A Model for Constructing Classroom Assessments
Author(s): Truus Dekker
Published by: National Council of Teachers of Mathematics
Stable URL: http://www.jstor.org/stable/20876031
Accessed: 24/05/2013 13:12

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of
content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms
of scholarship. For more information about JSTOR, please contact support@jstor.org.
Just like their American colleagues, teachers in the Netherlands design their own classroom assessments. However, they also design the tasks for the school exam. The final examination at the end of a student’s school career consists of a school exam, designed by the teacher, and a national exam, designed by the Dutch Institute for Educational Measurement, CITO.

During the mathematics teachers’ team meetings, vehement discussions about the content of these tests often take place. Teachers know that students and their parents will look very carefully at the tasks and at the scores and grades that are given, since much depends on the results. Low scores on classroom tests during the school year could mean a student has to take the course a second time and thus, in the final year, might fail his or her school exam. Usually, colleagues divide the workload of designing classroom assessments for their students; at times, they use some of the assessments provided by the textbook publishers. Often, tasks taken from the national examinations are added to classroom tests as well. Because the national examinations are published each year right after they have taken place, the tasks are freely available for use. Passing the final exam means a student gains access to higher levels of education. For teachers, the national exams form an important guide toward the necessary level of understanding required of their students.

A classroom assessment needs to represent the teaching that preceded it. On the other hand, tests influence the process of teaching and learning. For
example, if students need to be able to reason mathematically when solving a problem, they must be taught how to do that before they take a test. Teachers should discuss students’ work with them after an assignment so they can learn from mistakes and reconsider misconceptions they may have. In team meetings, it is important to discuss the design principles that are used to compose a balanced, time-restricted test.

**A BALANCED TIME-RESTRICTED WRITTEN TEST**

A time-restricted test limits both the number and the type of questions one can ask. Student investigation, for example, does not fit into this format. Students need to show they have mastered the basic knowledge they were taught prior to the test, but the teacher will want to determine whether students gained insight and are able to apply basic knowledge in new and unfamiliar situations. Thus, short-answer questions, the type that might appear on a short quiz, are not sufficient for a balanced test.

In this article, I will restrict myself to a discussion about tests a teacher would typically give to students after they finished a unit or chapter of the textbook. Usually, such a written test takes one or two class periods. A balanced test, in my view, assesses not only whether students can reproduce the definitions or procedures they learned but also whether they are able to choose their own mathematical tools, to generalize, and to reason deductively. A balanced test should be part of a coherent assessment plan, made for the whole school year, which enables the teacher to make sure different test formats are used during the year.

When designing a balanced test, my colleagues and I use the design principles shown in the pyramid model discussed below. The model was developed in 1995 by researchers at the Freudenthal Institute when designing the Dutch National Option for the Third (now Trends) in International Mathematics and Science Study, TIMSS. This model, in an adapted form, was later used for the international Programme of International Student Assessment study (PISA) of the Organization for Economic Co-operation and Development (OECD).

**DESIGN PRINCIPLES FOR A BALANCED TEST**

A balanced test reflects assessment for conceptual understanding as well as factual and procedural knowledge. Students need to demonstrate their understanding of mathematics and to show mastery of a variety of competencies when solving mathematics problems. A few of these competencies are—

- thinking and reasoning mathematically, using arguments and finding proof;
- using formal symbolic and technical language; and
- using mathematical models. A realistic problem is translated into a mathematical problem by using a model that has oftentimes been simplified.

**A PYRAMID MODEL**

For the design of a balanced test, competencies are summarized in a pyramid consisting of three levels. Note, however, that these levels do not represent a hierarchy in competencies but rather show the proportion of questions or amount of time to be spent on tasks at each of the levels. Being “proficient” in mathematics means being able to solve mathematics problems at all three levels of the pyramid, depending on the demands of the grade level. See **Figure 1**.

The level of a problem or question depends mainly on the teaching that preceded it; in the end, it is only the teacher who can decide upon the level of the task. A type of problem that requires integrating mathematical tools and was practiced often in class would be at level I for one group of students; if not practiced, it might be level II or even level III. An overview of the levels in the pyramid will be given, and some examples follow.

**Level I: Reproduction, use of standard procedures and algorithms, knowledge of facts and definitions**

Level I questions assess basic knowledge and procedures that may have been emphasized and practiced during instruction. As shown in the pyramid model, a level I question is not necessarily easy; difficult questions appear across competency levels. A difficult level I question often requires more steps to find the answer, with each of the steps being a procedure or algorithm. Since students can make more mistakes if more steps are needed, the task is a more complicated one.

Questions on this level are mostly easy to design; questions from the textbook can be used as examples.

**Level II: Making connections, integration**

At this level, students are expected to combine information from different sources and from different
mathematical tools and strategies and explain how they solve a problem. These tasks are often posed within a context, and students need to “mathematize” the situation before they can solve the problem. Note that the context can be mathematics itself as well, as the following example shows. Often, students have seen the context or a similar one used during their lessons, but the mathematical content is new and unfamiliar. For example, students are familiar with the context of square root functions like \( f(x) = \sqrt{2x - 4} \). A problem with a new and unfamiliar content within this mathematical context could be something like this:

The function \( f(x) = \sqrt{2x - 4} \) is part of a “family of functions” \( y = \sqrt{px - 4p + 4} \). For \( p = 2 \) the given function \( f \) appears.

Prove that all graphs represented by 
\[ y = \sqrt{px - 4p + 4} \]
have one point \( G \) in common.

Students need to find their own mathematical tools to solve the problem. No known strategy is available. One strategy could be to find the intersection point of two graphs, choosing two “easy” values for \( p \). Next, students should show that this point is on all graphs represented by \( y = \sqrt{px - 4p + 4} \).

**Level III: Mathematizing, reasoning, generalizing, and insight**

On this level, students should see that many contexts contain the same mathematical content. For some students, this ability to generalize is a challenge; they look at each new (context) problem as entirely new, instead of saying, “Oh, this is like the taxicab problem we solved earlier.” Problems on this level require students to make and criticize mathematical models and adapt them when needed. Students should write mathematical arguments to support their views and prove that statements are always right or never right.

The pyramid model serves two purposes. It allows teachers to classify mathematical tasks at different competency levels, and it provides a rough indication of the distribution of the number of questions at the different levels that should be present in a balanced test. Since tasks at level I, reproduction, are often single-answer questions that take relatively little time to answer, the ratio sometimes used is


**EXAMPLES AT DIFFERENT LEVELS FROM THE PYRAMID**

There is no sharp distinction between levels in the pyramid; a task can be placed lower or higher within the level according to the teacher’s experience. Some examples will be shown below, collected from different sources. Most of the larger examples were taken from the national examinations designed by CITO for the Dutch HAVO school level. HAVO stands for Hoger Algemeen Voortgezet Onderwijs, or Higher General Secondary Education. HAVO is one of the three types of general secondary education in the Netherlands. Students passing the HAVO exam get access to higher-level vocational education. They may want to become middle school or high school teachers, for example. The other types of secondary education are VMBO (Voorbereidend Middelbaar Beroeps Onderwijs), which prepares students for an intermediate level of vocational education, and VWO (Voorbereidend Wetenschappelijk Onderwijs), or University Preparatory Education.

**Level I Examples**

Here are some typical level I examples:

- What is the name of the graph represented by this equation: \( y = -x^2 + 2x + 5 \)?
- Solve for \( x \): \(-x^2 + 2x + 5 = 0\)
- Use the Pythagorean theorem to find the distance between \( A(-2, 5) \) and \( B(3, 18) \).

The tasks on this level are often posed as questions with only one correct answer. A multiple-choice format can sometimes be used, although limited information about student abilities is provided by this type of format. The format allows for guessing, so it will not be clear whether a student really reasoned in an appropriate way. For example, Benno gave the right answer “A” to this multiple-choice task:

The tangent of the angle between the positive \( x \)-axis and the straight line represented by the equation \( 5x - 8y - 3 = 0 \) is:

\[
A. \frac{5}{8} \quad B. \frac{-5}{8} \quad C. \frac{8}{5} \quad D. \frac{-8}{5}
\]

When asked for the reason behind his choice, Benno explained, “I chose 5/8 because height is always shorter than distance, and the answer should be positive, since the introduction said the positive \( x \)-axis.”

**Level II Examples**

The usual format for level II is an open- or extended-answer question posed within a context.

For example, as a part of a larger task, students were asked to “solve the equation \((x + 6)^2 = -81\), if possible.” Responses may provide information the teacher can use. An expected answer might be, “A quadratic number cannot be negative, so there is no value for \( x \) that would make the statement true.” This student makes a connection to other
knowledge he or she has about the number system. A student who expands \((x + 6)^2\) first and uses the determinant or even the standard algorithm may still find the right answer but demonstrates less insight; the same student might also complain afterward that there was not enough time to finish the test, part of a penalty for not having higher-level knowledge.

**Large-Context Examples**

The larger-context examples that follow (taken from the national mathematics-B exam for Dutch HAVO schools) will enable the reader to compare tasks at the same grade level but at different levels in the pyramid model. Often, when large-context tasks are used in a test, we start with one or two level I questions to engage students in the context. Then a level II question is followed by a level III question. The example below only shows the level II question. A level I question that could be added might ask students to find the score \(S\) for a person who weighs 95 kilograms and covered a distance of 34 meters.

Athletes sometimes take part in several disciplines like the 100-meter run, long jump, or shot putting. In order to compare the results in these disciplines, formulas are used to convert a result in meters, or seconds, into a score.

When shot putting, if you are not a professional, your own weight will make a difference. An example of a formula that can be used to convert a distance \(D\) (in m) to a score \(S\) for a person who weighs \(W\) kilograms (kg) is

\[
S = D \times \left( \frac{50}{W} \right)^{\frac{2}{3}}.
\]

This formula is rather complicated, so a simplified formula is made for a group of students, all more or less the same weight of 80 kg. On average, they manage a distance of 15 m.

In the simplified formula \(S = D - k(W - 50)\), there is an undetermined factor \(k\).

Find the value(s) for \(k\) that, for this group of students, results in a lower score \(S\) in the simplified formula compared to the first formula.

In order to solve the problem, students need to combine information from the text and from both formulas, which makes it level II. They also need to find their own strategy. It is a difficult problem on this level, since it is a rather complicated one.

The following example shows different levels of competence. Often in an examination, mathematical models are already provided. At level III, however, students should be able to criticize a model and adapt it if necessary. This seldom is part of a written, time-restricted test but will fit in other test formats, such as investigations. In the example, a level I question is missing, but one, such as the following, could be added to start with: Suppose you want to make a new glass panel because the old one is broken. What are the dimensions (diameter and circumference) of the circular piece of glass of which you are taking one-quarter?

The table shown in **figure 2** consists of a right-angled block and a glass panel in the shape of a quarter circle. The glass panel is set up on a metal structure with three legs. One of these legs is fixed in the block. The distance from this leg to the two nearest edges of the block is 20 cm; the distance to the right sides of the glass panel is also 20 cm.

Measurements are shown in the top view in **figure 3**. A mathematical model for this situation, which can be used to solve the problems, is given in **figure 4**. In this model, we neglect the thickness of the metal legs. Points \(P\), \(Q\), and \(R\) are points on the glass panel, right above the three metal legs. The glass panel can be turned around the fixed leg, which is below point \(Q\) in the model.
Fig. 4 Diagram for a model of the problem situation

1. In figure 4, what minimum length should PQ and RQ have to enable the glass panel to turn around 360°? Round your answer up to whole centimeters. (level II)

Suppose you were to place the block against the wall, as shown. The same situation is shown on an activity sheet, scale 1:15. Point Q is the same point as shown in figure 4. There are two points, A and B, on the wall where the glass panel will hit the wall when the panel is turned around the fixed leg.

2. Make an accurate drawing of the placement of points A and B. Explain what you did to determine the location of points A and B. (level III)

3. Find the distance between A and B on the wall. Give your answer in whole centimeters. (level III)

Answers and maximum number of points awarded
1. The vertex of the square that is to the right and behind point Q has the largest distance to Q. (1 point)
   The distance from Q to this vertex is $\sqrt{40^2 + 40^2} = 40\sqrt{2}$. (2 points)
   The minimum length for both PQ and RQ must be 57 cm. (1 point)

2. The drawing should be accurate and include the following:
   The glass panel will touch the wall at each of the endpoints of the quarter circle shown in the drawing. (1 point)
   These endpoints describe a circle with center Q. (3 points)
   A and B are the points of intersect of this circle with the line representing the wall. (2 points)
   Note: If the drawing is correct but any description is missing, 3 points should be subtracted.

3. $QA^2 = 80^2 + 20^2$ (2 points)
   $QA = \sqrt{6400}$ (1 point)
   $6800 - 40^2 = 5200$ (1 point)
   $AB = 2 \times \sqrt{5200}$ (1 point)
   The distance between A and B on the wall is 144 cm. (1 point)

Teachers must be careful with context tasks. Some students might get bored if too many questions about the same situation are asked. Furthermore, if the context is not an interesting one for a group of students, they might have a disadvantage compared to other students. The total number of questions in one context problem is usually restricted to three to five.

DESIGNING YOUR OWN BALANCED TEST
Teachers might be concerned that many of their students cannot solve level II or level III tasks. Most students cannot reason mathematically, argue with other students about their solution strategy, find mathematical proof for a statement, or generalize about solutions if they are not taught to do so. If teachers never pose higher-level questions during classroom discussions, they cannot expect students to answer these types of tasks during a written test. A test should reflect the teaching and learning process that preceded it, and if students are not taught to think for themselves but only to repeat exactly what they were told, they cannot solve higher-level problems. So designing a balanced test starts within the classroom, by discussing higher-level problems and their various solution strategies with students. Students should not be given quizzes consisting of level I questions only; tests in a variety of formats should be part of an assessment plan.

One advantage Dutch teachers have over their American counterparts is that they design and grade the school exams themselves, and they mark the national tests students need to take for their final examinations. The work is then sent to a randomly chosen school, and a mathematics teacher from this second school, who also had a group of students taking the same national exam, will provide another opinion. So teachers assess their own students' work as well as the work from other groups of students taught by colleagues from another school. While there is no feedback for students, since they leave school immediately after the examinations, teachers get much valuable feedback on their own work.

Many teachers complain that making a balanced
test takes much more time and effort than designing a simple quiz. Finding a good context within which to work is not simple at all. Here is what we do to find good contexts:

- Adapt exemplary problems made by others.
- Find newspaper articles about interesting problems. Often the wording needs to be changed for use with students and the task needs to be simplified.
- Make photographs of interesting objects of art in the neighborhood.
- Use (statistical) information from local authorities.
- Ask the applied mathematics group at the local university for formulas that are used in different professions.
- Use international research tests like the PISA study to find examples of tasks at different competency levels.

When my colleagues and I agreed to try making balanced tests, we started by changing and adapting only one problem in each test. Seeing what our students showed they were able to do helped keep this process going, because students often found strategies that we did not think of to solve a task. During team meetings, the content of the test was discussed and the test was reviewed. After using the test in the classroom, the content was discussed and reviewed again so an adapted version might be used during another school year. We also discuss student answers in class. For example, we give students different solutions from other, unidentified students along with the scoring guide and ask them to score and grade solutions themselves. They see value different strategies and better understand what sort of answer is expected from them. Some students, for example, always provide answers that are too elaborate, whereas other students need to provide more of an explanation.

By using balanced classroom tests, we want to do justice to the teaching and learning process; give students the feedback they need; and prepare them, not only for tests and examinations, but also for their future careers and further education.