NCRMSE Research Review
from the National Center for Research in Mathematical Sciences Education

Volume 1, Number 1
October 1991

In This Issue:

NCRMSE Begins Five-Year Program

You are reading the first issue of the NCRMSE Research Review: The Teaching and Learning of Mathematics. Designed as a newsletter, it will inform the mathematics education research community, educational policy and decisionmakers, and educators about the research and related activities undertaken by the National Center for Research in Mathematical Sciences Education. The *Research Review* will highlight specific NCRMSE programs, provide summaries of the knowledge developed by the programs, and bring recent national and international developments in mathematics education to the attention of readers.

All NCRMSE research programs emphasize building unified paradigms of study. Research on classroom instruction, student learning, curriculum, and assessment, according to NCRMSE Associate Director Thomas P. Carpenter, often have been conducted in ways that isolate the disciplines, creating separate and distinct streams of inquiry. The NCRMSE plan of research insures that teaching, learning, assessment, and curricular reform are approached from an integrated perspective.

Seven working groups carry out the NCRMSE activities: Thomas P. Carpenter and Elizabeth Fennema, University of Wisconsin-Madison, direct a group on the Learning/Teaching of Whole Numbers; James Kaput of Southeastern Massachusetts University directs the group on the Learning/Teaching of Algebra and Quantitative Analysis; Walter Secada of the Univesrity of Wisconsin-Madison directs the group on the Implementation of Reform; Judith Sowder of San Diego State University directs the group on the Learning/Teaching of Quantities; Richard Lehrer of the University of Wisconsin-Madison directs the group on the Learning/Teaching of Geometry; and Thomas A. Romberg of the University of Wisconsin-Madison directs the group on Models of Authentic Assessment. A Statistics group will begin its work next year.

Each working group involves a small group of productive researchers who share the same field of study, set priorities for research, recruit and train students, communicate with one another, and monitor the rapidly changing structure in a field.

Two of the working groups—Implementation of Reform and Models of Authentic Assessment—cut across or interrelate with the five content working groups. Each content-focused working group has developed preliminary research goals.

While the two cross-cutting working groups have designed preliminary research goals, additional tasks will develop through their relationships with the content-specific working groups. The Implementation of Reform Working group will examine how educational reform becomes integrated into classroom practice. It is undertaking specific studies of efforts to alter curriculum and practice in mathematics classes. These studies will identify the kinds of experiences, resources, and support systems teachers need if they are to carry out the reforms called for in the NCTM *Standards*.

The Models of Authentic Assessment Working Group is identifying a variety of models of assessment practices that are aligned, or in agreement, with the reform goals set out by the NCTM *Standards*. It will develop procedures for rating the validity, reliability, and utility of the models. The procedures will be used to judge models designed for program evaluation by states and schools or for instructional decisions by teachers. After potentially useful procedures are identified, the group will construct or adapt the assessment models, examine them for sensitivity to cultural and linguistic diversity, prepare aggregation and reporting procedures, and demonstrate the viable features of the assessment procedures to educators.

Thomas A. Romberg is responsible for the overall direction of the NCRMSE. Joan Daniels Pedro is assistant to the NCRMSE director, and Donald Chambers is director of dissemination. A National Advisory Panel of seven members advises the NCRMSE on the management of its research programs and reviews its work. Members of the Advisory Panel include Merlin Wittrock, chair, University of California-Los Angeles; Robert Davis, Rutgers University; Audrey Jackson, Parkway School District, St. Louis; Harvey Keynes, University of Minnesota; Jeremy Kilpatrick, University of Georgia; Mary Lindquist, Columbus College, Columbus, Georgia; and Edward Silver, University of Pittsburgh.

The NCRMSE is funded by grants from the Office of Educational Research and Improvement, United States Office of Education, Washington, D.C. In addition to publishing the *NCRMSE Research Review* and carrying out its 5-year research program, NCRMSE distributes research reports, publishes monographs relating to mathematics education, and provides informational programs to the mathematics education community.

For further information on NCRMSE contact Donald Chambers, Director of Dissemination, NCRMSE, Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison, 1025 W. Johnson Street, Madison, Wisconsin, 53706, or call (608) 263-0761.

New Book Cites Significant Advances in the Study of Teaching and Learning

"Connecting Mathematical Teaching and Learning," one of the chapters in Integrating Research on Teaching and Learning Mathematics, presents initial discussions on the development of a

unified paradigm for the study of teaching of mathematics that incorporates both cognitive and instructional research.

During the last decade, significant advances were made in the study of student learning and problem solving in mathematics, as well as the study of classroom instruction. Mathematics educators have been concerned that these two research efforts have been conducted as separate fields of inquiry. A number of them have agreed that there is an increasing need for an integrated research program that unites the two areas. Each of the book's eight chapters presents the perspective of its author on integrated research programs. The chapters include:

- "Research and Cognitively Guided Instruction"
by Thomas P. Carpenter and Elizabeth Fennema

- "Diversity, Equity, and Cognitivist Research"
by Walter G. Secada

- "Research on Learning and Instruction in Mathematics: The Role of Affect"
by Douglas B. McLeod

- "Curriculum and Teacher Development: Psychological and Anthropological Perspectives" by Paul Cobb, Erna Yackel, and Terry Wood

- "Connecting Mathematical Teaching and Learning"
by Magdalene Lampert

- "Methodologies for Studying Learning to Inform Teaching"
by James Hiebert and Diana Wearne

- "Intermediate Teachers' Knowledge of Rational Number Concepts"
by Thomas R. Post, Guershon Harel, Merlyn J. Behr, and Richard Lesh

- "Improving Research in Mathematics Classroom Instruction"
by Douglas A. Grouws

Editors of the book are Elizabeth Fennema and Thomas P. Carpenter, professors in the School of Education at the University of Wisconsin-Madison, and Susan J. Lamon, an assistant professor in the Department of Mathematics, Statistics, and Computer Sciences at Marquette University. It is part of a series, Reform in Mathematics Education, edited by Judith Sowder of San Diego State University.

Integrating Research on Teaching and Learning Mathematics
Elizabeth Fennema, Thomas P. Carpenter and Susan J. Lamon, Editors
1991, 142 pages, $12.95 paperback, $39.50 hardcover

Available from: State University of New York Press, c/o CUP Services,
P.O. Box 6525, Ithaca, NY 14851; phone (607) 277-2211.

(The book was developed at the National Center for Research in Mathematical Sciences Education with funding from the Office of Educational Research and Improvement, U.S. Department of Education [OERI/ED]).

Perspectives on Assessment

National and international attention increasingly is focused on the role of assessment in education. The test scores of American students, particularly the mathematics scores they achieve on standardized achievement tests, once were widely used to compare individuals, schools, and states. In the past, these comparisons—developed with psychometric sophistication—inspired public trust and were used to support varied political and educational agendas.

Today educational scholars in the United States and abroad are rethinking the educational value of tests made up predominately of multiple-choice items or of items not related to the instructional programs of schools. These scholars agree on the need to provide students, teachers, and parents with information about student performance in relation to a set of established standards. They also agree on the need to provide administrators and policy makers with information about how well the educational system is performing. Their professional attention is directed to developing alternate types of assessment items, tasks, response formats, and scoring rubrics. They are convinced that alternate assessment strategies also will require new psychometric approaches.

During the last year, national and international educational conferences have featured agendas with a primary focus on educational assessment:

- President Bush's education strategy is described in a publication, *America 2000*, released in March, 1991. The strategy calls for the development of national examinations, "American Achievement Tests," that will be used to assess student performance and to indicate the effectiveness of educational systems. The concern generated by this requirement led to the creation of a National Testing Council in July, 1991. The council consists of 32 members, 22 of them to be appointed by U.S. Secretary of Education Lamar Alexander.

- An International Commission on Mathematics Instruction (ICMI) held a conference of Assessment in Mathematics Education and Its Effects April 11-16, 1991, in Calonge, Spain. This conference provided a forum for discussion of the assessment tasks and strategies used by many countries, particularly Japan, Russia, and Australia.

- The Mathematical Sciences Educational Board (MSEB) held a National Summit on Mathematics Assessment in Washington, D.C., April 23-24, 1991. It was a political meeting designed to acquaint policy-makers with assessment issues and had the support of such prestigious associations as the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

- The National Center for Education Statistics (NCES) held and International Review of Preliminary Recommendations for the Third International Mathematics and Science Study (TIMSS) in May, 1991. The meeting was planned to enable the discussion of frameworks from which the questionnaires and items for the TIMMS would be developed. Experts from many countries participated. The results of the meeting may provide models of appropriate means of gathering data that will indicate how well educational systems are performing.

- The National Assessment of Educational Progress (NAEP) has been making state-by-state comparisons of levels of performance in mathematics and other instructional areas. The items they use have been viewed by many as an appropriate way of gathering data to indicate the quality of educational systems' performance. Those responsible for the revision of the test are weighing the questions of how well it presently is aligned with the NCTM *Standards* and how well it addresses the concerns regarding the limitations of multiple choice items.

- The New Standards Project intends to create a national examination system, as compared with a single examination. It will begin by involving a large number of people in establishing consensus on a framework. The framework will then be used to develop an examination and standards for grading. Developed by the Learning Research and Development Center and the National Center on Education and the Economy, the project will produce a "first draft" of a prototypic Grade 4 mathematics assessment framework that will have been reviewed by the Mathematical Sciences Educational Board's Study Group on Mathematics Assessment and by the MSEB Board. By the end of October 1991, the project will complete a "refined draft" of prototypic Grade 4 mathematics assessment materials and the needed ancillary materials such as scoring rubrics and teachers' guides.

The first issue of Research Review reports on some of the NCRMSE research on assessment. That research supports the need for some new thinking about assessment—particularly the kind exemplified by the New Standards Project. It also carries a related article by Susanne Lajoie that defines authentic assessment, develops a framework for its identification, and provides several examples of assessment strategies that hold promise as authentic forms for assessing mathematics learning.

Reviews of NCRMSE Research

Anecdotal evidence has long suggested that the tests required of their students by schools, districts, or states affect the instructional practices of teachers. Teachers, according to this evidence, "teach to the test." Several years ago the Mathematical Sciences Education Board held a national conference on The Impact of Testing on Mathematics Education. If tests influence what mathematics is taught and how it is taught, said conference participants, testing programs can be used to stimulate instructional change or they can become impediments to change. Three members of that group, Thomas A. Romberg, Jeremy Kilpatrick, and Tej Pandey, proposed a set of studies that would provide a research base and extend what is known about this issue. Several of the proposed studies on the influence of mandated testing on school mathematics instruction were recently completed by researchers at The National Center for Research in Mathematical Sciences Education. One of the studies of testing obtained information on the experiences and

the perceptions of a nationwide sample of 1,200 8th-grade mathematics tests with the NCTM *Curriculum and Evaluation Standards for School Mathematics*. The findings of the two studies are reported here.

Instructional Impact of Testing

The first study on the instructional impact of testing asked mathematics teachers whether mandated standardized mathematics achievement tests were given in their school settings. If tests were given, teachers were asked about their knowledge of the basic content and style of the mandated test taken by their students. It also asked about teachers' efforts to ensure that their students perform well on the tests, about adjustments they make in the curriculum to focus on the knowledge and skills on the tests, and about adjustments in their modes of instruction in response to the style and content of the tests. Teachers were also asked about the effects of testing on the use of problem solving and of calculators during instruction.

The use of mandated tests with 8th-grade mathematics students is wide-spread. Nearly 87 percent of the responding teachers indicated their students were given mandated mathematics tests. About 68 percent of teachers said their students took a state test and another 18 percent said their students were part of a state assessment program. Only 13 percent of teachers indicated their students were not required to take a mandated test in mathematics. Test results are used by districts and by teachers in a variety of ways. The results obtained by tests mandated by districts are used, said teacher respondents, by 80 percent of districts. They use them to group students by ability, compare them to national norms, or assign them to special programs.

Most teachers are provided with test data for individual students at an item or objective level. A majority of teachers think their districts are using the test results to stress what is tested or to deliberately affect teaching and the curriculum. When teachers think districts use test data to make decisions about students, teachers are much more likely to consider the style and format of the test as they plan instruction than when this is not the case.

Fewer than 20 percent of teachers say they make no instructional changes based on test results. Those teachers making instructional changes say that as a result of testing they increase their emphasis on some areas and decrease it on others. They spend additional time on basic skills, pencil and paper computation, topics emphasized on the test, and direct whole class instruction. They report a decreased emphasis on extended project work, problem solving, activities involving calculators, topics not emphasized on the test, activities involving computers, and cooperative learning. Finally, those teachers who give students practice test items and set aside time for students to prepare for the test are much more likely than others to reduce their emphasis on activities that involve computers and calculators.

Testing does have an impact on instruction, according to this national survey of 8th grade mathematics teachers. Teachers know what the tests call for and they plan accordingly. These results suggest that if districts adopt or prepare a test that sets standards of achievement they regard as suitable, a test with a style and format likely to promote the kind of mathematics the NCTM *Standards* seek, the use of that test to set standards and make decisions about students can also foster desired instructional changes.

NCTM Standards and Six Tests

Based on the first study of testing, we reported that tests have an impact on the instructional programs of a majority of 8th-grade teachers. A second NCRMSE study used standardized mathematics achievement tests that were identified in the first study. Only six of the tests—those identified as receiving the most widespread district usage nationwide—were examined.

Our purpose was to determine whether the widely-used tests reflected the recommendations of the NCTM *Curriculum and Evaluation Standards for School Mathematics* (1989). The NCTM *Standards* were developed as a means of improving the quality of school mathematics. The Boards of Directors of the National Council of Teachers of Mathematics established the Commission on Standards for School Mathematics. The Commission drafted a document that represents the consensus of NCTM's members about the fundamental content that should be included in the school mathematics curriculum and about key issues regarding student and program evaluation.

To complete the analysis of the tests, each item on each of the six tests was categorized by content, process, and level using the NCTM Standards for Grades 5-8. The content category included Numbers and Number Relations (nr), Number Systems and Number Theory (ns), Algebra (alg), Probability and Statistics (ps), Communication (com), Reasoning (rea), Connections (con), Computation and Estimation (c&e), and patterns and Functions (p&f). And the level category included Conceptual (conc) and Procedural (proc).

There were few differences across the six tests as shown in Table 1. When each item was identified according to content, process, and level the emphases of the tests did not vary. The percentages of the items on the individual tests fitting into the areas within each of the three categories were similar to the average percentages for all of the tests.

The results of this examination show that the six standardized mathematics achievement tests studied do not cover adequately the range of content called for in the NCTM *Standards* for Grades 5-8. A majority of items, 71 percent, fall into content area of Numbers and Number Relations. While 9 percent fall into the Measurement content area, between 3 and 6 percent fall into each of the remaining content areas. In the process category, a majority of items, 79 percent, are in the Computation and Estimation process area. While 20 percent fall into the Communication process, only 1 percent or less fall into the remaining process areas called for by the NCTM *Standards*, Problem Solving, Connections, Reasoning, and Patterns and Functions. An average of 89 percent of the items are classified as Procedural and 11 percent as Conceptual.

If tests are to reflect the new vision of the mathematics curriculum developed by Working Groups of the NCTM Commission on Standards for School Mathematics it will be necessary to vary content more than done at present; the processes of Problem Solving, Reasoning, Connections and Patterns and Functions will need to be added. An increase in the conceptual level of their items will also be needed. And if the NCTM *Standards* are to be implemented by schools, schools will need to select a set of standardized mathematics achievement tests different from those used by the majority of districts and states, or encourage the development of more

adequate tests by districts and states or by the developers of the standardized tests they currently purchase.

References

Romberg, T.A., Zarrinnia, E., & Williams, S. (1989, March). *The influence of mandated testing on mathematics instruction: Grade 8 teachers' perceptions.* Madison, WI: National Center for Research in Mathematical Sciences Education.

Romberg, T.A., Wilson, L., & Khaketla, 'Mamphono. (1991). *The alignment of six standardized tests with the NCTM Standards*. Madison, WI: National Center for Research in Mathematical Sciences Education.

A Framework for Authentic Assessment in Mathematics
By Susanne Lajoie

Vast differences exist between the tasks learned in school mathematics and those mathematicians or users of mathematics actually carry out (Pollak, 1987; Resnick, 1988; Lampert, 1990). Much of how we learn inside the classroom is different from how we learn outside of the classroom (Resnick, 1987). The focus inside the typical American classroom is on what the individual learner can accomplish independent of the group, or of any tools for learning such as calculators. In contrast, outside-of-classroom learning situations often are group situations where knowledge must be shared and where tools are available to enhance or extend our knowledge. Inside the classroom students are taught to manipulate symbols and abstract principles, but outside the classroom learning often is concrete and situated in the context in which it will be used. The term "authentic" has been used to suggest that some classroom activities are lacking realism and to conjure up an image of an alternate approach.

Requests for more authentic classroom activities have led to requests for authentic forms of assessment. These requests have come from several audiences. They range from student, teacher, district, and state sources to a national agenda on the integration of instruction and assessment. While the rhetoric is convincing, the images of authentic activities and assessment are still imprecise. This article is written to stimulate discussion on ways that authentic assessment can be operationally defined in the area of school mathematics.

The distinctions between in-school and out-of school learning have implications for defining authentic forms of assessment. In considering these distinctions we must also consider whether a framework for authentic assessment should incorporate a guide for authentic instructional activities in the classroom. Should we be concerned with mathematical knowledge that transfers to everyday uses of mathematics or should we consider authentic mathematics as something mathematicians do in their domain?

My primary focus in defining authentic assessment in mathematics is to provide a robust perspective of the individual learner's understanding of mathematics. Several audiences are considered as I define worthwhile mathematical tasks from the mathematical educator's

perspective, followed by a description of two interrelated theoretical perspectives on authentic activities as described in the literature on situated cognition and social constructivism.

Worthwhile Mathematical Tasks

The NCTM *Standards* represent goals for worthwhile or essential mathematics that are designed to make students mathematically powerful. These goals must be translated into tasks that exemplify authenticity. Only then can a framework for authentic assessment be developed.

A definition of an authentic mathematical activity emerges from the general assumptions of the NCTM *Standards.* One assumption is that knowing mathematics is doing mathematics. Doing mathematics refers to gathering and discovering knowledge in the course of solving genuine problems where knowledge emerges from experiences that are challenging but solvable. One way to increase such opportunities is to provide students with experiences in building mathematical models, structures, and simulations across multiple disciplines. Model building and discovering mathematical patterns are dynamic and constructive processes. Technology can be used to facilitate these cognitive processes as well as to record them. It can be used to assess developmental changes in reasoning, hypothesis formulations, verifications, and revisions. Technology can also serve as a medium for instructional manipulation where small changes in the instructional environment may account for changes in the learner's acquisition of knowledge.

The first four of the NCTM Standards were written as overarching goals that should be considered for all mathematics content at all levels. Any specific mathematical content, according to the four, should be designed to provide students with opportunities for mathematical problem solving, communication, reasoning, and connections.

Problem Solving

Activities that give students experience with problem solving can emerge from problem situations. These situations can be used to motivate students and serve as a context in which information is learned and knowledge is recreated across grades. Such situations imply complex, messy, and culturally-based problems that are open to multiple strategies and solutions (Zarinnia & Romberg, in press). Problems that are messy or ill-defined can provide more freedom for learners to pursue questions that reflect their personal interests. Such interests can be promoted by providing students with relevant, real-world applications. Real-world problems often include too little or too much information; they cannot be solved by applying a set of routine procedures.

Problem solving with mathematics involves modeling the problem and formulating and verifying hypotheses by collecting and interpreting data using pattern analysis, graphing, or computers and calculators. Technology is a powerful tool; it permits learners to manipulate data and see the consequences of their work in a few seconds.

Problem-solving activities need to include those that apply mathematics to the real world and those that arise from the investigation of mathematical ideas. Traditional curricula have emphasized mathematical ideas. The impetus for developing real and relevant problems stems from the need to contextualize mathematical concepts in a concrete rather than abstract manner.

These real world problems may take on cultural biases depending on the students who work with them. In addition to including applied and pure mathematical problem types, problem representations should be varied to provide for individual differences, i.e., verbal, numerical, graphical, geometrical, or symbolic, and permit several ways of reaching a solution.

Communicating

Communicating about mathematical ideas permits students to synthesize information about the ideas. There are a variety of modes of communication including reading, writing, discussion, and listening or concrete, pictorial, graphical, or algebraic methods. Activities which require students to communicate about mathematics provide them with opportunities to reflect on and clarify their own thinking and to develop a communal understanding of mathematical ideas and notations.

Students need opportunities to present ideas using language to insure that they understand words and their definitions and meanings. Teachers who structure classes to encourage communication provide students with opportunities to validate their thinking about mathematics. They can foster communication by asking questions, posing problems, or asking students to develop problems. Different levels of communication can be obtained by interviewing individual students, by using small groups, or by classroom discussions. These levels permit students to ask questions, discuss ideas, offer constructive criticism, and summarize discoveries in writing. Cultural and gender differences should be considered by those structuring activities to encourage communication.

Reasoning

Mathematics involves both inductive and deductive reasoning. Inductive reasoning is associated with mathematical creativity or invention. Deductive reasoning involves understanding the premises of a mathematical problem and reasoning logically with the given information. Challenging problem situations can provide opportunities for students to develop mathematical reasoning in a variety of contexts. The maturation of mathematical reasoning is a long process. Special developmental differences in reasoning, especially in Grades 5-8 where students move from concrete to abstract reasoners, must be planned for. The development of mathematical reasoning could be facilitated in both instructional and assessment settings if appropriate prompts, Why is this true? What if you changed this? Do you see a pattern?, and others are made available to learners.

Making Connections

A curriculum that integrates a broad range of mathematical topics rather than treating each topic in isolation is a connected curriculum. Number concepts, computation, estimation, functions, algebra, statistics, probability, geometry, and measurement become more useful to students when treated in an integrated fashion. Students can be helped to make connections between the topics if they are provided with contexts that require their integration when solving problems. It is not enough. However, to provide connections among mathematical topics; the connection of mathematics with other topics and such disciplines as science, music, and business is also necessary (Bransford, et. Al, 1988, Rosenheck, 1991). Teachers from other disciplines can help

to identify the mathematical ideals that can be explored in their domains. Geography, for example, provides opportunities for the use of scaling, proportion, ration, similarity, and other mathematical ideas. Using mathematics in contexts promotes attitudes of inquiry and investigation as well as sensitivity to the inter-relationships between formal mathematics and the real world.

Problem solving, communicating, reasoning, and making connections can be seen as curriculum goals that permeate the entire mathematics curriculum. Specific content areas also need to be addressed: number and number relations, number systems and number theory, computations and estimation, patterns and functions, algebra, statistics, probability, geometry, and measurement. In reviewing what the Standards deem worthwhile mathematical activities, it is important to realize that a single assessment of such activities will not provide a complete picture of a student's intellectual growth. Furthermore, different types of assessment are necessary to provide a complete picture of the learner's knowledge. In developing new forms of assessment, on must determine the types of assessment that are best for evaluating the various kinds of knowledge. Both individuals and small groups should be assessed, but for different skills. Small-group learning situations may be useful for measuring the ability to talk about and listen to ideas. Individual assessments may be better for assessing the learner's ability to synthesize knowledge.

Theories of situated cognition, social constructivism, and the influence of the group on the learning of individuals can be useful in defining authentic activities and authentic assessment. Although research on situated cognition is still in its infancy, there is evidence that certain activities described by its proponents are similar to those described as worthwhile by mathematical educators. Situated cognition refers to learning that takes place in the context in which one plans to use the knowledge. Problems must be realistic or authentic in the sense that the applications of knowledge are made apparent to the learner while the learning is taking place rather than outside of the context in which it could be used.

Situated Cognition

Situated cognition has developed out of the cognitive apprenticeship model of instruction (Collins, Brown, & Newman, 1989). The notion of a cognitive apprenticeship comes from traditional apprenticeships where novices learn their trade from a master. The masters share their knowledge with novices, assisting them in developing a product. Similarly, cognitive apprenticeships are designed around the notion that skilled learners can share their knowledge with less skilled learners to accomplish cognitive tasks. Cognitive apprenticeships, however, must model cognitive processes that are often difficult to externalize so that novices can observe or reflect upon the skills for a particular domain. In theory, the cognitive apprenticeship models offer suggestions for which skills to model for novices, how to provide scaffolding or assistance to less skilled learners, and when to fade such assistance when learners demonstrate they can construct their own meaning. Since the NCTM Standards call for an integration of instruction and assessment, the cognitive apprenticeship model has promise. It provides learners with ways to self-reflect and correct their performances based on assessment feedback. This theory does not provide specific guidelines for when and what type of feedback to offer or when to drop back

on the amount of assistance provided. If this theory were used to define authentic assessment in an operational way for mathematics knowledge, then such criteria would have to be developed.

Scaffolding or adaptive feedback is important in instruction and assessment. Vygotsky (1978) proposed that assessment consider both an individual's actual development or performance on a task without feedback and their potential development or performance on a task with feedback during test taking. With traditional assessment where learners' actual development is assessed, it would be difficult to differentiate between two learners who have the same score. The two learners could look quite different from one another if they were assessed in situations where limited feedback was provided in the test context. Assessment with feedback could measure the learners' potential rather than their actual performance. Learners may not need feedback the next time they are tested; thus the test would have become a learning experience in and of itself. This is a dynamic and adaptive form of assessment. It is dynamic since learners can be retested; it is adaptive since learners can learn from the test. Dynamic forms of assessment can provide feedback to learners, giving them ways to improve themselves and opportunities to reach their potential. Tests that serve a learning function may also improve learners' motivation and sense of self-efficacy.

Social Constructivism

The cognitive apprenticeship model is similar to the theory of social constructivism (Vygotsky, 1978). Learning occurs, according to the theory, when one shares cognitions with more capable peers. The NCTM *Standards* emphasize learners construction, verification, and revision of mathematical models. They also stress the importance of fostering problem solving, communicating, reasoning, and making connections through small group or whole class discussions. Situated cognition and social constructivist theories fit the NCTM *Standards* well.

Several researchers have examined the construction of mathematical meaning using small groups (Lampert, 1990: Resnick, 1988; Schoenfeld, 1985). Lampert discusses the importance of finding a common mathematical language for learners to use when communicating ideas. The group helps facilitate reasoning about mathematics. Resnick describes the importance of viewing mathematics as an ill-structured discipline where problem representations can be discussed and argued before mathematical procedures are employed. Resnick is particularly clear on the necessity of having a common core of knowledge in order to promote the types of dialogues that Lampert refers to in her work. Small groups can also foster reflection or the metacognitive skills necessary to evaluate mathematical problems (Schoenfeld, 1985).

The theories reviewed here provide great promise for building authentic activities as well as authentic assessments. There is a gap in the literature on how to operationalize these theories. It is difficult to design groups that will ensure the sharing of cognition and optimize learning for each group member. If more capable peers assist the less able learners by articulating their cognitive processes, we need to know how to design problem solving situations that will allow for the articulation of such processes, yet provide opportunities for the less skilled to participate in the overall task.

Authentic Assessment

Authentic assessment must take place in the context of the learning process. It must consider both the learning and the situation in which the learner is assessed. Authentic assessment must provide information on what the learner knows or does not know and the developmental changes in such knowing. Repeated measures of appropriate learning indicators must be made in order to obtain a robust picture of the learner's knowledge. These indicators must include a range of cognitive and conative abilities so that multiple perspectives are available for a particular area.

Authentic assessment will require instruments that provide in-depth perspectives on learning. Collins, Hawkins, and Frederiksen (in press) have begun to address the best tools for obtaining these perspectives. They suggest that one picture does not mean a thousand words when assessing what learners know. At least three different assessment mediums, they suggest, out to be use to obtain an integrated picture of the learners. The benefits of such mediums as paper and pencil, video, and computers jointly provide a more authentic picture of the learner than a single medium. Paper-and-pencil tests, the traditional form of assessment, are used to measure students' knowledge of facts, concepts, procedures, problem-solving ability, and text comprehension ability. Collins et al. (in press) suggest broader uses of these tests. Paper and pencil can also be used to record how students compose texts and documents of various kinds. Students traditionally have been assessed on their essays, but other writing tasks such as letters, reports, memos, drawings, and graphs can also be used to supplement compositions. Paper and pencil can also be used to assess how well students critique the quality of other documents.

Video can be used as a medium for assessing students' communication, explanation, summarization, argumentation, listening, and question-asking and question-answering skills. Video can also be sued to assess student interactions in the context of cooperative problem-solving activities. Video records of dynamic interactions can be scored at a later time. They provide opportunities for scoring oral presentations, explanations provided in a small group setting, and joint problem-solving activities.

The computer can provide yet another view of the learner. It can effectively track the process of learning as well as a learner's response to feedback. It can also simulate realistic situations in the classroom. The computer provides opportunities for assessing the dynamic nature of problem solving and opportunities to systematically vary the instructional environment on the feedback dimension and observe the effects on learning outcomes. The feedback dimension provides us with a novel mechanism for assessing how well or how poorly individuals respond to certain learning environments. The ability to track student performance provides opportunities for assessing such strategic aspects of knowledge such as hypothesis formation, hypothesis verification, or motivational aspects of learning—how persistent students are at trying to solve the problem—as well as actual learning outcomes. Thus computers provide opportunities for dynamic forms of assessment of those indicators that are determined to be the criteria for a successful performance.

Collins et al. (in press) suggest that the use of these three mediums of assessment will provide a more robust picture of the learner. The assessment media, however, is only as authentic as the task that the learner is being tested on. Care must be taken to define the types of student records

that will be collected with each medium, and to insure that such records reflect the performance indices that are most relevant to that medium.

Finally, the purpose or use of the assessment must be considered. If the use of assessment is by the learner and/or teachers, then such assessment tools must be available in the classroom on a regular basis, weaving together instruction and assessment. Learners should be able to use the tools to reflect on their strengths and weaknesses. Tests or assessment tools should be transparent in the sense that those who are being assessed understand the criteria on which they are being judged so they can improve their performance (Frederiksen & Collins, 1989). Frederiksen and Collins suggest that one way to ensure that assessment criteria are transparent is to provide a library of exemplars for students to visit. This library provides copies of records of student performances that have been critiqued by master assessors in terms of the relevant criteria. Such a library would help students evaluate their own performance and perhaps provide landmarks of success for which to strive. In addition to self-assessment, feedback should be given to students after a test is taken to help them improve their performances. Teachers can be assisted in using the assessment tools to determine what concepts students have misinterpreted.

Principles for Operationally Defining Authentic Assessment

We seeek to define and operationalize authentic assessment in order to improve learning. Thus, students should find undertaking an assessment task a learning experience. And teachers should learn what their students know or do not know as a result of the assessment task. Some tentative principles for operationally defining authentic assessment grow out of the theories and literature reviewed:

> 1) It must provide us with multiple indicators of the learning of the individual in the cognitive and conative dimensions that affect learning. The cognitive dimensions should include content knowledge, how that knowledge is structured, and how information is processed with that knowledge. The conative dimensions should address students' interest in and persistence on tasks as well as their beliefs about their ability to perform. Student interest in a topic often increases in conjunction with a deeper conceptual knowledge of that topic. Student choices may reflect their level of engagement and interest. These indicators must be examined repeatedly if they are to provide us with information on learning transitions or developmental maturity. Multiple mediums of assessment are necessary if we are to obtain valid indicators, i.e., that which we define as authentic. One measure, obtained by one medium, is unlikely to provide us with sufficient information on an individual. Varied types of procedures are necessary for gathering assessment information (Collins et al., in press; Romberg, in press).
>
> 2) It must be relevant, meaningful and realistic. It must be instructionally relevant, as indicated by its alignment with the NCTM standards. It must relate to pure and applied tasks that are meaningful to students and that provide them with opportunities to reflect, organize, model, represent, and argue within and across mathematical domains.
>
> 3) It must be accompanied by scoring and scaling procedures that are constructed in ways appropriate to the assessment tasks.

4) It must be evaluated in terms of whether it improves instruction, is aligned with the NCTM *Standards,* and provides information on what the student knows.

5) It must consider racial/ethnic and cultural biases, gender issues, and aptitude biases.

6) It must be an integral part of the classroom. Because teachers appear more likely to teach the information to students that appears on the tests, assessment tasks should be aligned with authentic activities such as those outlined in the NCTM *Standards*. Teachers need to be an integral part of the assessment loop so that they can learn from the assessment information and structure their instruction accordingly.

7) It must consider ways to differentiate between individual and group measures of growth and to provide for ways of assessing individual growth within a group activity.

Alternatives to paper-and-pencil multiple-choices tests do exist. Those listed here incorporate several principles of, and hold promise as authentic forms for, the assessment of mathematics learning:

Australian IMPACT Project

A set of studies were conducted in Australia to facilitate communication within the college level mathematics classroom (Clarke, Stephens, & Waywood, in press). Journals were kept by students and used by both teachers and students to foster a dialogue about what the students were learning. The quality of student journals progressed from simple narratives that described concepts to summaries that integrated mathematics knowledge, to dialogues regarding what questions should be addressed, what meaning could be constructed, as well as the connections of their work with other mathematics knowledge. These journals were beneficial to both teachers and students since they provided opportunities for dialogues that were not possible during a regular classroom session. They demonstrate that instruction and assessment can be integrated in the classroom. Student journals could provide us with new techniques for authentically assessing mathematical communication skills by providing the mechanism for examining transitions in developmental maturity in these skills.

Vermont Portfolios

Portfolios are promising as an assessment tool since they provide multiple examples of student work and provide students with experience in generating mathematical ideas, seeing mathematics as part of the culture, and being encultured into the mathematics experience. What is particularly intriguing about portfolios is the multiple audiences that can use them to obtain knowledge of the learners, teachers, and curriculum. Guidelines are needed, however, on how to score such materials.

California Assessment Program

The California Assessment Program (CAP, 1989) has addressed the concerns of the NCTM *Standards* with providing students opportunities to demonstrate their construction of mathematical meaning consistent with their mathematical development. Open-ended questions are provided that give students opportunities to think for themselves and to express their ideas. Communication is fostered in classroom discussions as well as in writing tasks. The data from this project provides a wealth of information regarding students' misconceptions and reasoning abilities.

Cognitively Guided Instruction

In the Cognitively Guided Instruction project (Carpenter, Fennema, Peterson, & Carey, 1987; Carpenter & Fennema, 1988) instructional decisions are based on careful analyses of student knowledge and the goals of instruction. Problems are selected that closely match the student's knowledge level. The assessment emphasis is on the learning processes of students. Individual and group data are collected.

Problem Situations

De Lange (1987) has designed mathematical problem situations comprised of multiple items with varied levels of difficulty. In his assessment of the Hewet Mathematics Project in the Netherlands, five different tasks were used to gather information: a timed written task, two-stage tasks, a take-home exam, an essay task, and an oral task. These provide a multifaceted evaluation of the learner. The two-stage tasks are especially interesting, in light of our principles of authentic assessment. Stage one includes open-ended questions and essay questions. These items are scored and returned to the student. In stage two, students are provided with their scores from stage one, allowed to take the stage one tests home, and given as long as three weeks to answer the same questions. The final assessment includes scores from stage one and stage two. Students can learn from their mistakes and from the feedback regarding their mistakes, making the testing process an interactive one that assists students in reaching their potential.

Superitems

Superitems are designed to elicit mathematical reasoning about mathematical concepts (Romberg & Collis, in press). The items are built to assess four different levels of mathematical maturity. At level four, the most mature level, the learner must articulate some understanding of the mathematical concepts either in words or symbols. The tasks can be used to obtain measures of developmental reasoning and serve as a first step in the identification of learning transitions in mathematical content areas.

I have laid out a tentative framework for the development of authentic forms of assessment. These and other alternative forms of assessment that incorporate new technologies hold promise for fitting within the operational definition of authentic assessment. Several parts of the framework require additional discussion or additional research. We will need to determine how cognitive and conative learning indicators can be operationalized in the context of an assessment task. We will need to study how to obtain frequent and valid measures of learners' performances. And we will need to define what we are assessing in individual and group

situations. Finally, when we are considering the multiple audiences that may use measures obtained by authentic means, we must keep equity issues in focus.

References

Bransford, J., Hasselbring, T., Barron, B., Kulewicz, St., Littlefield, J., & Goin, L. (1988). Uses of macro-contexts to facilitate mathematical thinking. In R. Charles, R. & E. Silver (Eds.), *The teaching and assessing of mathematical problem solving*. Hillsdale, NJ: Earlbaum.

California Assessment Program (1989). *A question of thinking: A first look at students' performance on open-ended questions in mathematics*. Technical Report, CAP, California State Department of Education, Sacramento, CA.

Carpenter, T.P., & Fennema, E. (1988). *Research and cognitively guided instruction*. Madison, WI: National Center for Research in Mathematical Sciences Education, Wisconsin Center for Education Research.

Carpenter, T.P., Fennema, E., Peterson, P.L., & Carey, D.A. (1987). *Teachers' pedagogical content knowledge in mathematics.* Paper presented at the American Educational Research Association, Washington, D.C.

Clarke, D., Stephens, M., & Waywood, (in press). Communication and the learning of mathematics. In T. A. Romberg (Ed.), *Mathematics Assessment and evaluation: Imperatives for mathematics educators*. Hillsdale, NH: Erlbaum.

Collins, A., Brown, J.S., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Cognition and instruction: Issues and agendas*. Hillsdale, NJ: Erlbaum.

Collins, A., Hawkins, J., & Frederiksen, J.R. (1991, April). *Three different views of students: The role of technology in assessing student performance*. New York: Center for Children and Technology, Bank Street College of Education.

de Lange J. (1987). *Mathematics insight and meaning*. Vak groep onderzoek Wiskundeonderwijs en Onderwyscomputercentrum, Rijksuniversiteit, Utrecht.

Frederiksen, J.R., & Collins, A. (1989). *Visualizing algebra: The function analyzer*. Pleasantville, NY: Sunburst.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, 27(1), 29-63.

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

Pollak, J. (1987, May). Notes from a talk given at the Mathematical Sciences Education Board Frameworks Conference, Minneapolis.

Resnick, L.B. (1987). *Learning in school and out. Educational Researcher*, 16(9), 13-20.

Resnick, L.B. (1988). Teaching mathematics as an ill-structured discipline. In R. Charles & E.A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 32-60). Reston, VA: National Council of Teachers of Mathematics.

Romberg, T.A. (in press). Evaluation: A coat of many colors. In T. Romberg (Ed.), *Mathematics assessment and evaluation: Imperatives for mathematics educators*. Hillsdale, NJ: Erlbaum.

Romberg, T.A., & Collis, K. (in preparation). Superitems.

Rosenheck, M. (1991). *The effects of instruction using a computer tool with multiple dynamically and reversibly linked representations on student's understanding of kinematics and graphing.* Unpublished Doctoral Dissertation, University of Wisconsin-Madison.

Schoenfeld, A. (1985). *Mathematical problem solving*. NY: Academic Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Zarinnia, E.A. & Romberg, T.A. (in press). A framework for the California Assessment Program to report students' achievement in mathematics. In T.A. Roberg (Ed.), *Mathematics assessment and evaluation: Imperatives for mathematics educators.* Hillsdale, NJ: Erlbaum.