Assessment and Mathematics Examinations in the CDIO project

Thomas Lingefjärd
Göteborg University and Chalmers University of Technology
Sweden

> A teacher says:
> *It is the examination that controls everything, and it is the examination that directs the unofficial syllabus. … The examination is so important that it is meaningless to do anything in terms of new pedagogical methods if you don't do something about the examination at the same time. The examination represents the entire pedagogy; it is the examination that decides what and how people learn. If I want students to learn more deeply—and I really do—then I have to change the examination; that's the only way that really works.*
> (Högskoleverket [National Agency for Higher Education], 1997, p. 21)

Teaching and Learning

In modern education an increasing emphasis is put on learning through problem-oriented or problem-based educational methods. The underlying idea is to improve the quality of students' learning about complex problems or phenomena in the world through assignments that give rich opportunities for active investigation, analysis, and reflection. Such methods entail an increased use of a wide variety of different information sources. When studying mathematics at the tertiary level, many students use tools like graphing and symbol-manipulating calculators and a variety of sophisticated computer programs like Maple, Mathematica, MatLab, and others. Students also use assorted textbooks and other reference books, and many of them are likely to turn to their family members, friends, colleagues, or maybe neighbors as a reference group.

One could argue that if the students do seek information in a variety of ways, then the way these students study is close to the way people ordinarily work. In many walks of life, people are valued for the everyday jobs or projects they do, their ability to work with others, their responses to problem situations, and their capacity to find tools or information that will help them to complete an assignment. In occupations as well as in modern studies, it is important to be open and flexible in one's approach. It is desirable and would be natural if the examinations in mathematics could mirror that fact.

Educators in many countries have expressed a desire to change the teaching and learning of mathematics. It is both obvious and sad that this change is moving forward slowly:

> Developments in society will force education, including mathematics education, to change in its focus, in its organization, in its use of technology and its content. The focus will change from teaching to learning, its organization will change from rigid class based learning to flexible team based learning, technology will be integrated into the learning process and will support both this new organization of learning and the learning tasks of the individual student.
> (van Weert, 1994, p. 621)

> Over the last decade, we have seen enormous progress in computer-based technologies for mathematics education, relative computational power of machines and software, friendliness of interfaces, and efficiency of connectivity. . . . Despite all this progress and promise, the penetration of these technologies in educational practice proves to be very slow and with great disparity from place to place. (Balacheff & Kaput, 1996, pp. 494-495)

Today the slogan "teaching for understanding" is something many, maybe most, teachers and educators would support, but few would agree on how to put it into action. And how could it be otherwise? The word *understanding* means different things in different contexts. A full explanation would probably depend upon a complete study of all aspects of mathematics education, because to measure or even identify understanding (whatever its meaning) would lead us into the complex area of assessment, and any attempt to teach for, through, or with understanding requires a detailed analysis of both teaching and learning.

Mathematical understanding

To try to describe and discuss knowledge of mathematics or mathematical understanding is indeed a formidable task. "Understanding a mathematical proposition—that is a very vague concept" (Wittgenstein, 1956, p. 5). Yet understanding is used in many ways in mathematics education, often without elaboration. Almost any book about the teaching or learning of mathematics will have its own description or definition of "understanding mathematics." Even if the definition is not there, there will definitely be a statement about the understanding of mathematics.

Since ancient times, people have been concerned about understanding (and lack of understanding) in connection with mathematics. In the *Phaedo* of Plato, Socrates challenges his own understanding:

> I cannot satisfy myself that, when one is added to one, the one to which the addition is made becomes two, or that the two units added together make two by reason of the addition. I cannot understand how when separated from the other, each of them was one and not two, and now, when they are brought together, the mere juxtaposition or meeting of them should be the cause of their becoming two. (Boyer, 1991, p. 83)

Henri Poincaré (1952) underlined the ambiguity of the meaning of the verb:

> What is understanding? Has the word the same meaning for everybody? Does understanding the demonstration of a theorem consist in examining each of the syllogisms of which it is composed in succession, and being convinced that it is correct and conforms to the rules of the game? In the same way, does understanding a definition consist simply in recognizing that the meaning of all the terms employed are already known, and being convinced that it involves no contradiction? (Quoted in Sierpinska, 1994, p. 72)

Sierpinska (1994) explains that researchers in mathematics education have different objectives when discussing the question of understanding mathematics. Some objectives are more pragmatic (to improve understanding), others are more diagnostic (to describe how students understand), and still others are more explicitly theoretical or methodological. What unites researchers is that they all have a theory of what understanding is, explicitly expressed or not. According to Sierpinska, there are at least four different theories or models of understanding in mathematics. To begin with, we have theories that are centered on hierarchies of levels of understanding. One such example is the van Hiele levels (van Hiele, 1986), but there are others.

Second, we have models that describes understanding as a growing ”mental model,” ”conceptual model,” ”cognitive structure,” or something similar. The term *cognitive structure* comes from Piaget, and several authors refer to Piaget when constructing their model for the understanding of mathematics.

Third, Sierpinska mentions models that look at the process of understanding as a dialectical game or interplay between two ways to apprehend understanding. The dialectical dualism may be illustrated by a concept considered as a tool in a problem-solving process and at the same time viewed as an object to study, analyze, and develop in a theoretical way. One well-known example is Skemp's (1978) discrimination between instrumental and relational understanding. According to Skemp, an instrumental understanding is what it takes to reach the right answer, while relational understanding means that you understand both what to do and why. Another way to describe this is as operational versus structural understanding (Sfard, 1994).

The fourth type of understanding is the historical-empirical perspective in which the epistemological obstacles are united by today's students (Sierpinska). Robert and Schwarzenberger (1991) claim that from a psychological perspective, it is meaningful to focus on tertiary students' growing ability to reflect on their own learning of mathematics. They argue that advanced mathematical thinking includes the ability to separate knowledge of mathematics from meta-knowledge of mathematics, which includes, for instance, how correct, relevant, or elegant a solution is. They further advocate that students at this advanced level should have a great amount of mathematical knowledge, experience of mathematical strategies, and well-functioning methods together with aptitude for communicating those skills at least on a basic level. According to Robert and Schwarzenberger (1991), research shows that students vary greatly in this respect.

Examination—Knowledge, Control, and Grading

Several explanations are commonly given when mathematicians discuss control over the examinee, questions of security and cheating, and the difference between teaching or learning and examining. Many mathematicians simply view the examination as a test and not as an important opportunity to learn. Logistics and tradition have kept assessment in mathematics relaying heavily on formal examinations. The discussion in Topic Group 6: Distance Learning (http://mcs.open.ac.uk/icme/) at the Ninth International Congress on Mathematical Education in Makuhari, Japan, in 2000 confirmed that the assessment situation seemed to be totally separate from the teaching situation in many mathematics courses around the world.

It seems that there are significant difficulties in changing the examination of mathematical skills in the undergraduate years and maybe even little educational reason to do so as long as mathematicians are happy with the results of the tests. However, examinations often test a narrow range of skills, and there is a growing effort to find ways to broaden the base of skills that are tested. In addition, textbooks are very often virtual clones of each other, with long sets of repetitive exercises, which unfortunately encourages a surface approach to learning.

> It has become clear from numerous investigations that:
>
> Many students are accomplished at complex routine skills in science, mathematics and humanities, including problem solving algorithms.
>
> Many have appropriated enormous amounts of detailed knowledge, including knowledge-specific terminology.
>
> Many are able to reproduce large quantities of factual information on demand.
>
> Many are able to pass examinations.
>
> But many are *unable* to show that they understand what learned, when asking simple yet searching questions that test their grasp of the content.
>
> In summary, the research seems to indicate that, at least for a short period, students retain vast quantities of information. On the other hand, many seem to forget much of it and do not appear to make good use of what they do remember. (Ramsden, 1992, pp. 30-31)

Clearly this is not a new phenomenon. For many years, perhaps as long as teaching has been practiced, teachers and educators have recognized these facts and have attempted to influence students' learning by a number of methods such as changing the style of their teaching, attempting to give clearer explanations, giving more examples, or preparing better lecture notes, all on the assumption that mathematics need only be presented logically in order to be learned. However, some research has shown that that students are often more motivated to learn material or methods that are of direct relevance to passing, and therefore willing to adapt their learning styles and to do what they perceive is necessary to pass assessment tasks (Ramsden, 1992). This research indicates that changing teaching methods without due attention to assessment methods is not sufficient.

Responsibility in Relation to Learning and Achievement

During their schooling, students inevitably try to identify, interpret, and follow authority. One interpretation of this social behavior is that the search for trustworthy authority is part of the human survival instinct. That instinct does not disappear when students begin their university studies, although the search for authorities or survival structures may be more hidden the older and more sophisticated they get. Figure 1 illustrates the didactical situation.

> A didactical situation is a ”game” in which the teacher negotiates with students' specific situations allowing an interaction with a milieu, which is likely to lead them to the construction of a given piece of knowledge. (Brousseau, 1997, p. 31)

In most study situations, there are naturally other sources of knowledge besides the student's personal knowledge that the student can rely upon.

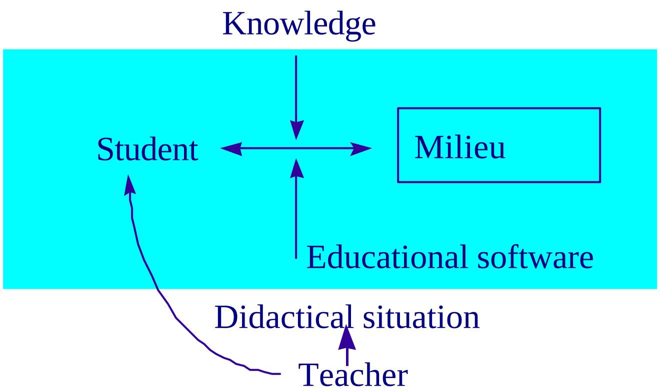


Figure 1: An illustration of the didactical situation. (Adapted from Balacheff, 1993, p. 156.)

Students who work on open-ended problems need feedback from other students or from a teacher if they are to progress and to stretch the limits of the activity and their own mathematical knowledge (Blomhøj, 1993). That observation leads to the question of the didactical contract in the classroom: to the responsibility to learn that all students should have and to the sources of authority the students are likely to identify when working as much with computers as with textbooks.

In the complexity of a situation in which students are mostly away from the teacher when reflecting on and learning mathematics at their own convenience, using calculators and computers from time to time, it is hard to describe all the relations that occur. The discussions in which the students take part nearly always have a third, silent partner: the calculator or computer software and its result. The third partner in the discussion changes the didactical contract between the students and the instructor.

Assessment and Taxonomies

Contrary to past views of learning, the cognitive psychology of today (Marton & Booth, 1997) suggests that learning is not linear but proceeds in many directions at once and at an uneven pace. People of all ages and ability levels constantly use and refine concepts. Furthermore, there is tremendous variety in the modes and speed with which people acquire knowledge, in the attention and memory capabilities they can apply to knowledge acquisition and performance, and in the ways in which they can demonstrate the personal meaning they have created. Current evidence about the nature of learning makes it apparent that instruction that strongly emphasizes structured drill and practice on discrete, factual knowledge does students a major disservice. Acquisition of knowledge skills is not sufficient to make one into a competent thinker or problem solver. People also need to acquire the disposition to use their skills and strategies, as well as the knowledge of when and how to apply them. These are appropriate targets for assessment.

If one adds the component of existing technology, assessment becomes even more complicated. The support to be provided by technology when students are being assessed is a difficult issue and the subject of ongoing discussion in several places around the world. A phrase often mentioned together with the use of technology is *authentic assessment* or *authentic performance assessment,* which, according to Clarke (1996), refers to mathematical tasks that are meaningful for the student, represent applications of mathematics, and include activities that are, in some sense, also carried out by mathematicians. The use of technology such as computer programs and graphing calculators naturally affects the evaluation situation and also what we mean by *assessment* (Webb 1992). An essential consideration is whether students using, say, a computer program when they are learning should therefore be allowed to interact with that program when being assessed in mathematics. We have to find ways of assessing what is looked upon as important, rather than assessing what is easily measurable. In other words, we have to deal with the truism that, in mathematics education, what is assessed is what is valued, and what is valued is what is assessed (Arnold, Shiu, & Ellerton, 1996).

The involvement of the students in the assessment is likely to shape the educational process. Any advice or instruction to a student on how to express the intended outcome will undoubtedly affect the way in which that student and his or her peers present the solution. It is essential to students' learning that they are well informed about the critical points that will be assessed and about the grading system to be used by the instructors. When students become more involved in the process of evaluation, it may be seen as a substantial part of the didactical contract being negotiated between student and teacher. Through this interplay, the students can learn to identify the criteria for qualitatively good performance. Further, they can also learn what is regarded as unsatisfactory, fair, good, or very good performance. It makes sense to give learners opportunities to analyze strong and weak answers to more open-ended problems (Moran, 1997).

If mathematics teachers allow group work, discussion, and information gathering in libraries and over the Internet, and also want students to learn more mathematics in collaborative work, then they face great demands on what types of problems they should pose. Silver and Kilpatrick (1989) argue for the use of open-ended problems in the assessment of mathematical problem solving, thereby moving from facts and procedures to concepts and structures. A relevant problem should encourage students to make various assumptions and use various strategies in which technology can serve as an aid but never as a goal. The problems teachers choose also need to provide the students with opportunities to express what they have learned in the course and in previous courses. At the same time that the problem should remain nontrivial in the presence of technological tools, their use should not be the only performance component that is essential and leads to success (Lingefjärd & Holmquist, 1999).

Important questions for assessment are the following:

- What mathematical content is most appropriately tested in a technology-enriched environment?
- How do problems on tests in a technology-enriched environment differ from those on tests not allowing the use of calculating technology?
- Should calculators and computers use be optional or required during testing?

Assessment Methods

Assessment drives what students learn and to some extend also what teachers lecture about. It controls the students approach to learning by directing them to take either a surface approach or a deep approach to learning (Ramsden, 1992). The types of questions we give show students what we value and how we expect them to direct and use their time of study. The educational system would benefit significantly if we could create and use questions that would help to build concepts, alert students to misconceptions and introduce applications and theoretical ideas. But is it possible to create and classify exam problems according to such objectives?

A strong tradition within mathematics is to classify and order problems in terms of difficulty and student errors. Would it be possible to use a taxonomy to classify a set of tasks ordered by the nature of the activity required to complete such a task successfully instead? Such a classification could have activities which need only a surface approach appearing at one end, and those activities requiring a deeper approach appearing at the other end. There are a variety of taxonomies that one could use, depending on the purpose. One of the best known is Bloom's taxonomy, which gives a hierarchy of concepts (Bloom, 1956).

Benjamin Bloom (1956) and his colleagues managed to organize goals of instruction into a taxonomy with the objectives to reflect the distinctions teachers make, to be logical and internally consistent, to reflect then-current psychology, and to be both neutral and comprehensive. The Bloom taxonomy was also designed to fit all school subjects. In the taxonomy, objectives were separated by "domain" (cognitive, affective, and psychomotor, with the cognitive receiving most attention), related to "educational behaviors," and arranged in ascending order from simple to complex:

a) knowledge
b) comprehension
c) application
d) analysis
e) synthesis
f) evaluation

Bloom and his colleagues quoted research evidence to support the claims for the cognitive domain that "as the behaviors become more complex, the individual is more aware of their existence" (p. 19) and that the individual abilities and skills to be found in the upper levels of the taxonomy were more efficiently learned than the knowledge at the lowest level (p. 42).

It did not take long before contrary opinions about the usefulness of Bloom's taxonomy began to emerge. In the First International Mathematics Study (Husén, 1967, ch. 1), Robert Thorndike claimed that he found striking differences within each country in the pupils' performance between items classified at different cognitive levels (p. 36).

> Research studies that have shown differences between process levels have found such differences only within topics, and complexity of process has then been equated with difficulty of test item. (Hill, 1984, p. 227; See also Seddon, 1978)

In his article *The Chain and the Arrow: From the History of Mathematics Assessment*, Jeremy Kilpatrick gives an excellent description of how Bloom's taxonomy has been used and misused over the last century:

> The Bloom taxonomy has often been seen as fitting mathematics especially poorly. As Ormell (1974) noted in a strong critique of the taxonomy, Bloom's categories of behavior
>
> > are extremely amorphous in relation to mathematics. They cut across the natural grain of the subject, and to try to implement them – at least at the level of the upper school – is a continuous exercise in arbitrary choice. (Ormell 1974, p. 7) (Kilpatrick 1993, p. 36)
>
> Additional criticisms have questioned the validity of the distinction between cognitive and affective objectives, the independence of content from process, and the meaning of objectives isolated from any content (Freudenthal 1975; Kilpatrick, 1979). Nonetheless, the view of mental abilities and consequently of mathematical thinking and achievement

> as organized in a linear, hierarchical way has been powerful in 20th century assessment practice. It has deep roots in our history and our psyches (Romberg, et. al., 1990, White, 1977). (Kilpatrick, 1993, p. 36)

In summary, it seems that Bloom's taxonomy is good for structuring assessment tasks but has severe limitations when used for mathematics. A group of mathematicians and mathematics educators at the University of Technology in Sidney have therefore constructed what they call a MATH taxonomy (mathematical assessment task hierarchy) for the structuring of assessment tasks (Smith, Wood, Coupland, Stephenson, Crawford, & Ball, 1996).

The MATH taxonomy uses eight different descriptors, gathered in three different groups (see Figure 2).

Group A	**Group B**	**Group C**
Factual knowledge	Information transfer	Justifying and interpreting
Comprehension	Application in new situations	Implications, conjectures and comparisons
Routine use of procedures		Evaluation

Figure 2: MATH taxonomy. (From Smith et al. 1996, p. 67.)

It is expected that students enter their tertiary institutions with most if not all of their mathematical experience with group A tasks and with just some experience with group B tasks. Consequently, their experience with group C tasks is severely limited or even nonexistent. Nevertheless, one of our aims at the university or tertiary level of mathematics education should be to develop skills in all three categories (groups).

Smith et al. (1996) recommend the use of a grid (Figure 3) that combines subject topics with the descriptors of the MATH taxonomy. The grid entries represents a reference to particular questions on the examination paper. This enables the examiner to more readily determine the balance of assessment tasks on the paper. In their article, Smith et al. claim that most of the mathematics examination papers they analyzed were heavily biased towards group A tasks.

Topic / **MATH Taxonomy**	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Factual knowledge					
Comprehension					
Routine use of procedures					
Information transfer					
Applications in new situations					
Justifying and interpreting					
Implications, conjectures, comparisons					
Evaluation					

Figure 3: Grid for MATH taxonomy and subject topics. (From Smith et al. 1996, p. 67.)

It is very important to realize how hard it sometimes is to decide what skills a particular problem assesses. If a student is asked to prove a theorem that was presented in a lecture, she or he might very well present a correct answer based upon a true understanding of the theorem. The student could have a full understanding of the theorem and its significance, and maybe also be able to apply the theorem in relevant situations or prove similar theorems. On the other hand, it could also happen that the student only knows how to reproduce the theorem in a given style. This style of assessment cannot discriminate between different types of learning that can lead to the same response. Naturally, if we are comfortable with this, then the assessment pattern is satisfactory. But if we wish to be sure that the student understands the theorem and has not merely learned it by rote, then we should ask more probing questions. It is indispensable to be clear about the desired outcomes of the assessment we construct and to be able to identify the types of assessment tasks that are reliable indicators of these outcomes.

The list of descriptors in Figure 3 seems to force all different types of mathematical assessment into one of eight categories, which could be seen as a rather stereotypic, quantitative way to classify qualitative differences. There will definitively be borderline cases or maybe cases that do not fit comfortably in any category or that fit into more than one category. Smith et al. (1996) conclude the following:

> It is not our aim to be able to uniquely characterize every conceivable assessment task. Rather, the aim of the descriptors is to assist with writing examination questions, and to allow the examiner's judgment, objectives and experience to determine the final evaluation of an assessment task. (p. 68)

The next step in the taxonomy process is to give a list of examples to illustrate the descriptors further. What is implicit in any such list of categorized problems for the tertiary level is the assumption that the student has prior knowledge in many areas of mathematics from his or her former studies. Even more important is the fact that when a students succeeds in proving a theorem for the first time, she or he is also demonstrating an ability to apply knowledge to new situations (Group B) but may only be demonstrating a factual recall (group A) when proving the same theorem for the second time.

The variety and complexity of a tertiary student's mathematical knowledge are illustrated in Figure 4. In this very schematic and simple 3-D knowledge profile of a student's knowledge development in mathematics, we can see the different areas of mathematics such as linear algebra, real analysis, discrete mathematics, et cetera, beginning from the left at the bottom of the figure. We can also picture different taxonomy descriptors such as Factual Knowledge, Information Transfer, and Justifying and Interpreting, just to mention some of the eight descriptors in the MATH taxonomy. The height of the bars illustrate different levels of quality in the students' knowledge such as poor, good, and excellent or whatever scale is used. The ascending bars illustrate the fact that the students accumulate mathematical experience, procedures, routines, and skills from one course to another. It is obvious that the term *Quality of Knowledge* must be seen as constructed, in a very complex way, by different sorts of skills, procedures, and conceptual understanding – all represented by different types of knowledge at different levels.

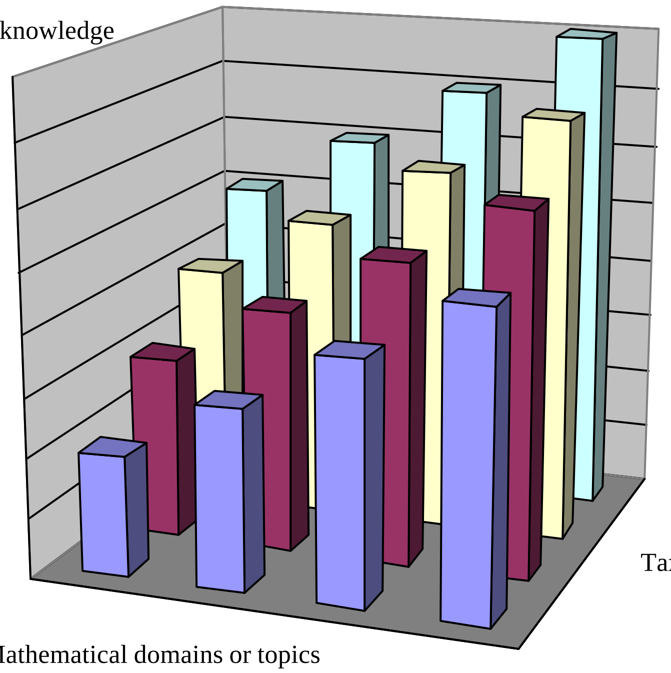


Figure 4: The variety and complexity of a tertiary student's mathematical knowledge.

It is very easy to believe that we must teach students procedures and factual knowledge first in order to be able to teach and assess deeper understanding at the other end of the taxonomy scale. By accepting this hierarchical view of how knowledge in mathematics (or any other subject) is built, we also encourage ourselves and our students to trust in and use unfortunate education practice:

> The idea that knowledge must be acquired first and its application to reasoning and problem solving can be delayed is a persistent one in educational thinking. "Hierarchies" of educational objectives, although intended to promote attention to higher order skills, paradoxically feed this belief by suggesting that knowledge acquisition is a first step in a sequence of educational goals. The relative ease of assessing people's knowledge, as opposed to their thought processes, further feeds this tendency in educational practice. (Resnick, 1987, pp. 48-49)

Constructing Assessment Tasks

Let us begin with two examples of the type of question we can use to test mathematical knowledge including communication skills. The first question requires a relatively short answer, while the second question is more extended.

1 Two students Carla and Johan are discussing whether or not the matrix equation **MAX** = **MB** representing a linear system has the same solutions as **AX** = **B**. Carla thinks they have the same solutions: she says that if you start with AX = B and multiply by M, then you do not change the solutions. Johan, knowing that matrix equations do not always behave like ordinary equations, is more suspicious.

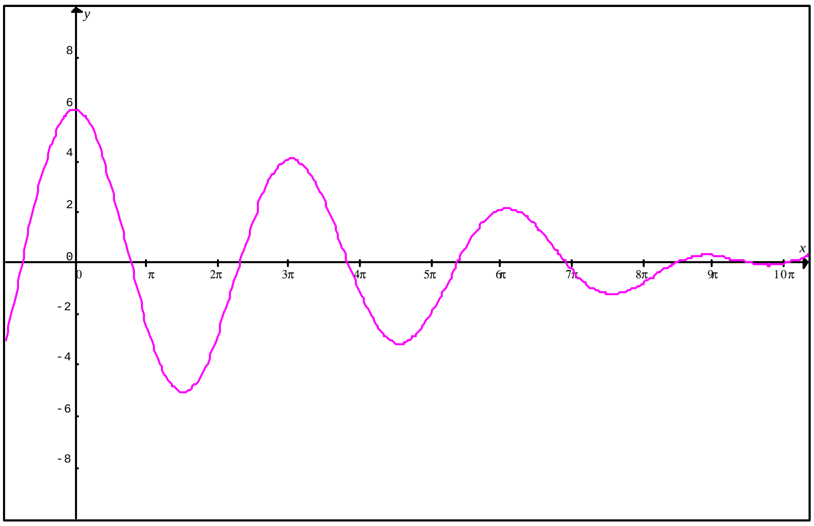
What do you think?
Write a clear justification of your position and illustrate your argument with unambiguous examples.

The next question tests a variety of proficiencies like the transfer of information from graphical to numerical representation, factual recall, the ability to select for relevance, and the ability to communicate ideas.

2 The graph of a sine curve of the type

$$y = (a - b \cdot x)\sin(c \cdot x + \frac{p}{d})$$

is plotted in the figure below for particular values of the constants *a*, *b*, *c*, and *d*.



i) What are reasonable values for the parameters *a*, *b*, *c*, and *d*?
ii) Describe a physical situation to which the sine curve could apply.
iii) What can be said about the roots to the equation $y = 0$?
iv) What can be said about the domain for the function *y*?

How would you characterize these two questions, according to the MATH taxonomy? Clearly either Question 1 or Question 2 are more than purely factual although they do require factual knowledge from the students in order to fully answer or even start to discuss them. Let me give some further examples of mathematical questions in different categories of the MATH taxonomy.

Factual knowledge

Example 1 What is the formula for the area of a circle?
Example 2 State Cramer's rule for solving a system of equations.

Comprehension of factual knowledge

Example Answer true or false. Let ***A*** and ***B*** be 2x2 matrices.

[ ] Determinant (***A*** + ***B***) = determinant (***A***) + determinant (***B***)
[ ] Determinant (***A*** · ***B***) = determinant (***A***) · determinant (***B***)
[ ] Determinant (***A*** · ***B***) = determinant (***B*** · ***A***)

A student might very well be able to correctly quote and use the relevant definitions but nevertheless unable to answer these questions correctly. This indicates that the ability to quote a definition may often be a meaningless skill.

Routine use of procedures

Here we must assume that the students have done drill and practice in task similar to the ones assessed.

Statement An n x n matrix $\boldsymbol{A}$ is singular if and only if the determinant $(\boldsymbol{A}) \neq 0$.

Evaluate $A = \begin{vmatrix} 2 & 1 & 3 \\ 4 & 2 & 1 \\ 6 & -3 & 4 \end{vmatrix}$

Information transfer

Example Here is an attempted proof of L'Hôpital's rule:

Statement

$$\textit{If } f(a) = g(a) = 0 \textit{ then } \lim_{x \to a} \frac{f(x) - g(x)}{g(x) - g(a)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Proof

a) $$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f(x) - f(a)}{g(x) - g(a)}$$

b) $$= \lim_{x \to a} \frac{(f(x) - g(a))/(x-a)}{(g(x) - g(a))/(x-a)}$$

c) $$= \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

Explain carefully what is happening in each of the two steps, labeled a), b) and c). Explain where there could be difficulties with the proof. What conditions should be added to the statement in order to make the proof valid?

Applications to new situations.

Our assumption here must be that the students have not met any of the results that they are asked to prove or situations they are asked to apply the results to.

Example

In a certain town 30 percent of the married women get divorced each year and 20 percent of the single women get married each year. There are 8000 married women and 2000 single women. Assuming that the total population of women remains constant, how many married women and how many single women will there be after 1 year? After 2 years? After n years?

Justifying and interpreting

Example

Here are two arguments, one to show that $\sin^{-1} x - \cos^{-1} x = \pi/2$ *and the other to show that* $\sin^{-1} x + \cos^{-1} x = \pi/2$. *They cannot both be correct (and may both be wrong). Find and explain the error(s) in the reasoning.*

We know that

$$\cos y = \sin (y + \pi/2)$$

for all y, so suppose

$$x = \cos y = \sin (y + \pi/2)$$

Then

$$y = \cos^{-1} x \qquad (1)$$

and

$$y + \pi/2 = \sin^{-1} x \qquad (2)$$

Subtraction of equation (1) from equation (2) gives the result

$$\sin^{-1} x - \cos^{-1} x = \pi/2$$

On the other hand, we also know that

$$\cos (\pi/2 - y) = \sin y$$

for all y, so suppose

$$x = \cos (\pi/2 - y) = \sin y$$

Then

$$y = \sin^{-1} x \qquad (3)$$

and

$$\pi/2 - y = \cos^{-1} x \qquad (4)$$

Addition of equations (3) and (4) gives the result

$$\sin^{-1} x + \cos^{-1} x = \pi/2$$

Implications, conjectures, and comparisons

The following problem suggest that students use a computer program or a sophisticated calculator to multiply given matrices and then make conjectures based on the results they obtained.

Example *This problem investigates the similarity of nth powers of matrices. You are given two square matrices* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *and a nonsingular matrix* $\boldsymbol{P}$ *that satisfies the relationship* $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$.

a) *Check that* $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$ *as claimed.*

b) *Calculate* $\boldsymbol{B}^2$ *and* $\boldsymbol{P}^{-1}\boldsymbol{A}^2\boldsymbol{P}$.

c) *Calculate* $\boldsymbol{B}^3$ *and* $\boldsymbol{P}^{-1}\boldsymbol{A}^3\boldsymbol{P}$.

d) *Calculate* $\boldsymbol{B}^4$ *and* $\boldsymbol{P}^{-1}\boldsymbol{A}^4\boldsymbol{P}$.

e) *Let* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *be any two similar matrices. Make a conjecture about the similarity of* $\boldsymbol{C}^n$ *and* $\boldsymbol{D}^n$, *for* $n = 1, 2, 3, \ldots$

f) *Prove your conjecture.*

Evaluation

Example 1

If S = {$\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_n$} *is a set of linearly independent vectors in an n-dimensional space, then S is a basis of V. Is it possible to prove this statement? Give reasons for your answer.*

Example 2

Here are two definitions of a complex number:

1 *The equation* $x^2 = -1$ *has no real roots, but we may invent an imaginary unit* ***i*** *for which* $i^2 = -1$. *We may then define a complex number as a combination p + iq formed from the two real numbers p and q and the imaginary unit i.*

2 *The complex numbers can be defined as the set* $\mathbf{C} = \{(x, y): x, y \in \mathbf{R}\}$ *together with certain standard arithmetical operations defined on this set.*

Compare the two definitions. Your answer should include:

The circumstances under which each definition would be appropriate.
The relative merits of each definition from a mathematical point of view.
Historical aspects of these definitions.
A demonstration of the equivalence of the definitions.

References

Arnold, A., Shiu, C. & Ellerton, N. (1996) Critical issues in the distance teaching of mathematics and mathematics education. In A. J. Bishop K. Clements, C, Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 701-754). Dordrecht, The Netherlands: Kluwer.

Balacheff, N. (1993). Artificial intelligence and real teaching. In C. Keitel & K. Ruthven (Eds.), *Learning from computers: Mathematics education and technology* (pp. 131-158). Berlin: Springer-Verlag.

Balacheff, N., & Kaput, J. (1996). Computer-based learning environments in mathematics. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 469-505). Dordrecht, The Netherlands: Kluwer.

Blomhøj, M. (1993). Modelling of dynamical systems at O-level. In J. de Lange, C. Keitel, I. Huntley, & M. Niss (Eds.) *Innovation in maths education by modelling and applications* (pp. 257–268). Chichester: Ellis Horwood.

Bloom, B.S. (Ed.) (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: McKay.

Boyer, C. B. (1991). *A history of mathematics* (2nd ed.). New York: Wiley.

Brousseau, G. (1997). *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970-1990* (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. & Trans.). Dordrecht, The Netherlands: Kluwer.

Clarke, D. (1996). Assessment. In A. J. Bishop K. Clements, C, Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 327-370). Dordrecht, The Netherlands: Kluwer.

Hill, P. W. (1984). Testing hierarchy in educational taxonomies: A theoretical and empirical investigation, *Evaluation in Education*, 8, 179-278.

Husén, T. (1967). *International study of achievement in mathematics* (Vol. 2). New York: Wiley

Högskoleverket [National Agency for Higher Education]. (1997). *Uppfattningar om examination-en intervjustudie av högskolelärare* [University teachers' beliefs about examinations] (Report No. 1997: 3 S). Stockholm, Sweden: Author.

Kilpatrick, J. (1979). Methods and results of evaluation with respect to mathematics education. In B. Christiansen & H. G. Steiner (Eds.), *New trends in mathematics teaching* (Vol 4. pp. 162-179). Paris: UNESCO.

Kilpatrick, J. (1993). The chain and the arrow: From the history of mathematics assessment. In M. Niss (Ed.) *Investigations into assessment in mathematics education* (An ICME Study, pp. 31-46). Dordrecht, The Netherlands: Kluwer.

Lingefjärd, T., & Holmquist, M. (in press). Mathematical modeling and technology in teacher education: Vision and reality. In W. Blum, K. Houston, S. J. Lamon & J. F. Matos (Eds.), *Modelling, applications and mathematics education*.

Moran, J. J. (1997) *Assessing adult learning: A guide for practitioners*. Malabar, FL: Krieger Publishing.

Marton, F., & Booth, S. (1997). *Learning and awareness*. Mahwah, NJ: Erlbaum.

Piaget, J. (1978). *Success and understanding*, London: Routledge and Kegan Paul.

Poincaré, H. (1970). *Science and method*, New York, Dover.

Ramsden, P. (1992). *Learning to teach in higher education*. London: Routledge.

Seddon, G. M. (1978). The properties of Bloom's taxonomies of educational objectivities for the cognitive domain, *Review of Educational Research*, 48, 303-323.

Sfard, A. (1994). Reification as the birth of a metaphor. *For the Learning of Mathematics, 14*, 44-45.

Sierpinska. A. (1994). *Understanding in mathematics*. London: Falmer Press.

Skemp, R. (1978). Relational and instrumental understanding. *Arithmetic Teacher, 2*6, 9-15.

Silver, E. A., & Kilpatrick, J. (1989). Testing mathematical problem solving. In R. Charles & E. Silver (Eds.), *Teaching and assessing mathematical problem solving* (pp. 178-186). Hillsdale, NJ: Erlbaum.

Smith, G. H., Wood, L. N., Coupland, M., Stephenson, B., Crawford, K., & Ball, G. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal for Mathematical Education in Science and Technology*, *27*(1), 65-77.

Romberg, T. A., Zarinnia, E. A., & Collis, K. F. (1990). A new world view of assessment in mathematics. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 21-38). Washington, DC: American Association for the Advancement of Science.

Van Hiele, P. M. (1986). *Structure and insight*. Orlando: Academic Press.

Van Weert, T. (1994). Education and computers: Who is in control? In K. Brunnstein & E. Raubold (Eds.), *Thirteenth World Computer Congress 94* (Vol. 2, pp. 619-626). Amsterdam: Elsevier Science B.V.

Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 661-683). New York: Macmillan.

Wittgenstein, L.(1956). *Remarks on the foundations of mathematics*. Oxford: Basil Blackwell.