大韓數學敎育學會誌 數學敎育學硏究 第 13 卷 第 2 號
The Journal of Educational Research in Mathematics
Vol. X III, No.2, 107-122. May 2003

A Study on Alignment of Assessment with Curriculum and Instruction in Mathematics Education: A Synthesis of Research

Oh, Young Youl*

The present study aims at theoretically reviewing studies on the need of alignments among assessment, curriculum and instruction in order to implement the goals of mathematics education. This paper discusses about issues and models for alignment between instruction and curriculum, and then mainly focuses on alignment of assessment with instruction and curriculum at two levels, classroom and large-scale assessment. It was pointed out that both many published standardized tests for large-scale assessment and classroom assessments failed to meet curricular goals and contents, and instructional methods. The findings imply that reform in mathematics education should not be driven by assessment, but be embedded in instructional practices.

I. Introduction

Acknowledging the strong impact of assessment on implementing the goals of mathematics curriculum and instruction, mathematics educators (e.g., Webb & Romberg, 1992; Jones, 1996) have come to agreement on the need for redefining the roles and relationships of assessment under the current reform umbrella. The Curriculum and Evaluation Standards for School Mathematics (1989) and Assessment Standards(1995) by the National Council of Teachers of Mathematics [NCTM] are the products of those needs from the community of professional mathematics educators. In particular, the NCTM Standards give special emphasis on the importance of aligning assessment with curriculum and instruction both internally and externally. In other words, the new reform stresses the methods and tasks of assessment tools coherent with the curriculum and teaching practice.

The demands for changing how to assess whether mathematics education meets the goals of the mathematics curriculum seem to be a reflection of problematic results of assessments focused mainly on the products of learning. One of the critical issues surrounding such assessments is whether they sufficiently reflect curriculum goals and instructional process. For instance, Jones(1996) points out that a nationwide standardized assessment does not meet the goals of the mathematics curriculum, and test items are not aligned with actual classroom teaching. However, the more serious problem is that the results of such assessments tend to be used negatively as a judgement of teaching ability.

This paper begins the discussion by clarifying

* Gwangju National University of Education, yyoh@gnue.ac.kr

the goals of the mathematics curriculum recommended by the NCTM's Curriculum Standards(1989), and then, addresses issues of alignment of assessment with curriculum and instruction, followed by issues of alignment between curriculum and instruction. Assessment is typically classified as two types: classroom and large-scale assessments. Issues of assessment alignment with curriculum and instruction, therefore, will be discussed both in micro classroom contexts and in macro large-scale situations, including the relevance of reform, standardized testing, and the use of technology.

II. Goals of Mathematics Education

The NCTM's Curriculum and Evaluation Standards for School Mathematics(1989) provides a new vision for the K-12 school mathematics curriculum. This vision is articulated through five general goals for students: appreciating the value of mathematics, becoming mathematically confident, becoming a mathematical problem-solver, communicating mathematically, and learning to reason mathematically. All these goals are regarded as necessary factors for students to gain mathematical power. Mathematical power is clearly described by the NCTM Standard(1989) as "students' abilities to explore, conjecture, and reason mathematically, as well as the ability to use a variety of mathematical methods effectively to solve nonroutine problems" (p. 5).

The new mathematics curriculum proposed by the NCTM(1989) has been greatly influenced by the failure of past reforms and by new perspectives of how mathematics should be taught. For instance, the back-to-basics movement in the 1970s' focused on skill mastering has been considered as failure in general. Hence, other reform documents, such as Agenda for Action (NCTM, 1980) and Everybody Counts (National Research Council, 1989), propose that problem solving should be central to school mathematics. Simultaneously, behaviorists and cognitive scientists debated about whether teaching and learning mathematics should be focused on mastering skills or understanding mathematical concepts and structures(e.g., Gagné, 1983; Wachsmuth, 1983; Steffe, 1983). The results of these debates seemed to play a significant role in expanding both the new curriculum and the methods of teaching and learning appropriate to the new demands of the society. Therefore, it seems that the demands of the new mathematics curriculum and instruction are based on the reaction to prior experiences in the mathematics education community.

III. Alignment of Instruction with Curriculum

The NCTM's(1991) Professional Standards for Teaching Mathematics provides standards for how mathematics should be taught coherently in alignment with the goals and objectives of the curriculum. In other words, the focus is on how to improve problem solving, mathematical reasoning(e.g., inductive and deductive reasoning) mathematical communication, and dispositions toward mathematics. To do so, heavy emphasis is given to selecting worthwhile tasks, redefining the teachers' and students' roles in discourse, using

tools(e.g., computers) for improving students' understanding, and creating a good learning environment.

As is well known, the NCTM's standards series are reform-oriented documents. The emphasis from such documents is on radically changing what to teach and how to teach mathematics. However, one of the serious constraints in implementing the current reform curriculum arises from the conflict between the goals and objectives of the curriculum and the traditional ways of teaching still wide in practice. The findings from the sixth NAEP data demonstrate that the most popular way of teaching is still dominated by traditionally skill-oriented instruction, heavily dependent upon textbooks(Lindquist, 1997). As the goals and principles of the curriculum have changed, instructional methods should similarly follow such curricular principles.

Some reform initiative programs have attracted a great deal of attention for their great success among mathematics educators, since they demonstrate how mathematics should be taught to be reform-oriented teacher. They are, for instance, CGI(Cognitively Guided Instruction), QUASAR (Quantitative Understanding: Amplifying Student Achievement and Reasoning) and the Second Grade Classroom Teaching Experiment. Although each program has a different theoretical perspective on learning mathematics, instructional methods from those three programs are highly aligned with the goals, objectives, and content of the reform curriculum. One of the common features from all three programs is found in the change of teachers' and students' roles; that is, transition from teacher-centered to student-centered teaching and learning. A brief description of the central features of these three programs follows.

Cognitively Guided Instruction(CGI) aims to improve teaching practices by providing in-service teachers(K-4) with research-based knowledge of the domain of addition and subtraction about students' thinking in problem solving(e.g., Carpenter & Fennema, 1992; Fennema, Franke, Carpenter, & Carey, 1993; Fennema, Carpenter, Franke, & Carey, 1993). The CGI assumes that children construct knowledge by themselves; therefore, teachers should facilitate children's mathematical thinking and create a learning environment. In this program, learning usually takes place as the teacher and students interact with each other, while focused on problem solving, conceptual understanding, and mathematical reasoning. An important feature of CGI is that teaching practices are qualitatively improved as teachers' knowledge of students' learning increases.

Another successful program implementing reform-based instruction at the middle school level is QUASA(e.g., Silver, Smith, & Nelson, 1995). The theoretical perspective of this program is based on situated learning, in which learning comes through participating in a community of practice(Lave & Wenger, 1991). While learning in CGI is focused on individual's constructing knowledge, students in the QUASAR site are expected to learn mathematics wholly through social interactions within a community.

The last program of note is a longitudinal Second Grade Teaching Experiment conducted by Cobb and his colleagues(e.g., Cobb, Wood,Yac-

kel, Nicholls, Wheatley, Trigatti, & Perlwitz, 1991; Wood, Cobb, & Yackel, 1990, 1991; Wood, 1995). In implementing this experiment, Cobb and his colleagues take a combined perspective on learning of mathematics, including both radical constructivist and sociocultural perspectives. One of the important features in this experimental study is the assumption that learning emerges through classroom discourse. Students in this classroom are probably expected to gain a positive disposition toward mathematics.

IV. Alignment of Assessment

The NCTM's Curriculum and Evaluation Standards for School Mathematics(1989) clearly state the principle of assessment. According to the alignment principle, methods of assessing students' learning should be aligned with the goals, objectives, and content of the curriculum as well as with instructional approaches and activities, including the use of technology. The discussion on the topics of assessment alignment will be founded on the alignment principle, described above. The relevant issues will be discussed in light of both classroom and large-scale assessments and will include consideration of the roles and effects of the reform, tests, and technology.

1. Classroom assessment

A. Purpose. The primary purpose of classroom assessment twofold: improving instruc- tional practices by providing to teachers useful information on students' mathematical developm- ent and promoting and evaluating students' achievement (NCTM, 1995). That is to say, classroom assessment is closely related with the core of education in that assessment in the classroom is about how to improve the quality of students' mathematical learning. Regarding issues of implementing assessment, the NCTM's Evaluation Standards(1989) and Assessment Standards for School Mathematics(1995) place a great deal of emphasis on the necessary alignment of assessment with curriculum and instruction. The NCTM's Assessment Standards clearly define assessment as "the process of gathering evidence about students' knowledge of, ability to use, and disposition toward, mathematics and of making inferences from that evidence for a variety of purposes(p. 3)." The reform standards for assessing students' learning require a radical change both in methods and in tasks.

B. Curricular alignment. The key point of curricular alignment regarding classroom assessment is how to sufficiently assess the goals, objectives, and mathematical content of curriculum, including problem solving, mathematical reasoning, communication skill, and mathematical dispositions. What does it mean to assess students' understanding of mathematics? How can we do that? Assessment of cognitive thinking is inherently complex, so it requires the use of a variety of assessment sources. Although a variety of forms of classroom assessment(e.g., teacher-made tests, quizzes, and open-ended tests) have been used in the U.S., teacher-made tests have mostly dominated classroom assessment(Dossey & Swafford, 1993). However, Dossey and Swafford point out that this is problematic, since

teachers typically provide pre-chapter or post-chapter tests by developing their own tests or by following tests provided by publishers. The assessment standards under the current reform require that teachers use a variety of assessment methods and tasks, including quizzes and interviews, since cognitive skills might not be adequately assessed with only tests. More detailed discussion about assessment alignment with the goals of the reform curriculum will be dealt with in another section.

C. Instructional alignment. One of the most well-known slogans regarding assessment under the current reform movement is that assessment should be an integral part of instruction(NCTM, 1995). This implies that assessment of students' understanding should focus on the whole process of learning as dynamic. The reform documents regarding assessment emphasize the necessary alignment of assessment with instruction as well as with the curriculum, as considered above. What does it mean to align assessment with instruction? Regarding this question, Webb and Romberg(1992) stress both the validity of assessment and the learning environment. The validity of assessment covers the breadth of the topics taught and the range of expected students' learning outcomes, while the learning environment is related to the use of instructional materials, such as calculators and computers. For instance, if calculators are allowed during learning recommended by the curriculum, then assessment alignment with instruction means that calculators should also be allowed for assessment. More specifically, if a calculator is allowed in solving $y = 3x + 2$ during instruction, then likewise it should be allowed for the same purpose during the exam. As in the case of curriculum alignment with assessment, instructional alignment will be also discussed in greater detail in the following section on models of classroom assessment.

D. Models of classroom assessment.

Model 1: Cognitively Guided Instruction(CGI) is an instructional program designed to increase the quality of teaching practices by providing inservice teachers with research-based knowledge about diverse strategies and of how students learn (Carpenter & Fennema, 1992). The fundamental assumption in this program is that if teachers have enough such knowledge, then they will improve their teaching by making instructional decision-making on the basis of children's thinking and finally result in belief change. Knowledge of students' thinking provides teachers with a framework for teaching and analyzing children's knowledge. In this model, the main purpose of assessment is to gain knowledge that would be able to be used for structuring learning environment for individual children. Hence, assessment is significant and considered differently from traditional perspectives. Common beliefs about assessment from CGI are that teachers need to have knowledge about how students are thinking, teachers require having their own framework for how to assess students' process of learning, and assessment should be an integral part of instruction(Fennema, Carpenter, Franke, & Carey, 1993).

Aligned with the goals of the NCTM's Curriculum Standards(1989), the focus of CGI is on

enhancing students' problem solving, conceptual understanding of mathematics, and mathematical reasoning. In this program, procedural steps for skill mastering are embedded in the process of problem solving. The principles of assessment are (1) that the process as well as the product should be assessed; (2) methodologically, observing how children solve and explain problems can be used effectively to examine the mental process of problem solving; and (3) to do so, assessment can be implemented during individual interviews, group or whole class interactions, or seatwork activities(Carpenter & Fennema, 1991). Formal tests are rarely used for CGI students. Further, there is general agreement about the good alignment of CGI programs with the NCTM's Curriculum and Evaluation Standards for School Mathematics(1989) and other reform documents (e.g., National Research Council, 1989). Here is a well-described example of how assessment is used:

> During the first week of school, "James challenged her and himself." He had been working on one of his big sheets on which he was adding and subtracting 2 one-digit numbers (7 + 8 = ?, 9 - 4 = ?). He went up to Ms. J. and told her that these problems were "easy for him, too easy." Ms. J., "Well, what can you do then?" James went on to describe situations in addition and subtraction that he felt he could do and also said that he could work with bigger numbers.
>
> During a structured interview, Ms. J. checked to see if James could indeed do more difficult problems. She found that James could solve: Join and Separate (Result Unknown), Join (Change Unknown), Join (Start Unknown), and Compare problems with numbers ranging from 1 to 25 by using a counting strategy. James could also write numerals to 1,000. Ms. J. took the information that James had given her and which she had confirmed, and worked at building on his knowledge. On his Big Sheets, she began to include large numbers and situations in which there was a missing addend(Fennema, Franke, Carpenter, & Carey, 1993, p. 573).

Model 2: QUASAR is reform project whose goal is to improve the quality of mathematics teaching practices for middle school students in economically disadvantaged communities(Silver, Smith, & Nelson, 1995). The study starts by recognizing that improving students' mathematical learning might be achieved by developing new models of instructional practices. From this program, students are expected to learn a broad range of mathematical content, gain deeper understanding of mathematical concepts and ideas, and demonstrate an ability to solve complex problems as well as reason and communicate mathematically(Silver & Lane, 1993).

The purpose of assessment in QUASAR is mainly to monitor and evaluate the impact of the program over a long period(Silver & Lane, 1993). To do so, an assessment instrument is developed from an approach supported by the NCTM's Curriculum and Evaluation Standards for School Mathematics(1989). Given the goals and perspectives of the QUASAR project, diverse methods and tasks for assessment are applied: for example, paper-and-pencil tasks for a whole group, assessment of collaborative and small-group working, students' performance assessment, and use of manipulatives or calculators for assessment. QUASAR assessment may not be used for individual students; rather, it is used for evalu-

ating and enhancing the quality of the program. One of the important features of assessment is that the method of "holistic scoring" is adopted for students' responses to tasks(Silver & Lane, p. 65). Silver and Lane describe the general procedure of holistic assessment as follows:

> With respect to mathematical knowledge, attention is paid to the extent to which students demonstrate their knowledge of mathematical concepts, principles and procedures, such as understanding relationships among problem elements; using mathematical concepts as a basis for their reasoning; using appropriate mathematical terminology or notation; executing procedures; verifying results of procedures; generating new procedures or extending familiar procedures(p. 64).

Based on the principles of assessment described on the above, a sample task for the sixth grade drawn from QUASAR assessment looks like as follows:

> Yvonne is trying to decide whether she should buy a weekly bus pass. On Monday, Wednesday, and Friday, she rides the bus to and from work. On Tuesday and Thursday, she rides the bus to work, but gets a ride home with her friends. Should Yvonne buy a weekly bus pass? Explain your answer.(Silver, Smith, & Nelson, 1995, p. 41)

Busy Bus Company Fares	
One Way	$ 1.00
Weekly Pass	$ 9.00

In simply looking at the problem, we might expect that students could clearly respond with little difficulty. That is, accustomed at the traditional ways of learning, we simply expect students to say "no," since weekly pass ($9) will cost one dollar more than one way pass ($8). It might be possible, however, that student could say "yes" with their logic based on situations. For instance, if Yvonne needs to take a bus during weekends to go shopping, or if her family might need to share the pass, the answer can be logically "yes." This task suggests that there can be more than one correct answer, and answers may be dependent upon context of the problem.

E. Issues of concerns. Although many issues are implied by the above discussions, issues of implementing reform-based assessment are of critical importance.

One issue of interest arises from teachers' difficulty in using new forms of assessment. Experienced teachers learned mathematics in traditional ways, but they are required to implement the current reform from a new and different perspective. Romberg(1997) suggests that the current reform needs a new model for experienced teachers to learn how to implement new forms of curriculum, instruction, and assessment.

Another issue is how teachers' knowledge impacts the assessment of students' knowledge. Regarding this issue, Chambers(1993) points out the importance of teachers' having both structured knowledge of the mathematics domain and knowledge of students thinking. For "measurement division" and "partition division" problems, for instance, teachers who have limited knowledge of these distinctions are then restricted in assessing how much students' knowledge about this and, consequently, are restricted in using the results of

assessment for instructional decisions(p. 19). Ma (1999) demonstrates that teachers' limited knowledge of subtraction with regrouping limits assessing the students' learning.

Thirdly, a dilemma under the current circumstances of schooling arises as assessment drives instruction in practice. Glaser and Silver(1994) indicate that externally mandated tests still play a significant role over teachers as accountable, but reform-minded teachers may not be able to accept such narrowed ways of assessing students' cognitive thinking. This circumstance in which assessment conflicts with curriculum and instruction, makes reform-minded teachers have two curriculums, one for teaching conceptual learning and one for preparing externally mandated assessment.

2. Large-scale assessment

A. Purpose. The general purpose of large-scale assessments is aimed at assessing whether the things are going well in schools, not at assessing students' level of mathematical achievement as is the case in classroom assessment. In addition, large-scale assessment may result in modifying reform programs to enhance their quality. In this regard, assessment of this kind is obviously different from classroom assessment. While the latter is directly about teaching and learning, the former seems to be more directed toward gaining a broad sense of progress of school mathematics programs.

A well-known example of nationwide assessment is the National Assessment of Educational Progress [NAEP], a program of the United States Department of Education. The purpose of NAEP is to "monitor achievement trends in the nation and in the participating states(Jones, 1996)." The analysis of the NAEP data, carried out six times over the last thirty years, provides mathematics educators with precious information about how students' achievement has changed and how well tasks are aligned with the goals and content areas of the curriculum recommended by the current reform. More importantly, however, the results of the data provide guidance for the direction of reform. Although mathematics achievement has significantly increased in general at all levels over the last thirty years, Lindquist(1997) indicates that students' low conceptual understanding of mathematics might result from the unsatisfactory effects of teaching practices, which have been heavily dependent upon textbooks and rule-based skill mastering rather than upon problem solving, mathematical reasoning, and mathematical communication.

As calls for reform in school mathematics have been gaining more acknowledgement, each individual state in the U.S. administers statewide assessment to schools as another form of large-scale assessment, primarily in order to monitor the state of schooling. The forms of statewide assessment seem to be depending on the curriculum demands and goals of each state. For instance, statewide assessments in Massachusetts, consisting of multiple-choice, open-ended, and performance tests, are used for the purpose of comparing school effectiveness and guiding curricular and instructional improvement, while yet other states(e.g., Michigan, Connecticut) allow students to use calculators(Dossey & Swafford,

1993).

To determine the alignment of large-scale assessments with curriculum and instruction, one can adopt the alignment standard for assessment mentioned earlier. That is, it will be enough to compare the goals, objectives, and content of assessment with a curriculum and instructional methods. Again, large-scale assessment will be considered through its separation into the two subsections, curricular and instructional alignment of assessment tools.

B. Curricular alignment. As indicated above, the NCTM's Evaluation Standards(1989) clearly emphasize the necessary alignment of assessment with the goals and objectives of the curriculum in assessing students' understanding. In order to provide understanding of the effects of schooling, assessment should match the stated objectives of curriculum. Many researchers(e.g., Pandey, 1990; Jones, 1996; Romberg, Wilson, Khaketla, & Chavarria, 1992) show concerns about the unproblematic use of large-scale standardized assessment tools.

The primary issue, therefore, is whether nationwide or statewide assessments truly reflect the objectives and mathematics content of the curriculum as presented by the NCTM Curriculum Standards (1989). With respect to this big issue, there seems general agreement among mathematics educators that most large-scale assessments do not sufficiently reflect the goals and contents of the current reform curriculum. Romberg and his colleagues (1992) compare the alignment of the following six most widely used standardized tests for the level of grades 5-8 with the current reform curriculum in terms of alignment of goals and content areas. Those large-scale standardized tests investigated are the California achievement test, Stanford achievement test, Science research associates survey of basic skills, Comprehensive test of basic skills, and the Iowa test of basic skills. Romberg et. al. considered both the seven categories of content area and the goals of learning based on the NCTM's *Curriculum Standards*(1989) as a basis of comparison. In general, they conclude that the items from most tests do not adequately cover the range of mathematics content. These tests usually heavily emphasize basic computation rather than the higher order thinking, such as problem solving and reasoning, highly recommended by the new reform movement. In particular, 82 percent of the test items were drawn from the content area of number and number relations, while very few percent were found to be in the area of problem solving, reasoning, and communication skills. Glaser and Silver(1994) raise the same concern about the effectiveness of the findings of large-scale assessments, since achievement assessment rarely reflect the important goals of curriculum. Although mathematical reasoning and problem solving ability are strongly emphasized by numerous reform documents, it seems that most large-scale assessments continue the traditional approach.

At grade levels K-4, the focus of the curriculum is on learning conceptually, actively participating in doing mathematics, enhancing mathematical thinking and reasoning abilities, improving application of mathematics into

everyday life, and appropriately using calculators and computers. Based on the formats of multiple-choice, regular constructed-response, and extended constructed-response, the items from the sixth NAEP assessment were drawn from five categories of content domains as follows: Number and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and Functions. For example, here is one of the items from the area of Number and Operations, used for both fourth and eighth grades:

> By how much would 217 be increased if the digit 1 were replaced by a digit 5(Kouba, Zawojewski, & Strutchens 1997, p. 90)?
>
> A. 4 B. 40
> C. 44 D. 400

Only 36 percent of the fourth grade and 72 percent of the eighth grade students correctly answered this problem of place value. Although the item takes the form of a word problem, it is a little different from a simple computation. If the item were to be aligned with the focus of curriculum mentioned above, it would have to be more related to students' real life contexts. Furthermore, some critics(e.g., Jones, 1996; Dossey & Swafford, 1993) attribute the significant increase of U.S. students achievement to relatively easy items such as the above, which casts doubt on the effectiveness of the assessment results and may indeed cause the ceiling effects of achievement scores. Dossey and Swafford (1993) note that the tendency to increase the stakes associated with the comparison between states from the NAEP may narrow down the curriculum goals into only the topics covered by the assessment.

C. Summary. Large-scale assessment must be aligned with the curriculum. Although large-scale assessment has potentially great advantages for implementing the reform move- ment, there is a lot of evidence for a serious lack of alignment of those tests with the goals, objectives, and mathematical contents of the curriculum at any level. The lack of alignment between the two categories may occur, as we are trying to view students' learning outcomes, educated by the reform-based curriculum, from the never-changing format of standardized tests, or as it is strongly believed among policy makers, parents, students, and teachers that it is easy to use as well as to give reliable information. Considering all these things, assessment should be aligned in order to make schooling more effective and the effects of large-scale assessment reliable and understandable, and then the results of large-scale assessments will play a more significant role for improving the core goal of the current reform, that is, improving students' conceptual understanding of mathematics.

D. Instructional Alignment. The last consideration of alignment is whether large-scale assessment reflects important aspects of classroom teaching practices, including activities and diverse approaches. The NCTM's Curriculum Standards (1989) recommends that teachers use manipulatives, calculators, and computers for helping students engage in classroom activities and construct conceptual understanding of mathe-

matics. However, the relationship between large-scale assessment and instruction seems to be different from other cases discussed in previous sections. For instance, there are close relations between classroom assessment and the goals of curriculum and instruction, but large-scale assessment needs to be aligned with these goals.

Although there is general agreement about the use of teaching materials(e.g., calculators) during tests, issues of instructional alignment of large-scale assessment are usually talked about how the results of tests have a positive or negative effect on teaching practices. Many concerns are raised by mathematics educators(e.g., Romberg, 1997; Jones, 1996; Dossey & Swafford, 1993) with respect to its negative effect on teaching practices against the directions of the reform movement. Dossey and Swafford claim that the negative effects of large-scale achievement assessments on teaching practices are caused by their roles as measuring for the effectiveness of teachers to address the issue of accountability. However, most of these concerns are actually originated by the features of standardized tests, to be discussed later.

E. Critiques of standardized tests. According to Dossey and Swafford(1993), a standardized test is defined as a "commercially prepared test that provides items for obtaining samples of students' behavior under uniform procedure(p.49)." Examples are the *Stanford Achievement Test, Iowa Tests of Basic Skills, and California Achievement Test.* Most standardized tests are constructed with norm-referenced and multiple choice items. Though many standardized tests in educational settings have been very widely used in elementary and secondary schools in the U.S., many concerns from the education community are raised over its use and malfunctions in the curriculum, teaching and learning.

One of the most important issues around standardized tests is content alignment with the curriculum. As noted earlier, there is general agreement about the lack of standardized tests' content and goal alignment with curriculum. There could be several reasons for this. Dossey and Swafford(1993) indicate that this lack of alignment is partly due to the structure of textbooks, which heavily emphasize computations. In other words, the problem of misalignment seems to originate from the difference between the structure of currently used textbooks and the goals and content emphasized by the reform curriculum. This concern also gives rise to another issue of validity and reliability of standardized tests. Another reason for content mismatch is that standardized tests do not reflect the current theory of learning (Putnam, Lampert, & Peterson, 1990). While most of such tests are constructed on the basis of the behavioral psychology theory of learning, knowledge in the current reform movement is usually considered as constructed by learners. As the theory of how learning takes place has changed, the assumptions of standardized tests have been losing their foundations.

Another issue of concern is over the role and use of the results of standardized tests. In most cases with respect to this issue, researchers show concerns about its negative effects on the classroom teaching climate and teaching practices.

Particularly, as they are used as accountable evidence for the effectiveness of teachers and schooling, it intrudes seriously into instructional practices. Notice the following criticism by Wise (1988):

> By mandating educational outcomes through standardized tests, content through curriculum alignment, and teaching methods through teacher evaluation criteria, states set in motion a chain of events that alter educational ends and means.... A standardized test would set the educational objectives for the teacher. Curriculum alignment would insure that the teacher would cover the material to be tested.... Less obvious, however, are that distortions introduced into the curriculum by testing. Some teachers begin to emphasize the content that they know will appear on the test. They begin to teach in a format that will prepare students to deal with the content as it will be tested. Some even teach items that are likely to appear on the test. Meanwhile, the rest of the curriculum is deemphasized(cited as in Pandey, 1991, p. 42).

Although standardized tests were originally created to assess students' level of understanding and to help teachers' instructional decision-making, it seems ironic that the tests which were most widely used are now criticized from all sides. One of the reasons seems to be due to the radical change of paradigm for the perspective of knowledge, from positivistic behaviorism to post-positivistic cognitive perspective. It seems making sense to say that standardized tests are not sufficient to assessing students' understanding and mathematical thinking, but it seems still necessary under the current reform movement as part of holistic process of assessment for those goals of the reform curriculum, only if it is appropriately used.

Information from standardized tests is useful in classroom teaching in several ways. For instance, teachers can provide items from standardized tests to students and have them react to the questions. At the secondary level, multiple-choice exams or other forms of standardized tests might be useful in understanding mathematical concepts. For another way of use, Romberg(1997) suggests that results of tests can be used to motivate or manage students' learning. Although the current reform heavily emphasizes problem solving drawn from the contexts of students' real life, in some cases(e.g., learning abstract algebra), students must explore area not at all related to their prior experience.

V. Conclusions

Assessment should be aligned with both curricular goals and instructional methods in order to successfully implement the current reform ideas in mathematics education. Also, the results of assessment should be used to help students' mathematics learning and to make instructional decision-makings(NCTM, 2000). Based on this theoretical review of papers, three lines of alignments among assessment, curriculum and instruction were examined in the present study. Alignments of assessment were investigated in two dimensions, classroom level and large-scale nationwide level, and appropriate assessment models to meet the current curricular goals were considered.

Above all, the results of the present study implies that dominated teacher-centered instructional methods for mathematics need to be changed toward learner-centered way to meet the current curricular goals, focused on the quality of teaching and learning mathematics. Regarding classroom assessments, it was pointed out that, in order to improve the quality of instructional practices, teachers are required to use a variety of assessment methods(including tests, quizzes and interviews) and assessment should be an integral part of instruction. This study exemplified CGI and QUASAR as desirable models of assessment to meet the current reform ideas, regarding alignment with curriculum and instructional methods. Assessment used in these two programs was focused on measuring the process and product, methodologically observing how children think and solve problems, and imple- menting during instructional practices, such as classroom interactions. With respect to large-scale assessment, the findings of this study pointed out two things as follows; (1) items from these tests do not adequately meet the range of curriculum content and instructional methods, and (2) the results should be carefully used. Especially, some important concerns over the role and use of the results of standardized tests were raised.

In summary, the findings of this study strongly suggest that currently dominated classroom assessments need to be reconsidered toward appropriately reflecting the current curricular goals and instructional methods, and the results of classroom assessment are required to be effectively used for improving the quality of teaching practices. Also, the results of nationwide assessment should be carefully used, especially, in the context of assessment-driven education such as in Korea due to its negative roles in implementing reform projects, in teaching and learning in classroom climate, and in evaluating the effectiveness of teachers and the quality of schooling.

References

Carpenter, T. P. & Fennema, E. (1991). Research and cognitively guided instruction. In E. Fennema, T. P. Carpenter, & S. J. Lamon (Eds.), *Integrating research on teaching and learning mathematics*. State University of New York Press.

Carpenter, T. P. & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Research in Education, 17*, 457-470.

Chambers, D. L. (1993). Integrating assessment and instruction. In N. L. Webb, & A. F. Coxford (Eds.), *Assessment in the mathematics curriculum* (pp. 17-25). Reston, VA: NCTM.

Cobb, P., Wood, T. Yackel, E., Nicholls, J., Wheatley, G. Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*(1), 3-29.

Dossey, J. A. & Swafford, J. O. (1993). Issues in mathematics assessment in the

United States. In M. Niss (Ed.), *Cases of assessment in mathematics education: An ICMI Study* (pp. 43-57). Dordrecht: Kluwer Academic Publishers.

Fennema, E. Carpenter, T. P., Franke, M. L., & Carey, D. A. (1993). Learning to use children's mathematics thinking. In R. B. Davis, & C. A. Mahe r(Eds.), *Schools, mathematics, and the world of reality* (pp. 93-117). Boston: Allyn and Bacon.

Fennema, E. Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). Using children's mathematical knowledge in instruction. *American Educational Research Journal, 30*(3), 555-583.

Gagné, R. M. (1983). Some issues in the psychology of mathematics education. *Journal for Research in Mathematics Education, 14*(1), 7-18.

Glaser, R. & Silver, E. (1994). Assessment, testing, and instruction: Retrospective and prospect. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 20, pp. 393-421). Washington, DC: American Educational Research Association.

Jones, L. V. (1996). A history of the National Assessment of Educational Progress and some questions about its future. *Educational Researcher, 25*(7), 15-22.

Kouba, V. L. Zawojewski, J. S., & Strutchens, M. E. (1997). What do students know about numbers and operations? In P. A. Kenney, & E. A. Silver (Eds.), *Results from the sixth mathematics assessment of the National Assessment of Educational Progress.* Reston, VA: NCTM.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, England: Cambridge University Press.

Lindquist, M. M. (1997). NAEP findings regarding the preparation and classroom practices of mathematics teachers. In P. Kenney, & E. Silver (Eds.), *Results from the sixth mathematics assessment of the national assessment of educational progress* (pp. 61-86). Reston, VA: NCTM.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States (studies in mathematics).* Mahwah, NJ: Lawrence Erlbaum Associates.

National Council of Teacher of Mathematics. (1980). *Agenda for Action: Recommendations for school mathematics of the 1980s.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation: Standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics (1995). *Assessment standards for school mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (1989). Everybody counts: *A report to the nation on the*

future of mathematics education. Washington, D.C.: National Academy Press.

Pandey, T. (1990). Power items and the alignment of curriculum and assessment. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 39-51). Washington, DC: American Association for the Advancement of Science.

Putnam, R. T. Lampert, M., & Peterson, P. L. (1990). Alternative perspectives on knowing mathematics in elementary schools. In C. B. Cazden (Ed.), *Review of research in education* (Vol. 16, pp. 57-150) Washington, D.C.: American Educational Research Association.

Romberg, T. A. (1997). Mathematics in context: impact on teachers. In E. Fennema, & B. S. Nelson (Eds.), *Mathematics teachers in transition* (pp. 357-380). Mahwah, NJ: Lawrence Erlbaum Associates.

Romberg, T. A., Wilson, L., Khaketla, M., & Chavarria, S. (1992). Curriculum and test alignment. In T. A. Romberg (Ed.), *Mathematics assessment and evaluation: Imperatives for mathematics educators* (pp. 61-74). SUNY Press.

Silver, E. A., & Lane, S. (1993). Assessment in the context of mathematics instruction reform: The design of assessment in the QUSAR project. In M. Niss (Ed.), *Cases of assessment in mathematics education: An ICMI study* (pp. 59-69). Dordrecht: Kluwer Academic Publishers.

Silver, E. A, Smith, M. S., & Nelson, B. S. (1995). The QUASAR project: Equity concerns meet mathematics education reform in the middle school. In W. G. Secada, E. Fennema, & L. B. Adajian (Eds.), *New directions for equity in mathematics education* (pp. 9-56). New York: Cambridge University Press.

Steffe, L. P. (1983). Seeking meaning in mathematics instruction: A response to Gagné. *Journal for Research in Mathematics Education, 14*(1), 210-216.

Wachsmuth, I. (1983). Skill automaticity in mathematics instruction: A response to Gagné. *Journal for Research in Mathematics Education, 14*(1), 204-209.

Webb, N. & Romberg, T. A. (1992). Implications of the NCTM Standards for mathematics assessment. In T. A. Romberg o(Ed.), *Mathematics assessment and evaluation: Imperatives for mathematics educators* (pp. 37-60). SUNY Press.

Wise, A. E. (1988). Legislated learning revisited. *Phi, Delta, Kappan, 69*(5), 328-333.

Wood, T. (1995). An emerging practice of teaching. In P. Cobb, & H. Bauersfeld (Eds.), *The emergence of mathematical meaning* (pp. 203-227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wood, T., Cobb, P., & Yackel, E. (1990). The contextual nature of teaching: Mathematics and reading instruction in one second-grade classroom. *The Elementary School Journal, 90*(5), 497-513.

Wood, T., Cobb, P., & Yackel, E. (1991). Change in teaching mathematics: A case study. *American Educational Research Journal*, *28*(3), 587-616.

평가와 교육과정 및 교수방법의 일관성에 관한 연구: 연구의 종합

오영열 (광주교육대학교)

본 연구는 최근 수학 교육 개혁 운동에서 중요한 이슈가 되고 있는 교육 과정, 교육 방법, 그리고 평가 사이의 상호 일관성(alignment)의 필요성을 이론적으로 고찰하는데 그 목적이 있다. 특히, 최근의 수학 교육의 흐름에 비추어서 교실 환경에서의 평가와 국가 수준의 평가의 시각에서교육 과정과 교육 방법과의 일관성에 대해 논의하였다. 본 연구의 결과는 학생들의 수학 학습에 대한 평가는 다양한 방법을 활용하여 수업과 통합적으로 이루어져야 한다는 것을 말해주고 있다. 특히, CGI와 QUASAR에서 적용된 평가 모델은 평가가 어떻게 교육 과정 및 교육 방법과 일관성을 이루면서 효율적으로 이루어질 수 있는지를 보여준다. 국가수준의 평가에 대한 연구 결과는 과거에 사용된 대다수의 이러한 종류의 평가들이 표준화 검사를 적용함으로써 많은 문제점들을 내포하고 있었음을 지적하고 있다. 특히, 평가 문항과 교육 내용 및 교육 방법과의 불일치성의 예들은 국가수준의 평가의 타당성 및 그 활용에 대해 많은 우려를 포함하고 있으며 표준화 검사는 그 대표적인 예라고 할 수 있다. 결론적으로, 수학 교육 개혁의 핵심 과제인 수업 관행의 질적 개선을 위해서는 평가의 효율적인 활용이 필수적이라고 판단되며 이를 위해서는 평가와 교육 과정 및 교육 방법과의 일관성은 그 전제조건이라고 할 수 있을 것이다.

핵심어: 교수방법(instruction), 교육과정(curriculum), 국가수준평가(large-scale nationwide assessment), 일관성(alignment), 평가(assessment), 표준화검사(standardized tests),