# Chapter 9
# Learning from assessment

Richard Askey*

There are many types of assessment and books can and have been written about them. My aim is modest, to point out a few instances of what can be learned by looking at assessments given in the past.

## General considerations

One important use of assessment is to find out where education has been weak. As an example, consider the results in geometry in the 1995 Third International Mathematics and Science Study (TIMSS) assessment of knowledge of advanced students in the last year of high school, (Mullis et al., 1997). The average score for the 17 countries was 500. The scores ranged from 424 to 548. The four countries with lowest scores on advanced mathematics and percentage of students in advanced mathematics courses are shown in Table 1.

Table 1. Countries with lowest advanced mathematics scores

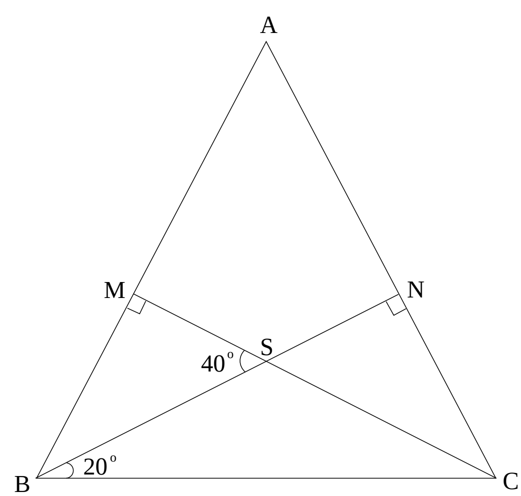
| Country | Score | Percent of students in advanced courses |
|---|---|---|
| USA | 424 | 14 |
| Austria | 462 | 33 |
| Slovenia | 476 | 75 |
| Italy | 480 | 14 |

*Acknowledgements: Suzanne Wilson found the California exams which Lee Shulman used in his 1986 article. I asked an educational historian, David Labaree, where one might find such exams. He suggested the Cubberley Education Library at Stanford. Barbara Celone and Yune Lee of this library found the California exams and some exams from Michigan and sent me copies. I am grateful to all of these people for helping to bring to light some very interesting documents on education in the United States in the 19th century.

As seen in Table 1, the percent of students taking advanced mathematics in the last year of secondary school could vary widely between countries. In general, the better students take advanced mathematics in their final year of school, so a higher percent of students taking advanced math would tend to lower the average score. Thus the gap between 424 and 462, large as it is, underestimates how much U.S. students are below the students in other countries which did not do very well. This gap is unacceptably large. Neal Lane, then director of the National Science Foundation, pointed this out in his comments in *Pursuing Excellence* (U.S. Department of Education, 1998). Unfortunately, this gap, which I consider strong evidence of a very poor geometry program in the United States, was not mentioned in *Principles and Standards for School Mathematics* (National Council of Teachers of Mathematics, 2000).

To illustrate some of the problems in the knowledge of geometry, consider a classic simple problem from synthetic geometry. See (TIMSS, 1995, problem K18).

> In the ΔABC the altitudes BN and CM intersect at point S. The measure of ∠MSB is 40° and the measure of ∠SBC is 20°. Write a PROOF of the following statement:
>
> "ΔABC is isosceles."
>
> Give geometric reasons for statements in your proof.



All that is needed to solve this problem is to know that the sum of the angles of a triangle is 180 degrees; and write that angle BSC is 140 degrees, so angle SCB is 20 degrees. Angle CSN is 40 degrees, so angles MBS and NCS are each 50 degrees. Thus the base angles are equal, so the triangle is isosceles. This is an easy example of a type of problem which is given in grade 8 in Hong Kong, Japan and Singapore. See, for example, problem 38 on page 9.51 in (Chan (2005)), problem 3 on page 121 in (Kodaira (1992)), and problems 1 and 3 on page 323 in (Lee (2002)). The first two are eighth grade books, the third is seventh grade. None of these East Asian countries participated in the 1995 TIMSS assessment of advanced students.

The international average of correct responses on this problem was not high, only 35%. The range was from 10% to 65% with the U.S. score the lowest at 10% correct. When partially correct answers were also considered, the U.S. percentage went up to 19% and the international average to 50%.

In the 2003 Trends in International Mathematics and Science Study (2003 TIMSS), it is interesting to compare the scores between the U.S. and Singapore. The absolute difference has been mentioned by many people, but I want to point out a relative difference. In grade 4 there are five categories. The results for both of these countries was similar in four of these categories, but in both cases there was an outlier; for the U.S. in Data and for Singapore in Number (see Table 2).

Table 2. 2003 TIMSS grade 4 results for the U.S. and Singapore

| Subject | U.S. | Singapore |
|---|---|---|
| Number | 516 | 612 |
| Patterns and relationships | 524 | 579 |
| Measurement | 500 | 566 |
| Geometry | 518 | 570 |
| Data | 549 | 575 |
| Overall average | 518 | 584 |

The Singapore program has a strong focus on arithmetic and multi-step word problems.

In *Beyond Arithmetic* the following claim is made:

> In the *Investigations* curriculum, standard algorithms are not taught because they interfere with a child's growing number sense

> and fluency with the number system (Mokros, Russell, & Economopoulos, 1995, p. 74).

The standard algorithms are taught in Singapore, and this is supplemented by other methods which work in specific instances. The standard algorithms do not get in the way of students learning mathematics. When taught well, they help develop knowledge of place value, the use of the distributive law when doing multiplication, and estimation when doing division. A focus on work with numbers and good word problems provides an excellent foundation for algebra. In grade 8 2003 TIMSS, (Mullis, et al, 2004) Singapore's average score went up to 605 while the U.S. average score went down to 504. The data in Table 2 and grade 8 TIMSS data showing a large drop in the U.S. geometry score suggests to me that we should rethink our priorities.

## Examinations for teachers

In his 1986 American Education Research Association Presidential Address, Lee Shulman started with some comments on content which he called "the forgotten part of education." To illustrate the importance of content long ago, he mentioned a test given to prospective grammar school teachers in California in 1875, (California, 1875). (Grammar school was grades 1 to 8.) The exam covered a wide range of skills and knowledge teachers need. It had 20 sections worth a total of 1,000 points. Three parts were on mathematics, written arithmetic (100 points), mental arithmetic (50 points), and algebra (50 points). Here are some questions from each of the mathematics parts. These questions are among the more complicated ones, but are representative of a substantial fraction of the questions on these exams.

> *Written arithmetic*
>
> Prove that in dividing by a fraction we invert the divisor and then proceed as in multiplication.
>
> If 18 men consume 34 barrels of potatoes in 135 days, how long will it take 45 men to consume 102 barrels? Work by both analysis and proportion. [The implied assumption is made that each man will eat the same amount each day.]

> *Mental arithmetic*
>
> A man bought a hat, a coat, and a vest for $40. The hat cost $6; the hat and coat cost 9 times as much as the vest. What was the cost of each?
>
> Divide 88 into two such parts that shall be to each other as 2/3 is to 4/5.
>
> *Algebra*
>
> Three stone masons, A, B, C, are to build a wall. A and B, jointly, can build the wall in 12 days; B and C can build it in 20 days, and A and C in 15 days. How many days would each require to build the wall, and in what time will they finish it if all three work together? [There is an implied assumption that A will work at the same rate whether working alone or with others, and the same for B and C.]
>
> Divide the number 27 into two such parts, that their product shall be to the sum of their squares as 20 to 41.
>
> A May-pole is 56 feet high. At what distance above the ground must it be broken in order that the upper part, clinging to the stump, may touch the ground 12 feet from the foot?

There were 10 questions in each of these parts. The time allowed was not stated, but at least two hours should be alloted to written arithmetic and to written grammar, two of three topics which were worth 100 points. "Applicants who fail to reach fifty credits [points] in written arithmetic, or in grammar, or in spelling" fail. "Applicants reaching less than sixty credits in any one of these must receive only a third grade certificate, no matter what average percentage they may make."

One comment on the written arithmetic part is worth quoting completely.

> Heretofore applicants have very seldom paid any attention to the special requirements of a paper. For instance, in written arithmetic, applicants are instructed that "no credits are to be allowed unless the answer is correct and the work is given in full, and, when possible, such explanations as would be required of a teacher in instructing a class; a rule is not an explanation." This

instruction has generally been disregarded. In future this will be a ground for the rejection of the paper. (California, 1875, p. 214)

Contrast these questions and remarks with some exams given now for teachers.

The Praxis tests seem to be the tests most commonly used by states as part of certification requirements. Here are two Praxis sample questions, (Educational Testing Service, 2005), for candidates who want to teach middle school mathematics. (Middle school students are generally between ages 11 and 15.) These questions are representative of the posted questions.

You are given corresponding values of x and y.

| $x$ | $y$ |
|---|---|
| $-4$ | $-2$ |
| $-3$ | $-\frac{3}{2}$ |
| $-2$ | $-1$ |
| $-1$ | $-\frac{1}{2}$ |
| $0$ | $0$ |

Which of the following is true about the data in the table above?

A As x decreases, y increases.

B As x increases, y does not change.

C As x increases, y decreases.

D As x increases, y increases.

If there are exactly 5 times as many children as adults at a show, which of the following CANNOT be the number of people at the show?

A 102

B 80

C 36

D 30

(Educational Testing Service, 2005, p. 5)

This question is a much easier one than the analogous question in 1875. (A man bought a hat, a coat, and a vest for \$40. The hat cost \$6; the hat and coat cost 9 times much as the vest. What was the cost of each?). The Praxis tests are designed to help states ensure an adequate level of knowledge of content and professional practice for beginning teachers. It seems that what is considered adequate now would not have been adequate in 1875.

For experienced teachers, the National Board for Professional Teaching Standards (NBPTS) has developed a certification program, which is used to encourage teacher growth and reward those who show adequate knowledge on an examination and an adequate portfolio presentation of their teaching. Here are two questions on algebra for teachers teaching early adolescents. NBPTS defines "early adolescents" as 11- to 15-year-olds, so these are students at the upper end of the ages of those taught by candidates taking California exams in the 1870s. These questions were posted on (NBPTS, 2005) when this paper was written. They and the NBPTS sample high school question below have been replaced by other sample questions on the NBPTS website.

Use the table of coordinates below to respond to the following prompts.

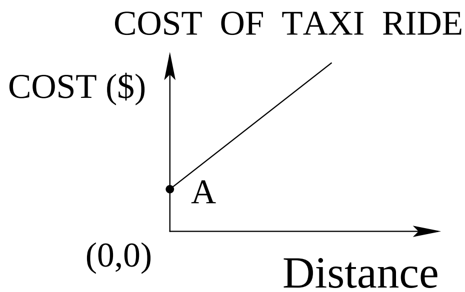
| | | | | | | |
|---|---|---|---|---|---|---|
| x | -3 | -1 | 0 | 1 | 4 | 6 |
| y | 3 | 1 | 0 | 1 | 4 | 6 |

a) Write an equation to model the given data set.

b) State the name of this function and describe the relationship between x and y.

c) Name and describe the specific type of transformation when y is increased by 5.

Use the graph below to respond to the following prompts.




a) Explain the significance of point A within the context of the price of a taxi ride.

b) Given that the slope of the line is 0.5, explain the meaning of this value within the context of the situation modeled above.

Not only are these much easier questions, the second one does not model the cost of a taxi ride, since the cost goes up in discrete increments for certain fractions of a mile, so the graph should be a step function.

To illustrate the NBPTS test for high school mathematics teachers, the following sample question was posted in (NBPTS, 2005).

> Use the following problem to respond to the prompt below. Solve $|x-3|+|x-2|=15$.
>
> A student submits the following response:
>
> $$\begin{aligned} x-3+x-2 &= 15 \\ 2x-5 &= 15 \\ 2x &= 10 \\ x &= 5 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x+3+x+2 &= 15 \\ 2x+5 &= 15 \\ 2x &= 10 \\ x &= 5 \end{aligned}$$
>
> a) Give a detailed explanation of the mathematical understandings and misunderstanding reflected in the student's response.
>
> b) Identify the mathematical knowledge and skills the student needs in order to understand why the response is incorrect. Be specific about what mathematics the student needs to

> learn to enable her to avoid the same error when solving problems requiring similar conceptual and procedural knowledge.

In two of the California exams given in 1873, and 1874, the following questions were in the section "Theory and Practice of Teaching":

> How would you teach a class the reason for inverting the divisor in division of fractions?
>
> (Dec. 1873, p. 224)
>
> How would you require your scholars to explain the process of subtracting 7,694 from 24,302?
>
> (Sept. 1874, p. 238)

Both of these are more important questions in elementary mathematics than the one involving absolute values is for high school mathematics. The California exams from the 1870s do not have high school questions, although some of the algebra questions deal with mathematics which is traditionally taught in high school. Some exams from Michigan (Michigan, 1899) have geometry questions which should be part of a good high school mathematics program. Here are two, one given in August, 1898 and the other in December, 1898. In both cases they are one of ten questions.

> Demonstrate: The volume of a triangular pyramid is equal to one-third the product of its base and altitude.
>
> (1899, p. 8)
>
> Demonstrate: Parallel transverse sections of a pyramidal space are similar polygons whose areas are proportional to the square of the distances from the vertex to the cutting planes.
>
> (1899, p. 17)

Nothing like this degree of knowledge seems to be asked on either the Praxis tests or the National Board of Professional Teaching Standards tests. Although only 50 points out of 1,000 in the California exams were in the section on the theory and practice of teaching, some of the questions in this section dealt with very important parts of teaching. Here are two. The second one was mentioned by Shulman (1986). I had to learn how to improve my oral reading when starting to read to our children. Either the skill mentioned in the question below had disappeared by 1940 or I was a poor student in this area.

> What is your method to teach children to discontinue the sing-song, or monotonous tones which many acquire in reading? Is the method original with yourself?
>
> How do you interest lazy and careless pupils? [Answer in full]

Multiple choice questions were not used 100 years ago. The National Board for Professional Teaching Standards does not use multiple choice items, but the Praxis mathematics content tests used by most states do. A method of asking questions so they can be machine graded yet cannot be solved by eliminating all but one of the answers will be described in the next section.

Shulman (1986) mentioned examining tests from Colorado, Massachusetts, and Nebraska in addition to tests from California and Michigan. Comparing such tests with current tests would be an important service which educators could do. State tests for teachers from the last part of the 19th century and the early part of the 20th century can be found in reports by State Superintendents of Education. State Historical Societies or major university libraries are places to look for these reports. I thank William Reese for mentioning this to me.

## Examinations for students

A new report (American Institutes for Research, 2005) compares and contrasts some aspects of mathematics education in Singapore with education in the United States. The comments on the level of questions asked of grade 6 students in Singapore and what is asked of grade 8 U.S. students in some state tests and of prospective teachers on the Praxis tests should be read by anyone who cares about mathematics education. The conclusions of this report, like the evidence from the test items cited above, make clear that U.S. expectations are set far too low.

Here is a problem mentioned in an article in the *Wisconsin State Journal* (Wilsman, March 20, 1998). The problem was written by a Board member of the Wisconsin Mathematics Council, and 26 other members of WMC helped in writing this article. It was written in response to an article which claimed mathematics education was not effective. The author claimed the following was on the state test. The problem (for eighth grade students) was:

> Eric hopes to sell 1/3 of a dozen paintings that he has finished for the art fair. Which equation should you solve to find the number of paintings that he wants to sell?
>
> A $1/3 = 12n$
>
> B $3n = 12$
>
> C $4n = 12$
>
> D $1/3n = 12n$

Few people would solve an equation to find how many paintings Eric hoped to sell. Just divide 12 by 3 to find the answer. I tried to find out what was on this test. The test developers told me this question was not on the test. It had been submitted, but was not thought to be good enough for their tests. Instead, it was used as a sample problem in (CTB McGraw-Hill, 1997, p. 62). I do not understand why the problem was not discarded rather than used as an illustrative example.

I wrote a response and mentioned a grade 6 Japanese problem which was available on the Web (Japanese-Online.com, 1995). Contrast this problem for grade 6 students in Japan with the one which was suggested as a good problem for grade 8 students in Wisconsin. (Again, an assumption is made about a uniform working rate of people even in different size groups.)

> A job takes 30 days to complete by eight people. How long will the job take when it is done by 20 people?
>
> A 1
>
> B 23
>
> C 21
>
> D 13
>
> E 12

Another source of Japanese problems is a collection of translations of Japanese university entrance exam questions (Wu, 1993). All of the problems are interesting, but I want to just describe how the response format used allows machine grading without being multiple choice. Consider a problem from an 1875 California exam.

> Bought 4 lemons and 7 peaches for 13 cents; 5 oranges and 8 lemons for 44 cents; and 10 peaches and 3 oranges for 20 cents; what was the price of one of each kind?

As a multiple-choice problem, if the price of any one of the three is given the problem becomes trivial. Yes, one could ask what would be the cost of 13 lemons, 10 oranges and 16 peaches as a multiple-choice problem, but here is a much better way to ask for answers. The cost of a peach is $\{A\}/\{B\}$ where next to A on the answer sheet is + - 0 1 2 3 4 5 6 7 8 9 and test-takers are told that fractions should be recorded in lowest form. In this case, $A = 1$ and $B = 2$.

The Japanese use this format on their analogue of the SAT, and ask multi-step questions with individual parts having answers which are recorded. This way they can learn at what step in a multi-step problem many students have trouble. We in the United States need multi-step problems on our exams. We have turned too much of our assessing over to measurement people, and the level of mathematics being assessed shows this. Even the National Research Council, when it did a study of tests for prospective teachers, did not have content people looking at the tests (2001, p. 416).

The picture one gets of education from old exams is consistent with the picture I had formed on the basis of looking at many old textbooks. The common view is that long ago all that students were asked to do is learn a procedure. This is not borne out in the exams given to prospective teachers. Recall that candidates were expected to give reasons for what was being done, both in these tests and in their teaching. In addition, the level of problems which teacher candidates were expected to solve was higher than we now ask of prospective teachers. This higher level of questions also shows up in many old textbooks. We have a lot to learn from the past, including past assessments.

Further information on tests for teachers was given by (Burlbaw, 2002) in a still unpublished paper.

# References

American Institutes for Research. (2005). *What the United States can learn from Singapore's world-class mathematics system (and what Singapore can learn from the United States): An exploratory study.* Washington, DC: Author. Retrieved March 26, 2006 from http://www.air.org/news/documents/Singapore%20Report%20(Bookmark %20version).pdf

Burlbaw, L. (2002). *Teacher test – could you pass the test?* Presented at Mid-America Conference on History, Univ. of Arkansas, Fayetteville.

CTB/McGraw Hill. (1997). *Teacher's guide to Terra Nova*, Monterrey, CA: Author.

California. (1875). *Sixth biennial report of the Superintendent of Public Instruction of the State of California, for the school years 1874 and 1875.* Sacramento, CA: G. H. Springer.

Chan, M. H., Leung, S. W. and Kwok, P. M. (2005). *New trend mathematics (S2B)*, Hong Kong: Chung Tai Educational Press.

Educational Testing Service. (2005). *Middle school mathematics (0069).* Retrieved March 26, 2006 from http://ftp.ets.org/pub/tandl/0069.pdf

Japanese-Online.com. (1995). *World math challenge volume 1.* Retrieved March 26, 2006 from http://www.japanese-online.com After registration, click on lessons.

Kodaira, K. (editor). (1992). *Japanese grade 8 mathematics*, (translation). Chicago, University of Chicago Mathematics Project.

Lee, P. Y. & Fan, L. (consultants); Teh, K. S. & Looi, CK̇. (authors). (2001). *New syllabus mathematics 1*, Singapore: Shing Lee Publishers

Michigan. (1899). *Annual report of the Superintendent of Public Instruction of the State of Michigan, with accompanying documents, for the year 1898*, Lansing, MI: Robert Smith Printing, State Printer.

Mokros, J., Russell, S. J., & Economopoulos, K. (1995). *Beyond arithmetic*, White Plains, NY: Dale Seymour.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *Mathematics achievement in the primary school* *years: IEA's Third International Mathematics and Science Study (TIMSS)*, Chestnut Hill, MA: TIMSS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth Grades.* Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College. Retrieved March 26, 2006 from http://timss.bc.edu/timss2003i/mathD.html

National Board for Professional Teaching Standards. (2005). *Early adolescence Mathematics scoring guide effective 2005.* Retrieved March 26, 2006 from http://www.nbpts.org/pdf/sg/55_EA_MATH.pdf

National Board for Professional Teaching Standards. (2005). *Adolescence and young adult Mathematics scoring guide effective 2005.* http://www.nbpts.org/pdf/59/56_AYA_MATH.pdf

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

National Research Council. (2001). *Testing teacher candidates: The role of licensure tests in improving teacher quality.* Committee on Assessment and Teacher Quality, K. Mitchell, D. Robinson, B. Plake, & K. Knowles (Eds). Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press. Retrived March 26, 2006 from http://darwin.nap.edu/books/0309074207/html/

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(1), 4-14.

TIMSS Mathematics and Science Items for the First Year of Secondary School, http://timss.bc.edu/timss1995i/items.html

U.S. Department of Education. National Center for Education Statistics. (1998). *Pursuing excellence: A study of U.S. twelfth-grade mathematics and science achievement in international context* (NCES 98-049), Washington, DC: U.S. Government Printing Office. Retrieved March 26, 2006 from http://nces.ed.gov/pubs98/twelfth/

Wilsman, M. (March 20, 1998). *Wisconsin State Journal.*

Wu, L. (Ed.). (1993). *Japanese university entrance examination problems in mathematics*, Washington, DC: Mathematical Association of America. Retrieved March 26, 2006 from http://www.maa.org/juee