- Analyzing and comparing cases of setting cut scores of achievement levels home and abroad.
- Reviewing the plans for resetting cut scores of achievement levels in academic achievement evaluations.

■ Reporting the Results and Plans for Their Utilization

- Reviewing the ways to report the results of the current NAEA.
- Analyzing the cases of reporting the results of other nations' academic achievement evaluations and the plans for utilizing the results.

- Proposing plans for reporting the results of the academic achievement evaluations from 2010.
- Exploring the ways to utilize the results of the academic achievement evaluations from 2010.

II. Need for Changes to the Scoring System in the National Assessment of Educational Achievement

The Scoring System of the Current Academic Achievement Evaluation

■ Academic Achievement Evaluation Design and Implementation

<Table 1> Design and implementation of the NAEA in 2009

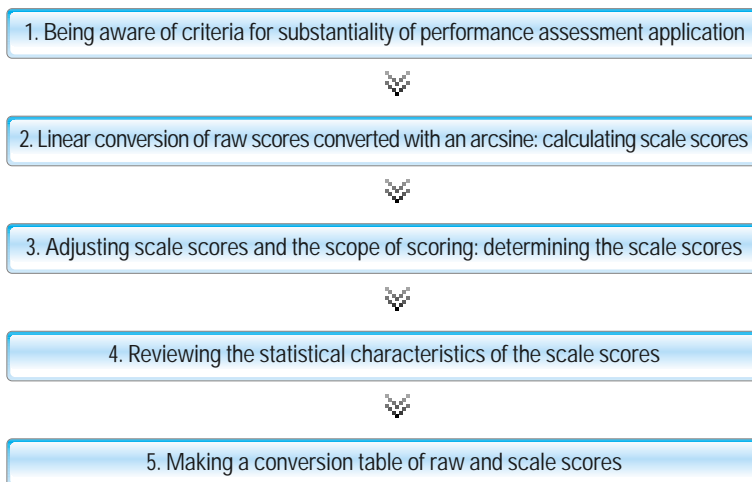
Subjects		Areas	Scope	Hours	Periods	Types of items	Tools
Elementary School 6th graders		Korean Social studies Mathematics	Entire curricula from E.S. 4th grade to 6th grade	40 minutes for each subject Survey : 15 minutes	October 13 and 14 (October 13 for vocational schools)	Multiple-choice and essay test	<ul style="list-style-type: none"> • Test for each subject : 2 types (A and B) • Questionnaire : 2 types (for students and for schools) Teachers are also included for the sampled schools and fill out a different questionnaire from the students.
Middle School 3rd graders			Science English (Survey)	Entire curricula from M.S. 1st grade to 3rd grade			
High School 1st graders	Academic school	Korean Social studies Mathematics English (Survey)	Entire curricula of H.S. 1st grade	70 minutes for each subject Survey : 15 minutes			
	Vocational school						

■ Goal Characteristics of the Evaluation Scale of the Current Academic Achievement

<Table 2> Goal characteristics of the evaluation scale of the NAEA (as of 2003)

Grade	Subject	Raw score characteristics		Goal characteristics of scale scores				
		Scope	Increment	Mean	Standard deviation	Minimum	Maximum	Increment
Elementary School 6th grade	Korean Social studies Math Science English	0-70	0.5	160	8.5	130	190	1
		0-70	0.5					
		0-70	0.5					
		0-72	0.5					
		0-60	0.5					
Middle School 3rd grade	Korean Social studies Math Science English	0-70	0.5	260	8.5	230	290	1
		0-70	1.0					
		0-70	0.5					
		0-72	0.5					
		0-72	0.5					
High School 1st grade	Korean Social studies Math Science English	0-70	0.5	360	8.5	330	390	1
		0-70	1.0					
		0-70	0.5					
		0-72	0.5					
		0-72	0.5					

■ Procedure of Developing the Scale of the Current Academic Achievement Evaluation



■ Stages of Achievement Levels in the Current Academic Achievement Evaluation

<Table 3> General achievement level skills of the NAEA

Achievement level	Characteristics
Advanced	Understands "most" of the basic content that students in the subject are expected to achieve.
Proficient	Understands "a large portion" of the basic content that students in the subject are expected to achieve.
Basic	Understands "parts" of basic content that students in the subject are expected to achieve.
Below basic	Below the basic academic abilities of students in the subject.

Potential Problems with the Scoring System after the Conversion to Full-Scale Implementation

■ Equating design:

Problems with gathering and designing equating data related to the difficulties with item security

■ Problems with scoring and scaling

- Problems with setting a set of criteria to connect curricular achievement level to scores of academic achievement evaluation.
- Problems with measurement science related to the scoring system.

■ Improving individual score reports:

It was pointed that the current reporting system did not provide detailed achievement information.

III. Improvement Measures for the Scoring System of the National Assessment of Educational Achievement

Plan to Improve the Gathering and Designing of Test Equating

■ Criteria for Selecting the Methods to Gather Test Equating

- Not using the data of “test equating for single groups” and “test equating for equivalent groups” in the gathering and designing process for test equating of annual tests.
 - When applying the non-equivalent group design with anchor items, external anchor items will be preferred to internal ones.
 - When using external anchor items in a non-equivalent group design with anchor items, the subjects should not be informed about which items will be included in the scoring process and which will not.
 - Feasibility should be considered when gathering and designing test equating data.
- Plan to Gather and Design Test equating Data in Academic Achievement Evaluation (security data)

■ Plan to Improve Scale Scores

- Proposing Ability Criteria based on Item Response Theory : Since the test equating process in the academic achievement evaluation is based on the item response theory, consistency can be achieved between the measuring model used in the test equating process and that used in the scaling process.
- Proposing Alternatives for Criteria Conversion based on Classical Test Theory

■ Resetting the Achievement Level

- The old way of combining the cut scores of achievement levels through test equating is no longer valid since the criteria of achievement levels were set in 2003.
- When adopting the ability scale of the item response theory, it is the right approach to apply the bookmark or IDM method based on the theory in order to set the achievement levels.
- When using the scale scores based on the raw scores, it is the right approach to adopt the Angoff method, which offers cut scores on the scale of raw scores, to set the achievement levels.

IV. Improving the Report of the Results of the National Assessment of Educational Achievement and Utilizing Them

Improving the Report of the Results of the Academic Achievement Evaluation

- **Basic Direction of Improving the Reporting of Students' Grades** : Visual reporting should accompany the current achievement levels for each subject and detailed information about their sub-areas.

<Table 4> Common features of grade reporting for each student

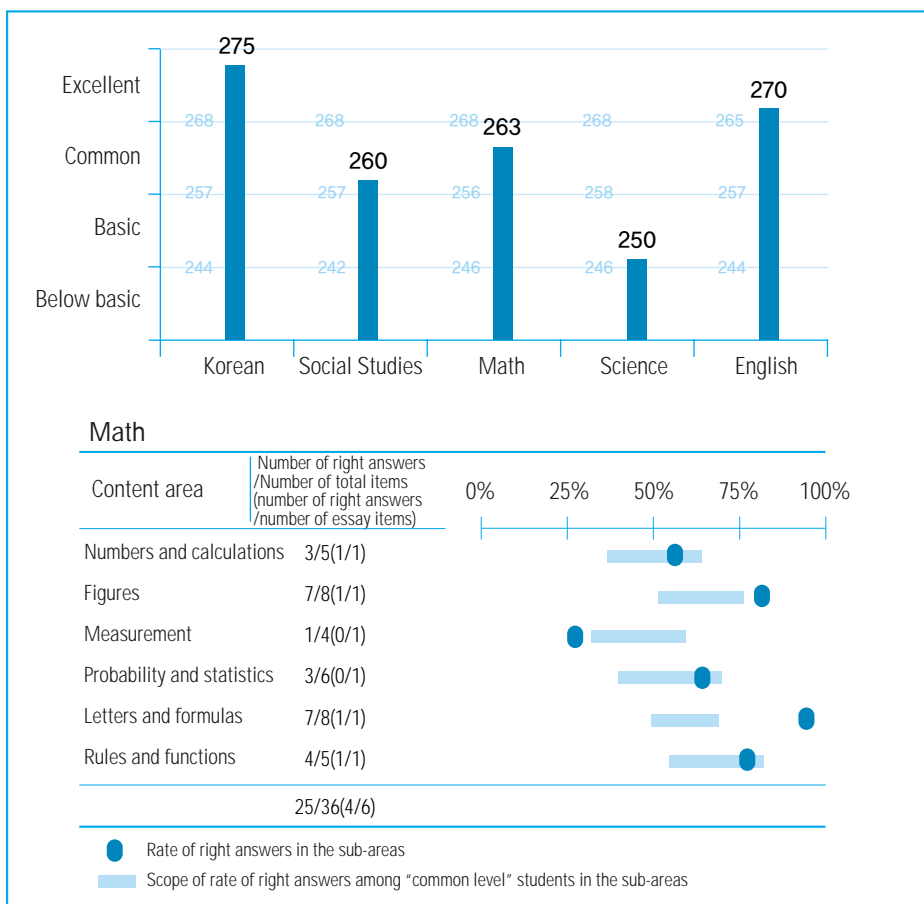
Category	Item	Content
Basic information	Student information	Name, grade, Student No.(student code)
	School information	Name and educational agency
Degree of academic achievement	Areas of evaluation	Title of subject
	Achievement level	Overall achievement level for each subject
	Subareas	Achievement level of the subareas included in the test
	Skills of achievement level	Explaining the ability and performance represented by each achievement level
	Goals of evaluation	Explaining the goals of the NAEA
	Graph	Visual reporting for achievement levels and places

- **Plan to Improve Student Grade Reporting Methods (Example)**

: Of the several improvement plans, Plan 2 determines the achievement level by subjects and offers the rate of right answers for each subarea.

Plan 2: Determines the achievement level by subject and offers the rate of right answers for each subarea.

Determines the achievement level (excellent, common, basic, and below basic) of individual students across five subjects and reports their scale scores, the number of total items in each subarea, and the number of right answers of multiple-choice and essay items.



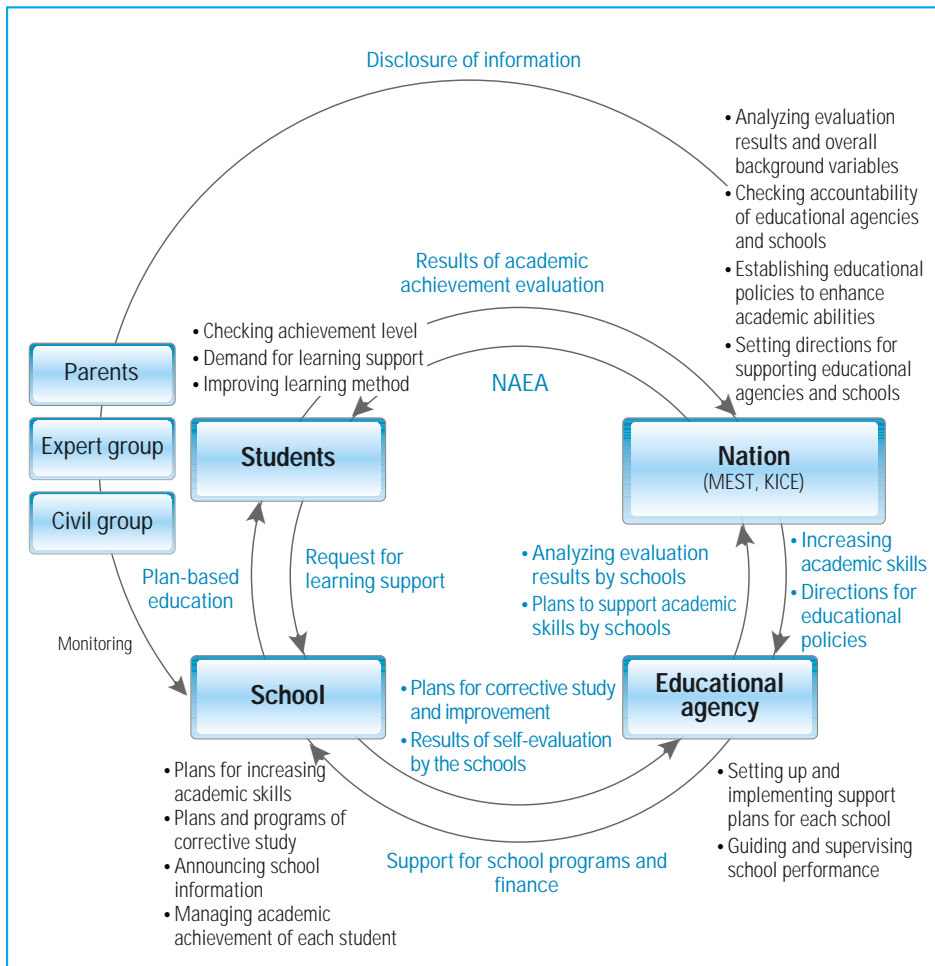
(Figure 1) Results of math achievement of Middle School 3rd graders in the academic achievement evaluation (Plan 2 was applied)

Plans to Utilize the Results of Academic Achievement Evaluation

■ Model to Utilize the Results of Academic Achievement Evaluation

: Disclosing education-related information will mark a revolutionary turning point in the history of information utilization. One will be able to compare schools and regions beyond any limitations and utilize the analysis results of relationships among the different variables of educational contexts when making decisions about educational policies.

- A model for utilizing the results will be suggested by making the links of a virtuous cycle to utilize the results of academic achievement evaluation in three dimensions.



(Figure 2) Model for utilizing the results of the NAEA

■ Implementation Plans for Each Subject of Utilization and Their Expected Effects

: Practical implementation plans for utilizing the academic achievement evaluation will be reviewed along with their expected effects by aligning such functions as feedback, planning, development, implementation, and application for the roles and results of those who are involved with academic achievement to each individual who utilizes them.

<Table 5> Implementation plans for each subject of utilization and their expected effects

Dimension	Subject	Goals	Utilization of results and implementation plans	Expected effects
Individual	Students and parents	<ul style="list-style-type: none"> Understanding individuals' academic achievement levels 	<ul style="list-style-type: none"> Utilizing the data on advancing to higher level education according to individual achievement levels Applying corrective programs Comparing achievement among subjects and understanding the advantages and disadvantages of subject content by area 	<ul style="list-style-type: none"> Examining school information and expanding choice Minimizing the number of students below the basic level
Educational agency	Teachers, headmasters and schools	<ul style="list-style-type: none"> Understanding individual students' academic achievement levels Understanding level differences among the subjects Understanding differences in areas of achievement in content among the subjects Understanding the degree of cross-sectional improvement of academic achievement 	<ul style="list-style-type: none"> Announcing the participation of individual schools in evaluation and their achievement levels by subject Applying the instructional programs to individual students Improving the teaching and learning methods by the contents of each subject Applying programs to increase academic skills or corrective programs Applying a support system by comparing the degree of improvement by year 	<ul style="list-style-type: none"> Checking and correcting teacher accountability to guarantee basic academic skills for each student Expecting overall improvement of students' academic skills
Region	Educational agencies	<ul style="list-style-type: none"> Understanding academic achievement level by schools Understanding academic achievement level in subjects within the region Understanding the degree of cross-sectional improvement of each school 	<ul style="list-style-type: none"> Supporting unit schools for their programs to increase academic skills Developing in-service training programs to help teachers to improve their teaching and learning methods Comparing schools and years in degree of improvement and providing support with budget and faculty talent? 	<ul style="list-style-type: none"> Checking and correcting unit school accountability within the region Inducing overall improvement of academic skills within the region
Nation	Government	<ul style="list-style-type: none"> Providing information about individuals' academic achievement levels Understanding academic achievement level of each educational agency and school across the nation Understanding the degree of cross-sectional improvement of each region Checking accountability of educational agencies for educational achievement 	<ul style="list-style-type: none"> Providing basic data for supportive policies by examining changes to academic achievement and comparing the degree of improvement Providing practical information to encourage study by examining each school within the jurisdiction of an educational agency for advantages and disadvantages Providing basic data to implement projects to increase basic academic skills 	<ul style="list-style-type: none"> Reinforcing quality control system of national education Checking the achievement of the National Common Basic Curriculum Implementing policies to guarantee national basic education on the national level Providing basic data to improve the curriculum and increase academic skills
Expert group	KICE and other concerned agencies and expert groups	<ul style="list-style-type: none"> Understanding the goals of the NAEA Understanding the meanings of the sample evaluation and full-scale evaluation Checking the framework of academic achievement evaluation and plans for utilizing the results Understanding the relationships among the background variables of academic achievement 	<ul style="list-style-type: none"> Analyzing the results of academic achievement evaluation in-depth Developing variety of educational indexes including one for academic achievement Developing criteria to evaluate agencies including the model to evaluate educational results and selecting models Providing a research framework for cross-sectional and longitudinal improvement Analyzing contextual factors to affect academic achievement Comparing and analyzing domestic and international cases for academic achievement? 	<ul style="list-style-type: none"> Proposing policies based on academic achievement and educational contextual variables Increasing school achievement and efficiency Expecting positive effects of the NAEA

V. Conclusion and Suggestions

Conclusion

- Each country in the world demonstrates its strong determination for accountability for public education by regularly checking the result or academic achievement of school education on the national level and providing corrective programs for individual students.
- They try to induce positive competitions and results among public educational agencies by disclosing information about the academic achievement levels of each student and school.
- The overall evaluation framework, development, implementation, and scoring should be done in a reliable manner, the legitimacy of the full-scale implementation should be secured, and checking each school's accountability should be understood in order to help bring about the model of a virtuous cycle for results for academic achievement evaluations to settle down in a stable fashion.

Suggestions

- There should be discussions with experts and independence should be given to the development of evaluative tools including the number of items, scope, and required time by subject in order to determine a new evaluation framework of academic achievement.
- There should be separate experiments and research on the sample size of preliminary studies in order to maintain test equating error at an appropriate level and minimize the possibility of item leakage.
- There should be a positive consensus for checking the accountability of school education and an effort to develop an effective evaluation model.
- There should be numerous follow-up studies within each research framework in order to devise meaningful educational policies with considerable data and applying them to schools by making active use of the expert group.
- KICE should play the role of a bridge between the expert group and the decision makers of educational policies and thus take active part in the establishment of educational policies to minimize trials and errors.