

Didaktik der Naturwissenschaften

Bernd Ralle, Ingo Eilks (Eds.)

Promoting Successful Science Education – The Worth of Science Education Research

A collection of invited papers inspired by the
19th Symposium on Chemical and Science Education
held at the University of Dortmund, 22-24 May 2008

Shaker Verlag
Aachen 2008

Developing and Implementing New National Standards in Science Education: The Role of Science Educators

Peter Labudde

University of Applied Sciences North-Western Switzerland, Switzerland

In many European countries, TIMSS and PISA have initiated the development and implementation of standards in science education. What, however, is the role of the science educators in this process? Four tasks can be differentiated: Firstly, the development of a competence model with different dimensions (e.g. as it is the case in Switzerland where a three dimensional matrix consisting of three axes – skills, domains and levels – has been introduced. There are eight different skills, eight domains and four levels for grade two, six and nine, respectively). Secondly, science educators have to define and propose standards. This includes the validation of the model by different tests (e.g. paper-pencil and performance tests). Thirdly, they have to ensure the implementation of those standards – this is a task that starts at the very beginning of the development of the model and which leads to different measures. Science educators particularly involve the teachers in every single phase of the project that deals with "standards". Fourthly, the educators have to be aware of the responsibility they carry in a project that has such a huge political, educational and pedagogical impact as the development and implementation of standards does. In the following, the four roles are described theoretically and illustrated by examples from the Swiss project HarmoS (Harmonisation of the Compulsory School).

New national standards: The Swiss example

In many European countries the results of international studies such as TIMSS (Third International Mathematics and Science Study) and PISA (Programme for International Student Assessment) initiated a broad discussion about the objectives of (science) education, scientific literacy, curricula and standards. The public and the politicians worried about the future of their society and the education of the next generation. In many countries, policy makers have decided on developing new curricula and performance standards – in combination with other strategies such as the development/enhancement of teacher training programmes, in-service training or nationwide tests.

The development and implementation of new national standards is a demanding task, it is a challenge not only for politicians and teachers but also for science educators. But what exactly is the role of the latter in this process? In this article, four of their tasks are described and analysed, both generally and specifically by referring to the example of Switzerland (for an overview of the Swiss project "HarmoS" see the

information in the text box below and Labudde, 2007a, 2007b). The role of science educators include the following tasks:

1. Developing a competence model,
2. Defining the standards in a cyclic process,
3. Supporting the implementation of the standards,
4. Being aware of the responsibility that comes with this role.

The development of new curricula and standards is nothing new; every generation has had to perform this task. But something has changed: the standards set by the curriculum and/or the politicians now have to be evaluated and validated. Policy makers, researchers and teachers do no longer ask for just a new curriculum with new objectives - they ask for more, namely standards. Furthermore, these standards should be realistic, *i.e.* students should be able to attain them and teachers should be able to support their students in doing so. As a consequence the standards - even before they come into effect - have to be validated, *i.e.* they have to be evaluated. For a description of the whole process of developing and implementing standards see the report of Klieme *et al.* (2004) which has become some kind of "Bible" for the development of standards, at least in the German speaking countries.

New Standards in Switzerland: The HarmoS Project

The Swiss Conference of Cantonal Ministers of Education (EDK, *Erziehungsdirektorenkonferenz*) initiated the project HarmoS (*Harmonisierung obligatorische Schule*: Harmonisation of the Compulsory School). The project has the highest priority in the Swiss educational policy. Harmonisation is urgently needed: Switzerland has 26 states, so-called Cantons, with 26 different school systems. With 26 different systems for a population of only seven million people, Switzerland may well have the highest density of education systems in the world.

What are the main goals of the project? "HarmoS intends to establish comprehensive competency levels in specific core areas for compulsory schools in Switzerland. The necessary work in this respect is being accomplished on two levels: The pedagogic-didactic level includes the development of competency models which will make it possible to determine which levels of competency can be expected at a specific time in the compulsory school (grades 2, 6 and 9). - The political and legal part relates to the conclusion of an inter-cantonal agreement on the harmonisation of the compulsory school". (EDK 2004a; EDK 2004b; Linneweber-Lammerskitten & Wälti 2005) HarmoS includes four subjects: mathematics, science, first language, and second language. In each of these four subjects, standards will be developed for the end of grades 2, 6 and 9, *i.e.* for students who are about 8, 12 and 15 years old.

The project is divided into different steps and phases:

- 2001-2004 Framing the project HarmoS by the EDK (influenced by Klieme *et al.* 2004)
- 2004-2005 Submission of the project, *i.e.* call for proposals of research consortiums, allocation of the project to four consortiums, each of which is responsible for one of the four subjects (first language, second language, mathematics or science)

"Scientific phase": Research taking place under the responsibility of the consortiums. A few milestones:

- 2005-2006 Development of a competence model
- 2006-2007 Elaboration of hundreds of problems for the assessment of students in grade 2, 6 and 9 in order to validate the model

2007/08	Accomplishment of the assessment with a representative nationwide sample
2007-2008	Processing and interpretation of the data
2008	Based on the data, a proposal of basic standards in science education was presented to the policy makers, <i>i.e.</i> to the Swiss Conference of Cantonal Ministers of Education (EDK)
	"Political" phase of the project:
2009	Political determination and agreement on the basic standards by the EDK
From 2009	Implementation of the standards including the development of a core curriculum for each of the Swiss language regions (German, French, Italian, and Romansh), establishment of an educational monitoring system and of a further support system for schools and teachers.

Task No 1: The development of a competence model

A competence model that deserves the title "model" should include more than just a list of a few skills, more than a description of the contents of a subject, as it can often be found on the first pages of a traditional curriculum. The development of a multi-perspective competence model has turned out to be a first demanding task for science educators - a scientific challenge.

Both in Germany (KMK, 2004) and in Switzerland, researchers have developed a three dimensional model, in Germany for each of the three sciences, *i.e.* for biology, chemistry and physics, in Switzerland - due to a different approach in science education that is more related to the idea of Science-Technology-Society (STS) - a competence model for science as a whole. The model includes three axes: skills, domains and levels (Figure 1).

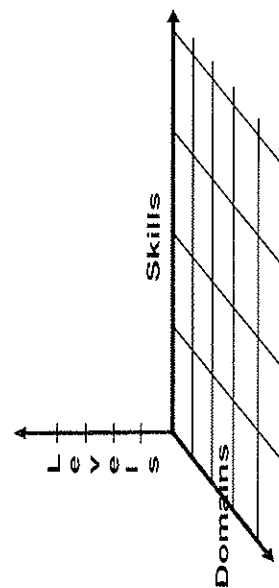


Figure 1. The three dimensional Swiss competence model for science education

The primary axis of the matrix designates the skills. It includes eight aspects: 1) displaying interest and being curious, 2) asking questions and investigating, 3) exploiting information sources (see figure 5 for an example), 4) organising, structuring and modelling, 5) assessing and judging, 6) developing and realising, 7) communicating and exchanging views, 8) working self-reliantly and reflecting on one's own work. The members of the HarmoS Consortium Science claim that this list of skills is final.

The axis of the domains includes 1) planet earth, 2) motion, force, energy, 3) communication, regulation, and control, 4) structure and properties of matter, 5)

plants and animals, 6) ecosystems, 7) human body, health, well-being, 8) nature, society, technology, perspectives. Here, in contrast to the first axis, the claim is not made that this list is final: A competence model is not the same as a curriculum that includes a final list of contents.

The third axis, the axis that designates the level, includes four levels (I – IV) for each of the grades 2, 6 and 9. There is an overlap between the levels of different grades, *i.e.* the highest level of grade 2 (titled IV/2) is the same as the lowest level of grade 6 (titled I/6), thus, IV/2 corresponds exactly to I/6 (Figure 2).

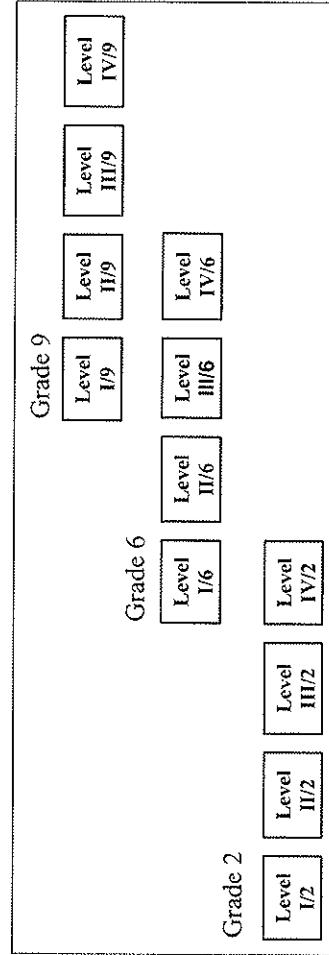


Figure 2. Four levels for each of the grades 2, 6 and 9, partly overlapping

Figure 2 illustrates that the competence model is a progressive one, *i.e.* that a student should/ can develop his or her competency from level I/2 to level IV/9. But it is not assumed that the model or the progression is a "hard" one. It is not assumed that students will progress at the same pace from one level to the next one or that they will remain at each level for an equal amount of time. On the contrary, this progression is "soft": Each individual progresses at his or her own pace, every student rests on different levels for a different amount of time.

Each of the eight skills - and their sub-skills - corresponding to the relevant levels have been described meticulously and sophisticatedly: Eight tables fill more than 20 pages with "can-do-descriptions". *E.g.* the skill "exploiting information sources" includes five sub-skills: Being able 1) to search for information, 2) to realise different kinds of information, 3) to read information, 4) to classify and validate information, 5) to apply information. Each sub-skill includes 12 levels, *i.e.* three (for grades 2, 6 and 9) times four; due to the overlap (see figure 2) there are only 9 entirely different levels. Figure 3 gives an example for the sub-skill "reading information": it contains one description for grade 2 (the lowest level of all, *i.e.* I/2), one for grade 6 (a "medium" one, *i.e.* II/6) and one for grade 9 (the highest level of all, *i.e.* IV/9).

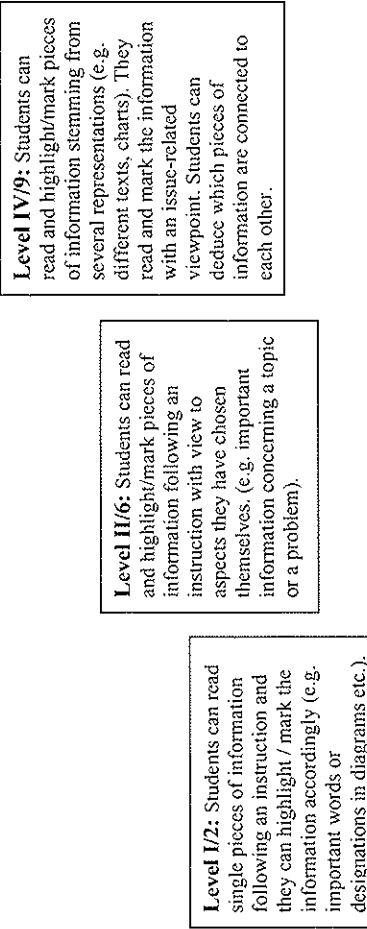


Figure 3. Three levels (out of nine) of the sub-skill "to read information".

This competence model has not been developed from scratch. It has - and it must have - its roots in the real, daily world (for more details see Labudde, 2007a). Thus, the following aspects have been taken into account and included in the competence model:

- Confirmed empirical results of research in science education,
- The current valid science curricula of a country (before developing the competence model all 26 curricula of the 26 cantons in Switzerland had been analysed),
- General objectives of education of the particular country,
- The structural frame of science education in the country (*e.g.* the formal and informal qualification of the science teachers, the possibilities of in-service training, the number of science lessons per week),
- International tendencies (like the concept of scientific literacy in PISA).

Task No 2: Defining the standards in a cyclic process

The frame of the future standards is set by politicians. In Switzerland, *e.g.* they have settled on performance standards that apply for students. The standards are basic ones (abroad they are also called minimal standards) and they will set the frame for a future curriculum, planned for 2012. The two main objectives of the implementation of standards in Switzerland are the following: Firstly the harmonisation of the school systems of the 26 different Cantons and, secondly, the idea that the standards will form the basis for a future system of monitoring of the educational system in Switzerland, *i.e.* a low-stakes assessment.

Within the given frame it is the job of science educators - in co-operation with other key stake holders - to develop and to propose the standards. The first step is the development of a competence model as described above. This is a process mainly

driven by normative decisions. However, setting standards also requires an empirical validation of the model and the proposed standards. For example, are the standards defined as basic standards - really basic or are they too demanding for most of the students? Thus, the standards have to be validated in a nationwide assessment with a representative sample (for examples of the assessment see figures 5 and 6). The results of this validation will lead to a revision of the competence model and the standards, which then again will be subjected to further empirical assessment: a continuous cyclic process (Figure 4).

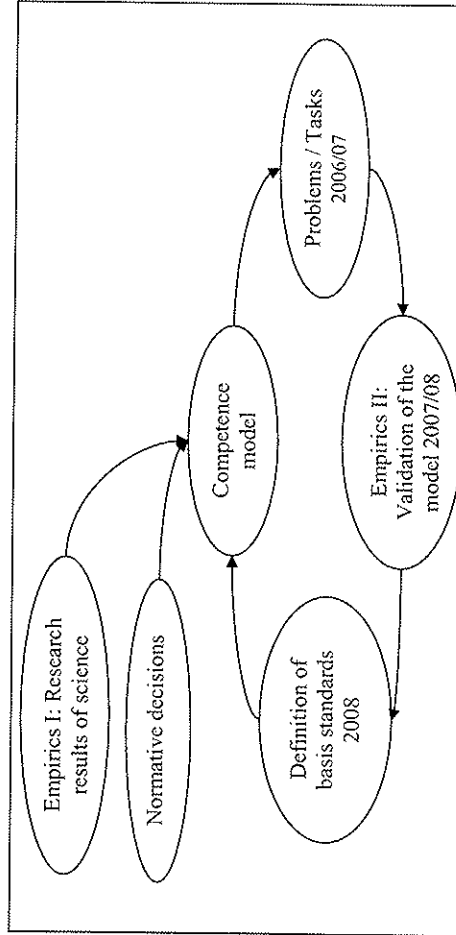


Figure 4. The cyclic process of developing standards (with years of the first cycle in Switzerland).

The consortium of HarmoS Science has developed a thorough test design in order to validate the competence model. The design includes the different skills, domains and levels (see "Konsortium HarmoS Naturwissenschaften", in preparation):

Grade 6 and 9, paper&pencil tests: Nationwide validation, representative sample for the German and French speaking part of Switzerland, N=4'600 for grades 6 and 9, respectively, 48 situations comparable to the PISA-problems for each of the two grades (some of the situations being the same for grades 6 and 9), each situation with 5-7 items measuring one of the three skills "exploiting information", "organising, structuring and modelling" and "assessing and judging". Each item was attributed to a specific level from I to IV, the items were either multiple choice ones or with open answers.

Grade 9, PISA-results: Re-categorisation of the PISA-problems from the perspective of the skills as defined in the competence model and further re-analysis of the SwissPISA-results.

Grade 9, performance test (hands-on-activities): 45 classes (N about 800) from the three linguistic parts of Switzerland, 12 experiments similar to those of the TIMSS-performance-test (Harmon *et al.*, 1997), each experiment lasted about 30 minutes and

tested different skills, in particular the skill "asking questions and investigating", but also the skill "developing and realising". Each student had to work on his or her own, had to project, perform and evaluate an experiment and had to fill in a test-sheet.

Grade 6, performance test (hands-on-activities): Similar to the performance test in grade 9, but only with 8 experiments and 30 classes from the German and French speaking part of Switzerland.

Grade 2, combined paper&pencil and performance test: 30 classes (N about 570) from the German and French speaking part of Switzerland, 16 paper&pencil situations (80 items) and 8 experiments (50 items) similar to those mentioned before, except for the fact that the children were guided through the test and were allowed to ask for a writing assistance.

All problems have been piloted in at least two classes some months before the main test. (Due to externally caused time restrictions and problems it was it not possible to perform a pilot phase with a sample big enough in order to evaluate the items in such a rigid way that would have allowed for a serious statistical analysis). The corrections of the tests, the data recording and analysis followed a well-defined procedure, similar to that of international studies like TIMSS or PISA: For all paper&pencil situations and for all experiments of the performance test a correction scheme has been developed. Native speakers (German, French and Italian) with a background in science, most of them university students, corrected the tests. About 15% of the items have been double-coded in order to check the interrater-reliability. By statistical methods such as Rasch analyses the validity and reliability of the items were controlled. This led to a reject of about 20% of all items due to translation errors, ceiling effects or other problems and faults. For further information see "Konsortium HarmoS Naturwissenschaften" (in preparation).

In order to provide an insight into the type of problem asked two examples are given below, one paper&pencil item for grades 6 and 9 and one experimental task for grade 2 (see Figures 5 and 6). For more examples we refer to *harmos.phbern.ch* and to Metzger & Labudde (2007).

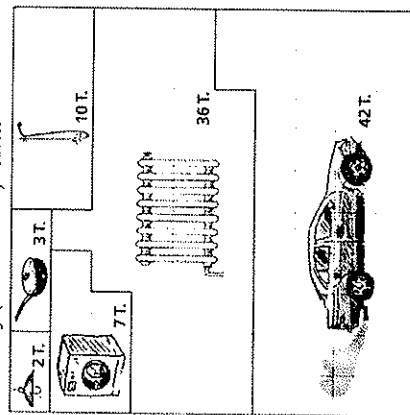
This item tests the skill "exploiting information sources". The experts have assumed that the level of this task is 9/I, i.e. the lowest one for grade 9 and - at the same time - level III for grade 6 (see the overlapping of the two levels in figure 2). A student gets 2 points if the answer is completely correct (1st car, 2nd heating, 3rd shower, 4th washing, 5th cooking, 6th light), 1 point if the 1st and 2nd rank were correct and only two other ranks were interchanged, 0 points for any other answer.

The number of completely correct answers was low: In grade 9 only 56% of N=320 students, in grade 6 about 45% of N=344 students have answered correctly. These results have surprised all experts, teachers and science educators; all of them had assumed a better performance of the students. Results like these lead to general questions: Can an item like this be seen as paradigmatic for illustrating the basic standards of the skill "exploiting information sources" in grade 9? Or is it still too difficult? Or is the item now, i.e. in 2008, too difficult, but should and could it be

used to illustrate the expectations and the level that students should reach in the far future? These questions may show how the definition of standards depends both on normative decisions and empirical evidence.

The Things that use a Lot of Energy

The diagram below tells you what we use our energy for. This is shown: Out of 100 parts (T) of energy in Switzerland we use so many (amount of T) for....



Develop a ranking (ranks 1 to 6, 1st rank: most energy is used on....)

light	washing	heating
cars, mopeds	cooking	showering, bathing

Figure 5. Typical paper & pencil item for classes 6 and 9.

A hands-on-activity for grade 2 deals with water: the pupils are asked to investigate the properties of water. Figure 6 shows the first part of the activity of totally seven parts. In all parts the pupils either have to observe something, have to plan and perform little experiments or have to describe and explain their observations. The pupils are guided through the whole of the hands-on-activity. If they have problems with writing and expressing themselves in sentences – and this happens sometimes with seven year old children – they were allowed to ask for a "writing assistance".

The two examples given above are paradigmatic for more than 700 items that students had to answer. The results of all the tests allow a huge variety of analyses e.g. regarding the difficulty of an item or a problem, gender differences, differences between the linguistic regions of Switzerland, comparisons between grades 2, 6 and 9, correlations between the skills and/or the domains. But above all, they allow for a validation of the competence model and they are a solid basis for the proposal of standards.



Figure 6. Investigating water, a pupil of grade 2 at work and the instruction given to the pupils.

What happens to the water?

Three types of materials are on the surface:

- a paper serviette at the top
 - rubber foam is in the middle
 - at the bottom is the poly pocket
- Pour exactly one spoonful of water at a time on each of those surfaces (paper serviette, rubber foam, poly pocket).

Carefully observe what is happening! Write down your observations and make a drawing of them.

Task No 3: Supporting the implementation of standards

From the very beginning of the development of standards their implementation should be kept in mind. Developing standards without already planning the implementation is worthless and condemned to failure. A first step and probably one of the most important ones is the involvement of practitioners, or better "reflective practitioners", from the start of the whole developmental process. Their expertise, together with the knowledge and skills of science educators, creates a beneficial condition for standards that are both realistic and innovative. Apart from the teachers other key stake holders (e.g. trade-unionists, policy makers and researchers) should be included since they come from a different background, e.g. in pedagogy, statistics or social studies.

In Switzerland, more than 100 people were engaged in the project HarmoS science. The overall team consisted of: a) A core team of 20 people, most of them science educators who had been or still are part-time teachers, b) more than 40 teachers from grades 1-9 from all linguistic regions of Switzerland, c) about 30 research assistants, most of them science teacher students, d) a guiding group of 12 key stake holders like mentioned above, e) a group of four experts in psychometrics. Thanks to this composition all relevant groups were represented in the project HarmoS. All those people did not only bring their expertise into the project, but they also became the "ambassadors", i.e. implementers of the standards.

After having defined and determined the competence model and the standards the following implementation of standards can be supported by different means such as:

- In-service training programmes and school development projects with a focus on the collaborative learning of teachers,
- An item bank with 'best-of-practice' examples,
- An item bank with paradigmatic problems or science projects,
- Tests with a representative national sample that monitor the system without benchmarking teachers or schools.

However, the most important issue is bringing the approaches that proved to be successful to "normal" classes.

Can standards be implemented without some form of centralised assessment? Some countries like England or France rely heavily on centralised assessment, others like Finland, Sweden or Switzerland do definitely not. If one wants to implement standards without centralised assessment three problems have to be dealt with. How can one achieve 1) a formal comparison of the learning results of one class with those of other classes of other teachers and/or of other schools, 2) a structured comparison between the outcomes of a class and the outcome intended by national standards, 3) a professional feedback for the teacher. There are several possibilities of how to approach these problems, one of them being the structured learning of teachers with, about and of each other. As described elsewhere (Labudde 2007a, p. 294) these options must be introduced in a formal, official way; and when doing so, suitable structures must be created for this purpose. This may be a difficult and long way, but it is a way that seems to be worthwhile and successful – as the examples of Finland and Sweden show (Waddington, Nentwig & Schanze, 2007).

Task No 4: Being aware of the responsibility

Developing a competence model and proposing standards to the politicians is different from normal research projects in science education that deal with studying learning processes in science or developing new teaching materials and units. Competence models and standards have a huge political, educational and pedagogical impact. As a science educator one should be aware of one's responsibility. A continuous dialogue between science educators and politicians, even after having "finished" the job of having proposed a model and standards, is mandatory. Science educators should inform policy makers about issues such as:

- The range and the limits of the proposed competence model and its validation,
- The advantages and disadvantages of basis, regular or excellence standards in science,
- The intended and the non-intended consequences of centralised assessment, e.g. its contribution to more equal opportunity but also its effect of closing down innovation,
- The limits of centralised assessments, e.g. the fact that a competency like conducting a complex science project cannot be assessed on a national level,
- The broad range of skills that are generally not covered by common standards, e.g. social skills such as being able to co-operate or being tolerant,
- The danger that teachers, students and parents could only focus on subjects where standards have been defined, e.g. first and second language, mathematics and science, while forgetting other subjects like music or sports.

This list is neither exhaustive nor scientifically neutral. The implementation of standards cannot be neutral, because its main drive is political. The task of science educators is not completed with the development and validation of a competence model and with the proposal of standards. Informing and counselling policy makers

and politicians are never ending duties for science educators working in the domain of competence models and standards. When doing so science educators show competency and set standards – for their own profession, for research in general and for the development of daily science instruction.

Acknowledgements

This article would not have been possible without the outstanding work of my colleagues in our consortium HarmoS Science. Without their support and our fruitful discussions about standards in science education, this publication could not have been written. Warm thanks to M. Adamina, L. Bazzigher, B. Bringold, U. Frischknecht, P. Gigon, F. Gingsins, C. Gut, B. Jaun-Holderegger, A. Jeitzer, B. Kuterim, S. Metzger, C. Nidegger, K. Raths, R. Stebler, P.-Y. Theurillat, M. Vetterli, U. Wagner, C. Weber and A. Zeyer.

References

- EDK (2004a). *HarmoS - Zielsetzungen und Konzeption [HarmoS – Objectives and Concept]*. Bern, Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK). http://www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html (last access 20/06/08).
- EDK (2004b). *Projekt HarMoS - Offertenausschreibung für Entwicklung von Kompetenzmodellen [Project HarMoS – Announcement for the development of competence models]*. Bern, Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK). www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html (last access 01/06/08).
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V., et al. (1997). *Performance assessment in IEA's Third International Mathematics and Science Study*. Chestnut Hill: TIMSS International Study Center, Boston College.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., et al. (eds.) (2004). *Zur Entwicklung nationaler Bildungsstandards: Eine Expertise [On the development of national educational standards – An expertise]*. Bonn: Bundesministerium für Bildung und Forschung.
- KMK. (2004). *Bildungsstandards in Physik für den Mittleren Schulabschluss [Educational standards in Physics for a medium level school degree]*. Bonn: Kultusministerkonferenz.
- Konsortium HarMoS Naturwissenschaften (in preparation): *Schlussbericht HarMoS Naturwissenschaften zu Händen der EDK [Final report HarMoS science for the EDK]*. Bern: Deutschsprachige Pädagogische Hochschule Bern. For more information (only in German and French) see <http://harmos.phbern.ch> (last access 20/6/08).
- Labudde, P. (2007a). How to develop, implement and assess standards in science education? 12 challenges from a Swiss perspective. In D. Waddington, P. Nentwig & S. Schanze (eds.): *Making it comparable: Standards in science education* (pp. 277-301). Münster: Waxmann.
- Labudde, P. (2007b). Naturwissenschaftliche Bildung: Quo vadis? [Scientific Literacy: Quo vadis?]. In P. Labudde (ed.): *Bildungsstandards am Gymnasium: Korsett oder Katalysator?* (pp. 283-291). Bern: hep.
- Linneweber-Lammerskitten, H., & Wälti, B. (2005). Is the definition of mathematics as used in PISA Assessment Framework applicable to the HarMoS Project? *Zentralblatt für Didaktik der Mathematik* 37, 402-407.
- Metzger, S., & Labudde, P. (2007). HarMoS Naturwissenschaften+: Bildungsstandards für die Schweiz [HarmoS Science+: Educational standards for Switzerland]. *Praxis der Naturwissenschaften Physik*, 56(6), 14-18.
- Waddington, D., Nentwig, P., & Schanze, S. (eds.) (2007). *Making it possible - Standards in science education*. Münster: Waxmann.