PRINCIPLED PRACTICE

In Mathematics & Science Education

NATIONAL CENTER FOR IMPROVING STUDENT LEARNING & ACHIEVEMENT IN MATHEMATICS & SCIENCE

In This Issue

FOCUS

Assessment in Mathematics & Science

FEATURES

Comments From . . . Jan de Lange, Freudenthal Institute, The Netherlands........12

What's Up on the Web 10

For Further Thought 8

Center Mission 11

ARTICLES

The Changing Face of Assessment Mary C. Shafer & Sherian Foster............................1

What Should We Assess?..........2

Assessment Tasks and Analysis of Student Work 3

The Center's Role in Assessment Reform 7

Explore our Web site!

http://www.wcer.wisc.edu/ncisla

The Changing Face of Assessment

by **Mary C. Shafer & Sherian Foster**

The biggest industry and the smallest of businesses, even the service professions, need people with mathematical and scientific understanding and skills vastly different from those needed as little as a decade ago. No longer is shopkeeper math or a little general biology sufficient to meet the demands of living and working in a technology-driven information age.

There is increasing need for workers and an informed public with the ability, for example, to judge the reasonableness of a calculation result given by a calculator or computer, to think critically and make decisions based on statistical information in the news or on the job, to evaluate the myriad financial offers bombarding us daily, or to consider the validity of advertising claims said to be the result of a scientific study.

As school mathematics and science change (see *Principled Practice*, vol.1, no.1) to meet contemporary needs, assessment must also change. No longer is it sufficient for students to memorize and recite isolated facts, execute memorized algorithms to do paper-and-pencil calculations, or carry out prescribed laboratory routines. No longer is it appropriate to test only these skills. This newsletter looks primarily at mathematics and addresses the following: (1) What Should We Assess? (2) Assessment Tasks and Analysis of Student Work, and (3) The Center's Role in Assessment Reform.

Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison • 1025 W. Johnson Street, Madison, WI 53706

What Should We Assess?

Students need to develop understanding and proficiency ranging from understanding and using basic skills to thinking analytically and solving complex problems. This must involve learning and working with **important ideas** *in four domains of mathematics: algebra, geometry, number, and statistics and probability. A complete assessment program,* **over time**, *must measure and describe a student's growth and achievement in* **all** *of these areas.*

To assess the development of student thinking and understanding, it is helpful to define three levels of thinking as shown in Table 1. Questions that elicit Level I thinking are easy to pose and easy to grade: Because Level I questions often call for performing specific calculations, solving a given equation, or reproducing memorized facts, student responses are either right or wrong. Level I questions are often multiple choice or fill in the blank and are usually posed in isolation with no connection to a real or imaginable situation.

TABLE 1 Levels of Thinking

Level I ***Reproduction***

- knowing basic facts
- applying standard algorithms
- developing technical skills

Level II ***Connections***

- integrating information
- making connections within and across mathematical domains
- deciding which mathematical tools to use to solve problems
- solving nonroutine problems

Level III ***Analysis***

- mathematizing situations
- analyzing
- interpreting
- developing one's own models and strategies
- making mathematical arguments
- generalizing

Adapted from *Mathematics in Context* (1996–1998)

Because Level II questions require integrating information, making connections within and across mathematical domains, or solving nonroutine problems, they are harder to design, and student responses are harder to evaluate. Level II questions are more likely to be posed in the context of a real or imaginable situation, and they engage students in mathematical decision making. Teachers must understand each student's thinking and strategies and make judgments about the level and mathematical soundness of each student's work. Students' reasoning and solution paths may show qualitative differences.

Questions that elicit Level III thinking are the most difficult to design, and student responses are the most difficult to evaluate. Level III questions call for students to mathematize situations (recognize and extract the mathematics embedded in a situation and use that mathematics to solve problems), analyze, interpret, develop their own models and strategies, and make mathematical arguments and generalizations. Level III questions are designed to be open ended. More than one response can be considered "correct," and all reasoning must be supported with mathematical arguments. Level III questions are most likely to be posed in the context of a real or imaginable situation, and students often come up with novel solutions and strategies. Teachers must judge the soundness of each student's strategy and arguments.

Jan de Lange of the Freudenthal Institute, one of the originators of *Mathematics in Context* (National Center for Research in Mathematical Science Education & Freudenthal Institute, 1996–1998), developed a pyramid assessment model which we have adapted in Figure 1. Every assessment question can be located in the pyramid according to the level of thinking called for, mathematical content domain, and degree of difficulty. The X, for example, locates a Level I geometry question of medium difficulty.

Because assessment needs to measure and describe a student's growth and achievement in all domains of mathematics and at all three levels of thinking, questions in a complete assessment program, over time, should "fill" the pyramid: There should be questions at all levels of thinking, of varying degrees of difficulty, and in all content domains.

Note that although Level I questions are often easy and straightforward, they can also be hard and relatively complex. For example, for middle-school students to find the rate of travel, given time and distance traveled, is easy. They apply the $d = rt$ formula and solve. Finding the relative lengths of all sides of a right triangle, given the tangent of one acute angle, is much harder for middle-school students. Nevertheless, *Because assessment needs to measure and describe a student's growth and achievement in all domains of mathematics and at all three levels of thinking, questions in a complete assessment program,* **over time**, *should "fill" the pyramid:*

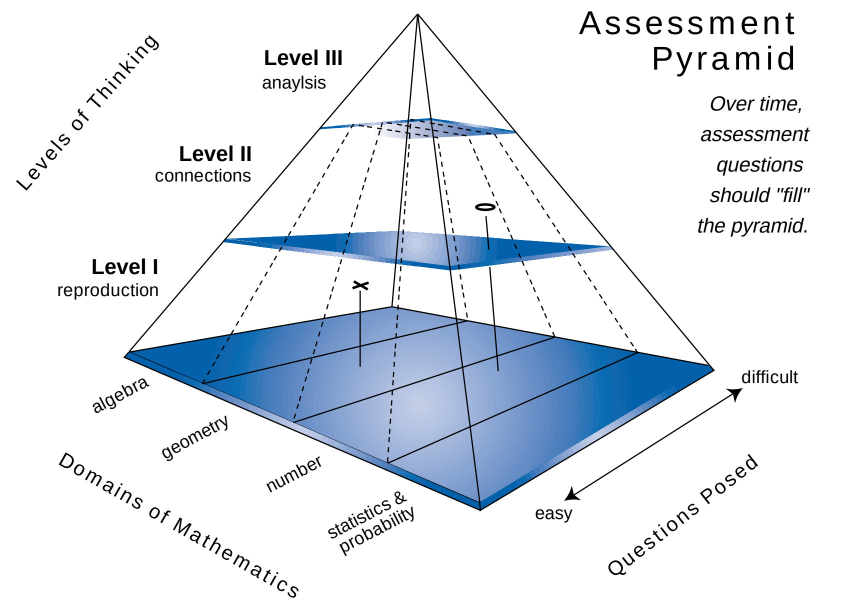


FIGURE 1 Adapted from Verhage and de Lange (1997).

There should be questions at all levels of thinking, of varying degrees of difficulty, and in all content domains. ◆

because the problem involves only computations using memorized relationships and algorithms, it is also a Level I question.

Look again at the Assessment Pyramid in Figure 1. Truus Dekker of the Freudenthal Institute, in an unpublished manuscript, pointed out that the three-dimensional model helps us see other important aspects of the nature of mathematics, the learning of mathematics, and the assessment of student understanding and achievement. When writing Level I assessment tasks, mathematical domains can be kept more distinct, and the difference between easy and hard questions can be great. As the level of thinking required increases, it becomes harder and harder to distinguish mathematical content domains and to devise questions that involve only one domain. Students must make increasingly more connections (and more complex connections) among domains. A geometry question, for example, may involve applying algebra knowledge; questions requiring interpretation of statistical information may require applying geometric knowledge. As the level of thinking required increases, the range between easy and hard questions becomes smaller.

The Assessment Pyramid enables us to visualize what is necessary for a complete assessment program. However, creating appropriate assessment tasks that do not test simply reproduction of facts and application of specific formulas and algorithms, but over time, also measure a student's growth and achievement will require retooling of assessment frameworks and tasks.

Assessment Tasks and Analysis of Student Work

Designing assessment questions, particularly those intended to elicit Level II or Level III thinking, is not easy.

The assessment question in Figure 2 (⬬ in Figure 1), from the *Mathematics in Context* Grade 6 unit, More or Less (Keijzer et al., 1997), is a number-domain question of medium difficulty designed to elicit Level II thinking. Specifically, it assesses the student's ability to find a certain percent of a number, to relate percents, fractions, and decimals, and the degree to which the student's number sense alerts her or him to unreasonable results.

continued on next page . . .

Assessment Tasks and Analysis of Student Work

Dale's store is having a sale on small fans that regularly cost $5.98 each.

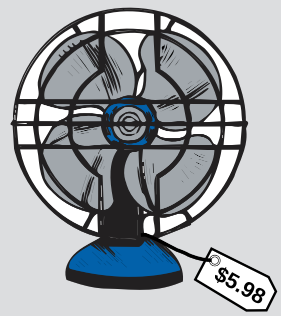


FIGURE 2 Used with Permission from Encyc

RESPONSE FROM STUDENT *A*

5% is $\frac{1}{20}$

$\frac{1}{20}$ x \$6.00 is .30

So the price is about \$5.70

50¢ off gives about \$5.50

$\frac{1}{5}$ x \$6.00 is \$1.20 So the sale price is about \$4.80.

Discount 3 is better.

STRATEGY

After rounding the price to $6.00, Student *A* did three sets of calculations then compared (approximate) final prices.

ANALYSIS OF RESPONSE FROM STUDENT *A*

After rounding the price to $6.00, Student *A* did the following calculations: (a) multiplied a fraction times a decimal, then subtracted; (b) subtracted; and (c) multiplied a fraction times a decimal then subtracted. Student *A* then compared (approximate) final prices. *A*'s response shows some understanding of the relationship between percents and fractions and demonstrates the student's ability to calculate using fractions and decimals.

Note that this written response alone might not be enough to completely assess the depth and sophistication of *A*'s understanding. Depending on what the teacher knows of *A*'s work and thinking from classroom activity and discussion, the teacher might need to interview *A* to get more information.

If asked "How do you know 5% is the same as $\frac{1}{20}$?" *A*'s response might indicate that this is an internalized fact or show whether *A* can explain, mathematically, *why* 5% is the same as $\frac{1}{20}$. Finding out *how A* calculated $\frac{1}{20}$ of $6.00 is also important. If *A* used a calculator, the keys used and the order in which they were pushed may reveal information about *A*'s understanding of multiplying fractions and decimals. For example, did *A* enter 1 ÷ 20 x 6.00, or 6.00 ÷ 20, or 6 ÷ 20, or 600 ÷ 20 (then use number sense to place the decimal)? *A*'s calculator procedure may yield some information about *A*'s understanding of the relationship between fractions and decimals. For example, did *A* enter .05 x 6.00? *A*'s response might indicate the use of number sense in knowing that .30 is a reasonable result for this calculation, and, if *A* did the calculation mentally, it could give the teacher information about the depth of A's number sense.

The written response alone might not be enough to completely assess the depth and sophistication of a student's understanding.

Depending on what the teacher knows of a student's work and thinking from classroom activity and discussion, the teacher might need to interview the student to get more information.

Customers can choose one of the following discounts.

DISCOUNT 1: 5% off

DISCOUNT 2: 50¢ off

DISCOUNT 3: $\frac{1}{5}$ off

Which discount gives the lowest price?

STRATEGY

After rounding the price to $6.00, Student *B* did two calculations then compared the discounts in terms of cents off.

RESPONSE FROM STUDENT *B*

Discount 3 is the best one.
I rounded $5.98 to $6.00
Then 10% is about 60¢ so
5% is about 30¢.
$\frac{1}{5}$ off of $6.00 is more than
$1.00 off.
So discount 3 is the best
one.

ANALYSIS OF RESPONSE FROM STUDENT *B*

Student *B* not only used a different strategy (comparing discounts) to solve the problem, *B* also apparently used strategies other than applying typical algorithms to make calculations. After rounding the price to $6.00, *B* used number knowledge to calculate 5% of $6.00. *B* first found 10% of $6.00, then (we surmise) took half of that to find 5% (30¢). Knowing two of the discounts, 30¢ and the given 50¢, *B* realized it was not necessary to calculate the third discount exactly. It was sufficient to know whether the third discount was more than 50¢. *B* knew that $\frac{1}{5}$ of $6.00 was more than $1.00 and, therefore, reasoned that the third discount was the best.

Again, the written response may not be sufficient to judge the depth and sophistication of *B*'s understanding. For example, asking, "How do you know $\frac{1}{5}$ of $6.00 is more than $1.00 off?" is particularly important. Perhaps *B* mentally calculated $\frac{1}{5}$ of $5.00 and, knowing $\frac{1}{5}$ of $6.00 is more than that, immediately chose discount three as the best discount. This could indicate deeper number knowledge than exactly calculating $\frac{1}{5}$ of $6.00 using paper and pencil or a calculator.

More Examples

The following examples, also from *Mathematics in Context,* are Level III questions. As you read these, imagine several possible student responses and analyze each as above. Consider also how you might design a Level III question and what student responses your question might elicit.

In the Grade 8 unit, Insights Into Data (Wijers et al., 1998), students are asked to cut a graph from a newspaper or magazine, including any caption or article that accompanies the graph. They are to "write a paragraph that explains how the graph more clearly represents the information in the caption or article, and whether you think the graph is fair or unfair." This Level III task is specifically designed to assess whether the student is aware of questions that should be asked when analyzing data sets and representations of data.

To complete this task, the student must first understand and make inferences about the information in the article: What information is there? Is it categorized? How? What graphical representations are appropriate for the data in the article? The student must consider whether the graph *fairly* represents the information in the article: Do the categories accurately reflect the article? Are both axes appropriately scaled? Are any bars or pictures proportionally drawn? Are comparisons accurately made?

continued on next page . . .

Assessment Tasks and Analysis of Student Work

In this task, students use Level III thinking as they develop their own strategies for analyzing the problem, decide what questions are important to ask, decide what representations of the data are appropriate, and select ways to judge the quality and fairness of the graphs presented. Students make mathematical arguments to support their inferences and conclusions based on their knowledge of statistics, number, and, possibly, algebra and geometry.

In the Grade 5 unit, Per Sense (van den Heuvel-Panhuizen et al., 1996), students are given the budget on the right for a fictitious country called Elbonia and, before being given a Level III question, are asked the following three questions: ▼

16. What percent of the total budget is this amount?

17. About $3 billion of the International Aid budget was missing at the end of the year.
What percent was missing?

18. It is possible that your answer to problem 17 differs from other people's answers. See if everyone got the same answer. How is it possible for there to be different correct answers to this question?

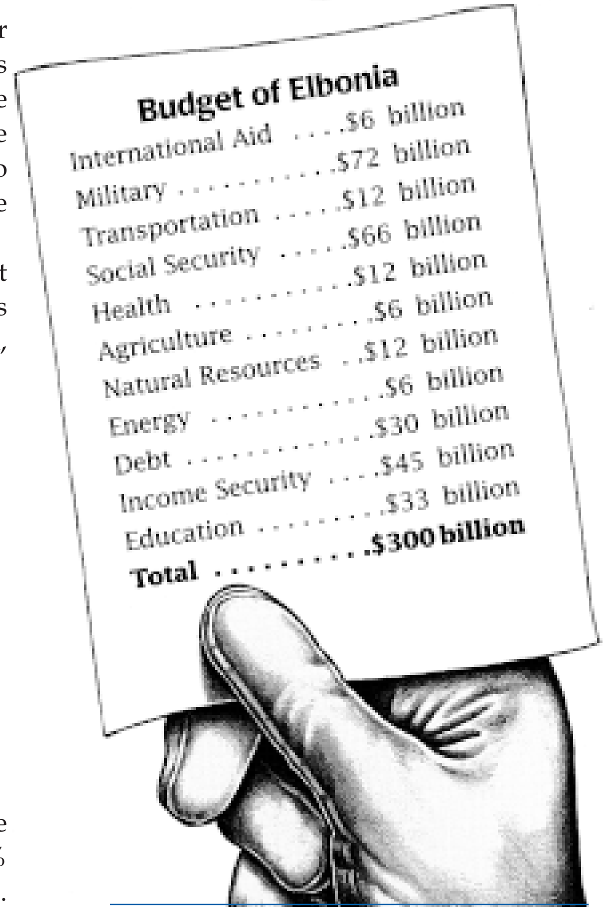
There is a problem, students are told, accounting for all of the money. An undercover detective takes Mr. Butler (who, for a 1% commission, delivers International Aid money for Elbonia), to dinner. Students are given the following information and questions: ▼

After dinner, the server brings the check to the table. The total is $20. Mr. Butler announces his intention to leave a 15% tip. First he gives the server a dollar. "That's 5%," he says. Then he adds a dime to the dollar. "This is another 10%, so altogether it is a 15% tip," he explains confidently.

The server is stunned and can't say a word. Mr. Butler looks at her with a smug expression. "You're welcome," he says. Suddenly, the detective jumps up and says "Aha! Now I know where the money went! You are under arrest!"

19. What did the detective figure out that could be used to convict Mr. Butler of fraud? In your notebook, write your answer as completely as possible so it can be used by the prosecuting attorney at Mr. Butler's trial. Include all the important information you know about percents so that the prosecuting attorney can convince the jury.

20. Is there any way that Mr. Butler could plan his defense? Explain.



Problems & Budget of Elbonia used with permission from Encyclopaedia Britannica Educational Corporation.

Question 19 is designed to elicit Level III thinking. Specifically it assesses the student's ability to determine whether percents are used appropriately in a decision-making situation and to use fractions, ratios, and percents as comparison tools. It also assesses the student's understanding of the relative nature of percent. Consider where you would locate this question in the Assessment Pyramid and why. Imagine possible student responses at different levels of thinking.

Analyzing and evaluating student responses to Level II and III questions can be as challenging and complex as designing the questions. Even though a task may be well designed to elicit Level II thinking, a student may give only a Level I response. For example, in responding to the question about the fans (Figure 2), a student might apply three different algorithms to calculate the exact price after each discount, then indicate which discount gives the lowest price. The student would not need to make any connections relating fractions, percents, and decimals nor would the student need to understand that comparing discounts (rather than final prices) is sufficient.

On the other hand, a student may go beyond what is called for and give a Level III response. For example, a student might explain why 5% and $\frac{1}{5}$ of something are different or explain how a discount of 50¢ might be more or less than 5% or $\frac{1}{5}$, depending on the total price of an item.

How then, other than practice and dialogue with colleagues, can we become proficient at analyzing and evaluating student work (as well as at designing assessment tasks)? The National Council of Teachers of Mathematics' *Curriculum and Evaluation Standards* (1989), *Professional Standards for Teaching Mathematics* (1991), and *Assessment Standards for School Mathematics* (1995) and the National Research Council's *National Science Education Standards* (1996) give guidance and examples. Fred Newmann, Walter Secada, and Gary Wehlage (1995) offer standards for assessment tasks and student performance which may also be of help. They include examples of how these standards were applied in their study of restructured schools. Examples of activities, units, and assessments used in Center research are also listed in our *Bibliography of NCISLA Research* (available on our Web site *http://www.wcer.wisc.edu/ncisla).*

Analyzing and evaluating student responses to Level II and III questions can be as challenging and complex as designing the questions.

The Center's Role in Assessment Reform

Traditionally, classroom assessment has focused on the factual information or algorithmic procedures a student can reproduce at the end of a unit of instruction, an arbitrary unit of time (week, grading period, semester), and judgments about students' knowledge have been based on whether their answers were either right or wrong.

Typically, assessment results have been collapsed into a single grade, assumed to describe the student's knowledge and achievement in comparison to other students. Research has shown that such grades reflect neither the full range, nor the depth and richness, of a student's knowledge and achievement and that almost all teachers take nonachievement factors (e.g., effort, disposition) into account when deciding a grade.

Current reform emphasizes a student's growth and development (over a much longer period of time) of a deep understanding of the "big ideas" of mathematics and science: Assessment is embedded in, and integral to, instruction. Thus, Center research focuses on classroom assessment and the ways teachers make judgments about students' knowledge, understanding, and progress. In particular, research will gather information about (a) what teachers use as evidence of student learning and understanding and whether they use multiple sources of evidence (e.g., classroom discussion, student's problem-solving strategies and procedures, student's oral explanations, student's reasoning and support of mathematical or scientific arguments, quizzes, tests); (b) whether teachers focus assessment solely on units of instruction (or units of time) or whether they also consider evidence gathered over longer periods of time (e.g., Grades 5–8); and (c) how teachers use information from multiple sources to determine grades and to talk to parents and others about what students know and can do.

Ultimately, according to Director Thomas A. Romberg, the Center's contribution to reform in assessment will be, primarily, in two areas. First, through *collaboration* with (rather than research on or dictates to) teachers in the field, the Center intends to produce a description of what it takes, not only to design and sustain classrooms that foster the development of deep and rich student understanding of the big

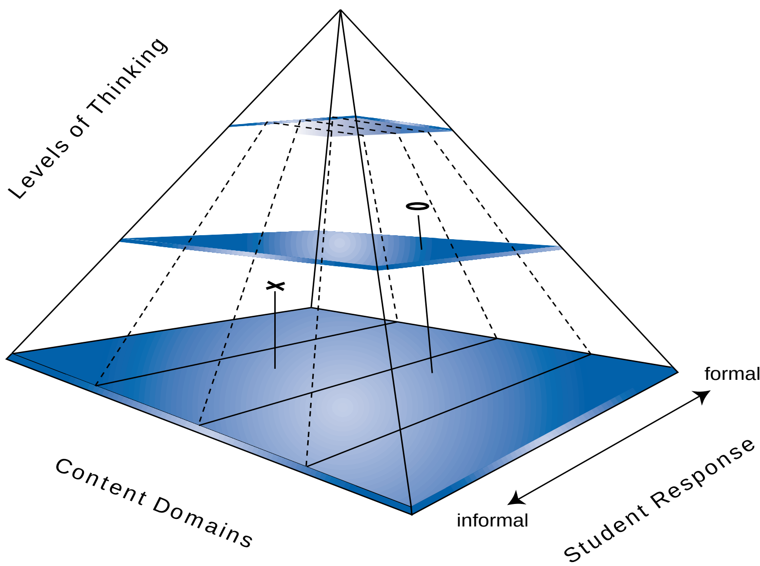
continued on next page . . .

The Center's Role in Assessment Reform

ideas in mathematics and science, but what it takes to assess that understanding. Second, the Center will disseminate its research findings to stakeholders and change-makers who can use it to make informed decisions. This will include teachers, parents, administrators, policymakers (e.g., local and state school boards and legislators), professional organizations in mathematics and science education, and agencies that help schools implement reform.

◆ ◆ ◆

For Further Thought



Pyramid adapted from Verhage and de Lange (1997).

The Assessment Pyramid can be helpful in looking at a complete assessment program. Assuming that assessment questions, *over time,* should "fill" the pyramid, perhaps this same pyramid can be helpful in analyzing student responses as well. Suppose, for example, we want to track a student's development from using informal and concrete reasoning and problem-solving strategies to more formal and abstract reasoning and strategies. Imagine replacing the dimension on the pyramid labeled *easy–hard* with *informal–formal*, then locating a student's response in the pyramid. We suggest that, *over time,* a student's responses should also "fill" the pyramid.

Can this be done for other dimensions (e.g., mathematical communication) that we want to measure or track? Can the pyramid be helpful in science assessment? Can the domains of mathematics be replaced with domains of science? Can the same three levels of thinking be used for science? At this time, we do not know. However, perhaps these and similar questions will help teachers, parents, administrators, or others think about assessing and tracking the development of deep and rich student understanding.

RESOURCE LIST

Keijzer, R., van den Heuvel-Panhuizen, M., Wijers, M., Shew, Julia A., Brinker, L., Pligge, M. A., Shafer, M. C., & Brendefur, J. (1997). More or less. In National Center for Mathematical Sciences Education & Freudenthal Institute (Eds.), *Mathematics in context.* Chicago: Encyclopaedia Britannica Educational Corporation.

National Center for Research in Mathematical Science Education and Freudenthal Institute. (Eds.). (1996–1998). *Mathematics in context.* Chicago: Encyclopaedia Britannica Educational Corporation.

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 Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Newmann, F. M., Secada, W. G., & Wehlage, G. G. (1995). *A guide to authentic instruction and achievement: Vision, standards and scoring.* Madison, WI: Wisconsin Center for Education Reasearch.

Verhage, H. & de Lange, J. (1997, April). Mathematics education and assessment. Pythagoras, 42, 14–20.

van den Heuvel-Panhuizen, M., Streefland, L., Meyer, M. R., Middleton, J. A., & Browne, J. (1996). Per sense. In National Center for Mathematical Sciences Education & Freudenthal Institute (Eds.), *Mathematics in context.* Chicago: Encyclopaedia Britannica Educational Corporation.

Wijers, M., de Lange, J. , Shafer, M. C., & Burrill, G. (1998). Insights into data. In National Center for Mathematical; Sciences Education & Freudenthal Institute (Eds.), *Mathematics in context.* Chicago: Encyclopaedia Britannica Educational Corporation.

Comments From . . . Jan de Lange, Freudenthal Institute, The Netherlands *(continued from back cover)*

An International Perspective on Improving Mathematics and Science Education in the United States

versions included a carefully selected set of anchor items, which appeared on both tests. Those items were selected on quality criteria and, as modeled by the assessment pyramid (discussed in feature article; see Figure 1), included questions from all four mathematical domains and at all three levels of thinking. Methodologically speaking, students' performance on the anchor items should have been the same for either version of the test, international or national. In reality, however, students seemed to fare much better on the anchor items when they took the national version. Until further research is done, we can only surmise that students felt more comfortable and more motivated when taking the national, culturally compatible, version of the test.

Second, standardized tests (including TIMSS) are *not balanced* in what and how they test: They have an overabundance of questions that require only lower-level thinking (refer to Assessment Pyramid in Figure 1), and they rely heavily on multiple-choice questions. For example, in the eight TIMSS rotation booklets used in The Netherlands, there were 429 multiple-choice questions, 43 short-response questions, and 29 extended-response questions. The bulk of the questions required little more than Level I thinking.

Finally, in the United States, changing the content and format of assessment may prove very difficult because it will mean changing what is valued. To an outside observer, many parents appear more data driven than child driven, and both parents and school boards seem to value students' scores on standardized tests more than students' intellectual development. As a result, when innovations are implemented, their effects are generally measured by the wrong instruments. Standardized tests, *as they are currently designed*, use multiple-choice questions to measure, primarily, low-level thinking. Such assessments should be avoided as much as possible. Parents, school boards, and administrators need to be educated about the strengths and weaknesses of any assessment program.

In 1989, then-President Bush and 50 governors articulated a national goal: that the United States would be first in the world in mathematics and science achievement by the year 2000. The time line for this goal seriously underestimates the problems involved and the time needed to make solid educational change. Curricula, textbooks, teacher preparation, classroom instruction, professional development, and assessment *all* must change, and many players are involved — parents, school boards, state and national policymakers, teachers, and the general public. Only if all players are willing to consider fundamental changes in U.S. mathematics and science education will change be possible. The process will take 20 years, *but it must start today.*

Principled Practice *in Mathematics and Science Education* is published by the National Center for Improving Student Learning and Achievement in Mathematics and Science, Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison.

This publication and the research reported herein are supported under the Educational Research and Development Centers Program, PR/Award Number (R305A60007), as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the supporting agencies.

Permission to copy not necessary. ◆ This publication is free upon request.

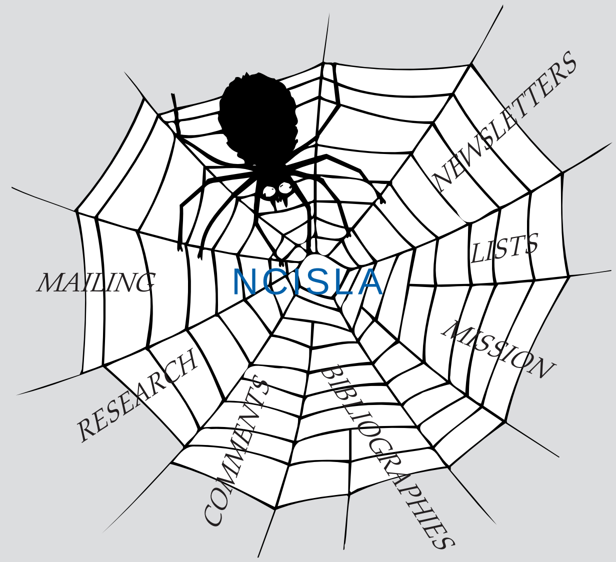
Address inquiries to the Center, Attention: Judith Peterson, or to our E–mail address.

What's Up On the Web?

NCISLA is on the Web !

http://www.wcer.wisc.edu/ncisla

- Find out about the Center and its mission.
- Read or download our newsletters.
- Read or download our bibliographies.
- Find a list of Center researchers.
- Offer your comments.
- Add your name to our mailing list.



Coming Soon !

- ***Searchable bibliographies.***
- ***General information about the work of our Design Collaboratives and Study Groups.***
- ***Links to affiliated institutions.***
- ***Links to other mathematics and science education Web sites.***

NCISLA UNDERLYING BELIEFS

Central to the Center's mission is the belief that there is a direct and powerful relationship between student understanding and student achievement. In fact, we have high expectations for all students, and we believe the way to high student achievement rests on students' understanding of important mathematical and scientific ideas taught in school classrooms by professional teachers.

More specifically, we believe the following:

1. All students can and must learn more, and somewhat different, mathematics and science than has been expected in the past. In particular, all students need to have the opportunity to learn important mathematics and science regardless of socioeconomic class, gender, and ethnicity.

   Our society has long underestimated the capability of all students to learn both mathematics and science.

2. Some of the important notions we expect students to learn in both disciplines have changed. This is in large part due to changes in technology and new applications for mathematics and science. Thus, at every stage in the design of instructional settings we must continually ask, "Are these *important ideas* in mathematics and science that students need to understand?"

3. Technological tools increasingly make it possible to create new, different, and engaging instructional environments. Technological tools include not only calculators and computers, but also a wide range of things such as manipulatives and other hands-on materials in mathematics, lab equipment in science, distance learning via satellite broadcast, audio- and videotapes, measuring instruments, building materials, access to natural resources, new ways of grouping students and new possible assignments (because technology gives teachers new ways to monitor student work and new things students can produce).

4. Student understanding develops as a result of students' building on prior knowledge via purposeful engagement in problem solving and in substantive discussions with other students and teachers in classrooms.

5. Real reform in the teaching and learning of mathematics and science will occur only when the advocated changes in content, work of students, role of teachers, and assessment practices become *common practices* in school classrooms.

6. Such reforms will happen only if teachers are professionally supported by other teachers, administrators, parents, and the public.

◆ ◆ ◆

CENTER MISSION

The Center's mission is to craft, implement in schools, and validate a set of principles for designing classrooms that promote student understanding in mathematics and science.

◆ ◆ ◆

To achieve this mission, we are conducting a sustained program of research and development in school classrooms *in collaboration* with school staffs to do the following:

1. *Identify a set of design principles.*
2. *Demonstrate,* **in classrooms,** *the impact of the design principles on student achievement.*
3. *Clarify how schools can be organized to support teaching for understanding.*
4. *Develop a theory of instruction related to teaching for understanding.*
5. *Find ways to provide both information and procedures for policymakers, school administrators, and teachers so they can use our findings to create, and sustain, classrooms that promote student understanding.*

CENTER INFORMATION

Director

Thomas A. Romberg
University of Wisconsin–Madison
Madison, WI

◆ ◆ ◆

Associate Directors

Angelo Collins
Vanderbilt University
Peabody College
Nashville, TN

Richard Lehrer
University of Wisconsin–Madison
Madison, WI

Walter G. Secada
University of Wisconsin–Madison
Madison, WI

◆ ◆ ◆

Associated Institutions

Michigan State University
East Lansing, MI

TERC
Cambridge, MA

University of Massachusetts–Dartmouth
North Dartmouth, MA

University of Pittsburgh
Pittsburgh, PA

University of Wisconsin–Madison
Madison, WI

Vanderbilt University, Peabody College
Nashville, TN

Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison, 1025 W. Johnson Street, Madison, WI 53706

◆ *Phone: 608-263-3605 / Fax: 608-263-3406* ◆

E–mail: ncisla@mail.soemadison.wisc.edu
http://www.wcer.wisc.edu/ncisla

NEWSLETTER INFORMATION
Volume 1, Number 2 ◆ Fall 1997

Dissemination Coordinator
Sherian Foster

Newsletter Production
Judith Peterson

Comments From . . . Jan de Lange, Freudenthal Institute, The Netherlands

An International Perspective on Improving Mathematics and Science Education in the United States

The performance of U.S. students in the Third International Mathematics and Science Study (TIMSS) has caused quite a stir. Although on a par with other major industrialized nations (Canada, England, and Germany), the United States was outperformed in mathematics and science by birds of a very different feather: three Asian countries (Singapore, Korea, and Japan), five Central and Eastern European countries (Czech Republic, Hungary, Austria, Slovenia, and Bulgaria), and one industrialized country (The Netherlands). The variety of this group makes it difficult to formulate recommendations about improvement of mathematics and science education in the United States, but some of the problems, focusing here on assessment, can be identified.

First, can any single test (in this case, TIMSS) measure the quality of mathematics and science education in all countries and across vastly different cultures? It is now generally agreed that mathematics is neither value nor culture free, as was long assumed. We cannot, however, quantify the *fit* between the TIMSS test items and a country's implemented or intended curriculum. A question meaningful in one culture might make no sense at all in another culture. This can affect both the results and the conclusions drawn from those results.

The Netherlands addressed this issue by designing a "national option" test, which more closely matched Dutch students' cultural experience and understanding. The national and international test

continued on page 9 . . .

National Center for Improving Student Learning and Achievement in Mathematics and Science
Wisconsin Center for Education Research
School of Education, University of Wisconsin–Madison
1025 West Johnson Street, Madison, WI 53706