On the way to open standards for education

Fritz Nestle, Ulm (Germany)
Nikolaus Nestle, Darmstadt (Germany)

Abstract: The newest answer of German education authorities to the problems of learning, the so called „Bildungsstandards“, is far away from any suitability. The more than forty years old dream of R. Mager in „Preparing Objectives for Programmed Instruction“ has got a great resonance in Germany too in that time. It was to early. Today we can realize this dream of modern learning environments in the world wide web: E-Testing as the base of successful E-Learning. The „Dortmunder Manifest“ presents the requirements. An adaptive basic program system in HTML/PHP with the respective properties to offer such checks will be given. A new form of evaluation is suggested.

Kurzreferat: Der Titel zitiert aus den am 4.12.2003 beschlossenen, verbal formulierten Bildungs-"Standards" der KMK. Sie machen den Lehrkräften viel unnütze, vermeidbare Arbeit. Im Vortrag wird gezeigt, wie man Bildungsstandards als interaktive Webseiten mit nicht all zu großem Aufwand (php) selbst erstellen kann. Damit wird ein Problem der Siebzigerjahre des vorigen Jahrhunderts elegant gelöst: Die Operationalisierung von Unterrichtszielen. Die Organisation von Lernen - und "Lehren" - muss nicht mehr künstlich auf dem Stand gehalten werden, den man bei der Schuhherstellung und in vielen anderen Bereichen schon vor zweihundert Jahren verlassen hat.

ZDM-Classification: B10, B50, C70, C80

1 Shoes and school organization

A comparison to the use of standards in other fields of society:

Two hundred years ago, shoes were manufactured individually by shoemakers. It was a complicated and cumbersome process in which the foot was measured, a last was carved according to the individual customers foot and about 150 separate parts were stitched together to a pair of shoes. When the shoe was finished like that, there was no choice for the customer but to wear it. There were no ready made shoes to choose from, and there was typically only one shoemaker serving a village or a small town.

While the making of shoes has moved to industrial mass-production since then, teachers still work pretty much in the same way as they've done 200 years ago. Lessons are hand-crafted one-offs, and the pupils have to arrange themselves with the teachers they happen to be assigned to. Not all of these teachers have the will and/or means to orient their work on the talents and needs of the individual pupil. The teacher's role is furthermore complicated by the fact that the educational system's goals are only poorly defined as it is for example the case in the new „norms“ for the 16 years old, fixed by German education authorities („Bildungsstandards für den Mittleren Schulabschluss“, Vereinbarung der Kultusministerkonferenz ((KMK), 05-12-03).

2 “Bildungsstandards” - impossible to extract useful conclusions

An example: „Pupils develop an understanding of natural numbers, integers, fractions and rational numbers and can use these types of numbers appropriately according to the requirements brought about by specific situations.“ („Die Schülerinnen und Schüler entwickeln sinntragende Vorstellungen von natürlichen, ganzen, gebrochenen und rationalen Zahlen und nutzen diese entsprechend der Verwendungsnotwendigkeit.“)

Such abstract phrases are not helpful for teachers, let alone for parents and pupils. On the basis of such norms, it is not really possible to deduce exact norms on the skills to be acquired in the respective math curriculum. Sill (Sill, H.D., Abstract to the GDM meeting 2005) has expressed this view on the „Bildungs-standards“: “The standards ... are provided on a very general level which makes it more or less impossible to extract useful conclusions for a well-founded curriculum development or the preparation of a lesson from them.” „(Die Bildungsstandards ... bewegen sich auf einer so allgemeinen Ebene, dass es kaum möglich ist, konstruktive Schlussfolgerungen für eine fundierte Curriculumentwicklung oder den Unterricht abzuleiten.“)

Our present organisation of learning in schools is obsolete. The training for future teachers is obsolete. It feels like training works for a modern shoe factory in the use of awl, cobbler's wax and twine – or even worse: in the theory how to use awl, cobbler's wax and twine.

3 Realizing the dream of Mager

The development of appropriate standards for skills in mathematics is still a great challenge for didactical efforts. This is especially the case if the standards are to be implemented in a way which is accessible to the learners themselves. The introduction of such transparent and democratic standards in math-learning could well be the greatest revolution in mathematics education within the last 200 years.

It would provide an incentive for the individual learners to reach a given standard at their own speed while classroom learning often leaves one part of learners behind while slowing down some others thus making ineffective use of the time budget of most learners in a class. Being exposed to lectures at the wrong learning speed may even spoil any innate motivation to learning in part of the pupils.

Are such transparent standards just a dream? Not at all. E-Testing in the way suggested by the „Dortmunder Manifest“ (http://www.bildungsoptionen.de/manifest.htm) can realize this vision. The „Dortmunder Manifest“ describes desirable properties of such standards; theses of one of the authors (FN) presented at the GDM spring meeting 2003 in Dortmund, Germany. They can be realized in the internet by means of simple PHP programs.

To some professionals in education, „teaching to the test“ is a horror vision. We don't understand this position.

What could be better than „learning for the test“, if the test itself matches the goals of education and is representative for the skills that should be mastered. The role of teachers working in a school system using such tests, would shift from lecturer-examiner to a coach helping individuals in optimal preparation for the challenges of learning.

Geering (Geering, P., abstract for 2005 spring meeting of GDM) notes „If there is marking, the learners will concentrate their efforts on what is rewarded by good marks.“ („Wo zensiert wird, erbringen Lernende diejenigen Leistungen, die durch Zensuren belohnt werden.“). It therefore seems to be a good idea to provide learners with a (freely accessible and transparent) objective standard provided in addition or instead of the standards set by an individual teacher. Such standards provide also a form of empowerment for self-evaluation. The access to test based on such standards could also open up a new choices for learning arrangements adapted to individual learners' needs (e.g. in continuing education or for mature students).

How can such standards be realized on an internet based testing platform?

This needs in a simple way a system composed of three main modules:

1. A sufficiently rich database of test items representative for each curriculum is required. From this database, adequate test items are selected and individualized in a random process for an individual e-testing event. The selection process can be guided by metadata on components of the curriculum covered by each test item. These metadata may also be used for providing additional feedback to the learner in self-testing situations.
2. A test platform which presents the selected items to the learner and accepts the students' solutions.
3. An evaluation program that analyzes the individual answers and computes the test scores. The most simple and reliable choice with respect to evaluation are test items with a single „correct“ answer. Sophisticated analyses of free answers may become more and more realistic in the future. However, in many cases this may prove unnecessary details as even complicated knowledge can often be examined by appropriately prepared questions with „simple“ answers.

After analysis and scoring, the learner is provided immediately with a score (and maybe also with a specific feedback on the answers to individual testing items).

In the framework of our present PHP system, these functionalities are realized by a range of different files generated for each e-testing event to a given subject XXXX (for instance “2min”):

the selection control file

xxxx.php (2min.php)

the yyy (yyy = 1, ..., y0) presentation files

xxxxayyy.php (2mina12.php)

the respective evaluation file

xxxxbyyy.php (2minb12.php)

a file to present the score list

xxxxsco.php (2minsco.php)

the respective data file

xxxxsco.txt (2minsco.txt)

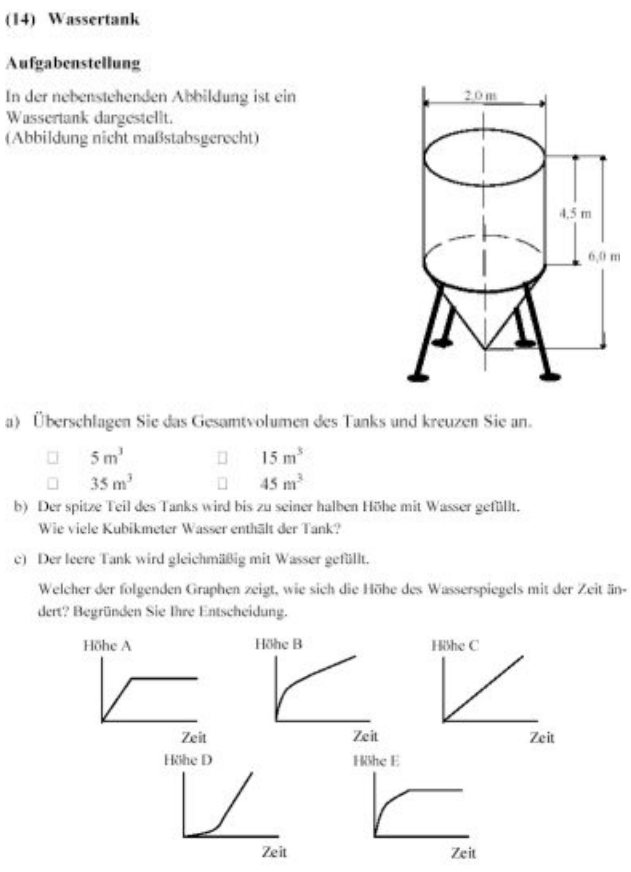
While the score list provides direct feedback to the user, more detailed information contained in the evaluation files may be used for a detailed statistical evaluation which may prove a new valuable tool for learning research (Nestle et al. 2005).

Examples where some of these functionalities are used can be found on internet, e.g. under
http://www.bildungsoptionen.de/dilli/fuchs.php (German for 4. form) or
http://www.bildungsoptionen.de/dilli/2min.php
(knowledge of fundamentals in economics)
http://www.bildungsoptionen.de/kmk/014.php KMK-example from “Bildungsstandards für den Mittleren Schulabschluss”
http://element.fkp.physik.tu-darmstadt.de/physik4bi/onlineuebung2005/uebung.php?idx=20 (physics for civil engineering undergraduates)

4 Back to mathematics:

In the German „Bildungsstandards“, some example problems are provided. Only one of these examples for 16 years old (example 14) allows a transparent evaluation und therefore is also accessible to automatic evaluation by means of simple PHP scripts or similar computer tools (See figure 1; in German). We will use this example to show the possibility to construct a class of equivalent problems:

Figure 1. Example 14 of KMK-Bildungsstandards für den Mittleren Schulabschluss



Math competences needed for solving this problem comprise:

1. Estimation of volume of an object composed of several simple geometric objects such as cylinder, cone and the like.
2. Calculation of a simple geometrical body's volume from data described verbally or in a sketch,
3. Qualitatively describing the kinetics of a process of filling or cutting determined by the geometry of the object to be manipulated.

Parts of these mathematical tasks occur in many areas of daily life or the workplace. A range of objects composed of a cylinder and a cone such as in the problem of figure 1 is given in figure 2. All of these situations could be used as settings for similar problems in a database of problems as required for an e-testing platform according to the „Dortmunder Manifest“. The random selection of the problem for an individual e-testing event would be done in two steps:

1. Random choice of a setting,
2. Random choice of parameters.

Figure 2. Other Settings

The range of different problems that can be generated in this subcategory of problem solving is so broad that memorizing doesn't seem like a practicable strategy of *„concentrating efforts on what is rewarded by good marks.“* Therefore, the presentation of one problem of this type is a valid way to test an individual's skills in solving this type of problem.

An appropriate strategy for producing a data base of problems is an open source/open content approach similar to the one pursued by Wikipedia. Such a work would be rooted in a community and does not need the establishment of new specialized institutions or councils.

5 Learning process and evaluation

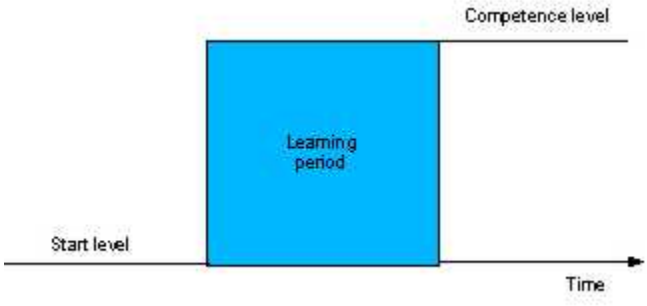
In addition to an assessment of the individuals' performance in the e-test, the solutions given by the students also provide a feedback on the viability of the test items themselves and the „difficulty“ of each test item.

Many traditions in test development try to impose a Gaussian distribution over the test results so that items which are solved by nearly all participants in a pre-test are excluded in the development of a test data base. This approach seems questionable as it is not compatible with the requirements of mastery learning. The pool of test items in a viable e-testing data base should therefore include such problems, too. Giving up the axiom of a Gaussian distribution of results allows more holistic feedback on the actual performance of each individual taking the test and furthermore will help bolstering the learners' self-confidence during the test.

It might be useful to take a look on the roots of a Gaussian distribution of test results being produced by the superposition of individual learning curves:

The typical course of a learning process is given in Figure 3: First the learner is essentially ignorant concerning a certain topic. This is followed by a period of uncertainty. Finally the learner has gained a new knowledge or competence – in the ideal case persisting for life.

Figure 3. Three stages of learning



In a standard school learning situation, the assessment of students' achievement takes place while most students are still somewhere in their learning period. As this period takes different time for individual learners, some of them are evaluated before they have managed to embrace the new topic, the majority is tested in a phase of random stability of the new competence while only a minority has reached the level of a stable life-long skill.

Thus we get nearly a Gaussian distribution of test results, but it's the a superposition of incomplete passages on individual learning curves and actually achievable levels of skills. For the individual, learning is completed if the learner passed from start level to the respective competence level.

Forcing a Gaussian distribution over the results of test items therefore mirrors a principal misunderstanding of learning: Items are sorted out, if nearly all of the pre-test population or only very few are able to solve the problems. This sorting is made without expert rating of the importance of each item.

This effect may lead to very odd results in the composition of tests. Thus for instance from the multiplication table multiplication with one factor zero were sorted out in a study with some hundred test persons, for there were not sufficient correct answers – but in accordance with this inappropriate evaluation theory! (Fricke, A., Über Meßmodelle in der Schulleistungsdiagnostik, Schwann, Düsseldorf 1972, S.93 -95)

Mastery learning in its pure form is an alternative paradigm for teaching and learning which is not biased by the need to create a Gaussian distribution of results.

The development of the database with test items of course needs an evaluation. In order to avoid the pitfalls of forcing statistics over testing results, expert rating seems to be appropriate for this. However, this raises the question: Who should be the experts to do this?

Again, the internet offers the novel and truly democratic choice to make the learners' evaluation of the test items a part of the evaluation process. Various forms of such evaluation strategies have developed in internet communities such as E-bay or Wikipedia. Of course the type of evaluation must be complemented by classical expert ratings and by the input from education professionals.

Instead of a classical summary or conclusion we want to end this contribution by repeating the three key features necessary for e-testing-based „Bildungsstandards“:

1. It needs a sufficiently large data base of test items for each topic (so that memorizing doesn't help).
2. There must be free access to random selected items for self-testing.
3. Immediate feedback for the learned must be provided in self-testing.

References

Fricke, A., (1972). Über Meßmodelle in der Schulleistungs-diagnostik, Schwann, Düsseldorf 1972, S.93 -95

Geering, P. , (2005). Abstract to 2005 spring meeting of GDM

Mager, R. F. (1962). Preparing Objectives for Programmes Instruction, Fearon San Francisco ,1962

Nestle N., Petermann C., Gädke, A. (2005). Eine korrekturfreundliche Nebenfachklausur in Physik – Eingliederung in ein Lehrkonzept, Möglichkeiten zur weiteren Automatisierung und Bearbeitungsanalysen *DPG-Frühjahrstagung Didaktik der Physik Berlin 2005* (Proceedings on CD-ROM, ed: Nordmeier V, Berlin: Lehmanns)

Nestle, F. (2003). Dortmunder Manifest, http://www.bildungsoptionen.de/manifest.htm (BZM 2003)

http://www.bildungsoptionen.de/dilli/fuchs.php (German for 4. form) or

http://www.bildungsoptionen.de/dilli/2min.php (knowledge of fundamentals in economics)

http://www.bildungsoptionen.de/kmk/014.php KMK-example from “Bildungsstandards für den Mittleren Schulabschluss”

http://element.fkp.physik.tu-darmstadt.de/physik4bi/onlineuebung2005/uebung.php?idx=20 (physics for civil engineering undergraduates)

KMK (2003). „Bildungsstandards für den Mittleren Schulabschluss“, Vereinbarung der Kultusministerkonferenz ((KMK), 05-12-03

Sill, H. D. (2005). Abstract to 2005 spring meeting of GDM

Authors

Nestle, F., Prof. a.D., (Institut für Mathematik und Informatik, Pädagogische Hochschule Ludwigsburg), Judenhof 2, D-89073 Ulm

Nestle, N., Prof. i.V., Institut für Festkörperphysik, Technische Universität Darmstadt, Hochschulstraße 6, D-64289 Darmstadt