The Development, Implementation and Evaluation of Science Education Standards in Germany

Peter Nentwig

Leibniz Institute for Science and Mathematics Education at the University of Kiel (IPN)

Email: peter.nentwig@gmx.de

Abstract: As a consequence of the unsatisfactory results in international comparative studies such as TIMSS and PISA, the ministers of education in the 16 states of the Federal Republic of Germany agreed on the introduction of common national educational standards to ensure quality in schools. The German Conference of the Ministers of Education (Kultusministerkonferenz, KMK) published standards for the science subjects biology, chemistry and physics for the end of lower secondary school (grade 10; students 15-16 years old) in 2004. These standards were to be concretized on state level and are mandatory in school since the school year 2005/06. The new standards have shifted the focus from input- to output-control of the school system – from syllabus to be taught to performance to be observed. They also introduced "competences" as the new paradigm for teaching and learning. In order to monitor the target achievement of the new output-orientation, the Institute for Educational Progress (IQB) was founded with the major task to develop, conduct and evaluate tests, which have since been attached to the regular PISA procedures. The German process of developing, implementing and evaluating standards for science education is presented and discussed in the context of similar developments in other countries.

Keywords: standards, competences, learning outcomes, secondary education

1 Introduction

Germany's educational system had been seen as ambitious and efficient, not least in the system's self-perception. Therefore, the results of international comparative studies such as TIMSS (cf. Beaton et al. 1996; TIMSS 1998) and PISA (OECD 2001, 2004) came as a shock. Germany had not partaken in international comparisons for some time, and the educational authorities had not expected the country to rank below the top. According to the underlying PISA concept of scientific literacy (cf. OECD 2006), the ability to apply scientific knowledge in complex tasks had been the test emphasis. German students while demonstrating satisfying supplies of factual knowledge showed a disconcerting weakness in making use of this knowledge. These results were considered alarming particularly because application oriented knowledge and understanding was then and is now seen as crucial for citizens' lifelong adjustment to an ever changing technology-dominated world. Not least was the poor performance of young persons, who would soon leave school and enter a professional life, seen as risk for the country's economic future. Something had to be done.

Traditionally, science education in Germany was input-steered by the syllabus. It specified what teachers were required to teach grade by grade, subject by subject. The underlying implication was that students would learn what was taught. What common sense had long doubted was officially attested with the results of PISA and other studies: input and output were not congruent. German 15-year-olds performed below expectation. Therefore the "Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany" (Kultusministerkonferenz, KMK) decided to change the system from input-orientation to output-orientation. Rather than determine what was to be taught, it was now defined what was to be learned.

In 2003, achievement standards were issued for German language, Mathematics, and first foreign language (KMK 2003a, 2003b, 2003c, 2003d), followed in 2004 by standards for biology, chemistry and physics (KMK 2004a, 2004b, 2004c, 2004d). These standards specify what students are expected to have learnt by the end of lower secondary education (grade 10, students about 16 years old).

In this development, Germany was not alone at the time. While in some countries (e.g. the United States or the UK) a tradition of education standards had already existed, in several continental European countries a movement towards educational standards evolved in response to the results of international comparative assessments (cf. Waddington et al. 2007, Schanze & Nentwig 2008, DeBoer 2011).

2 Definition of Standards

In the past, science teaching in Germany had been ruled by the syllabus for each subject. These syllabi had strongly focused on content, mainly organized by the structure of the discipline as represented in textbooks. The new education standards now introduced "competence" as the new paradigm. This development was strongly influenced by a report presented by a group of renowned educational experts (Klieme et al. 2003) on behalf of the KMK, in which the concept and purpose of standards were laid out. In this report, the notion of competence played a dominant role, and it has since been the prevalent term in the German speaking countries and beyond in continental Europe. Together with "competences", the categories "basic concepts" and "attainment levels" are constitutive for the definition of education standards in Germany. Based on the above mentioned report and in the light of the contemporary state of the art in science education, groups of experts from biology, chemistry and physics education were engaged by the KMK to write up the standards for their respective subjects.

2.1 Competences

"Competence" as a term is used both in everyday language and in the educational discourse. In the German speaking countries, the term is most often used with reference to the explication of Weinert (2001) as clusters of cognitive prerequisites that must be available for an individual to perform well in a particular content area. According to Weinert, competences are individually available or learnable cognitive prerequisites and skills to solve particular problems. Beyond the cognitive aspects, Weinert's original definition also includes motivational, volitional, and social willingness to use this problem solving ability in varying situations. This extension, however, is mostly disregarded in the term's use for standards formulation, mainly for practical reasons. Motivational and volitional influences on performance as result of competence would be too difficult to measure in the assessment of student learning. Therefore, the definition of competence used in the priority program "Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes" of the German Research Foundation (DFG) describes competences as "context-specific, cognitive achievement dispositions" (Klieme & Leutner 2006, p. 879). It has been widely accepted to see motivation not as component of the competence concept, but rather as an important condition for specifying the relation between competence and performance. Thus, competence-related motivational attitudes should be measured separately (cf. Weinert 1999).

In the process of developing standards, the role of intelligence as a confounding factor in the assessment of competence has been debated. It appears obvious that the development of knowledge and skills is influenced by intelligence and that, from a psychological point of view, competence and intelligence are not completely separable. However, the restriction of competence to domain-specific knowledge and skills in the definition of Weinert is widely seen as functional and acceptable, because it places more emphasis on the learned and learnable aspects of education in school settings than on general cognitive abilities (cf. McClelland 1973).

Finally, competences are seen as domain-specific because they are acquired in specific situations and are linked to content-specific, task-specific, and demand-specific knowledge and experience. In the course of a cumulative learning process, these domain-specific competences can be generalized (cf. Klieme et al. 2003).

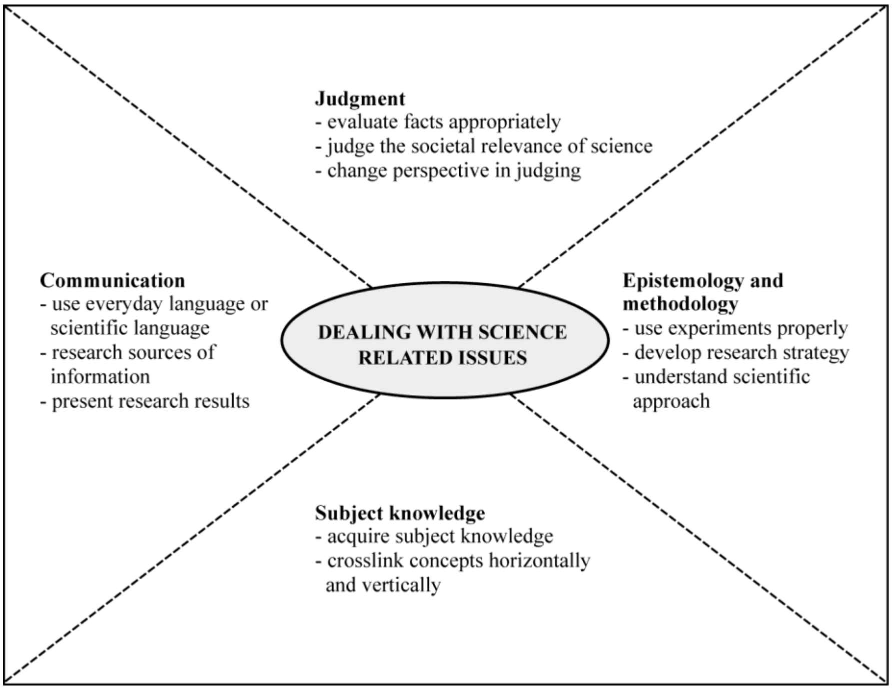
In summary, this concept of competence stresses the functional understanding of the term. Competence is directly linked to the application and use of the abilities, proficiencies, or skills that are necessary to reach a specific goal.

For the science subjects, the authors of the standards defined competences in four dimensions, so called competence areas:

- *subject knowledge*
  to know science phenomena, facts, laws, terms and concepts
- *epistemology and methodology*
  to use experimental and non-experimental research methods and conceptual models
- *communication*
  to retrieve, use and exchange subject related information
- *evaluation and judgment*
  to evaluate and judge subject related issues

While the content-oriented subject knowledge dimension still is a rather traditional one of knowledge recall, the other three are process-oriented and focus on the application of knowledge and skills. Figure 1 illustrates the interaction of the four competence areas. They will be further explicated below.

Figure 1 Interaction of the four competence areas in dealing with science related issues. (cited from Niedersächsisches Kultusministerium 2007 in this author's translation).



In the German science education community, a large variety of competence models are being developed that go beyond the one originally used for the development of standards. Most are more differentiated, many are more theory driven, some are empirically validated (cf. Bernholt et al. 2009, Kauertz et al. 2010). Particular competence models have been developed e.g. for the subject knowledge competence area (Neumann et al. 2010) and for the area of epistemology and methodology (Neumann 2011). Further models have been proposed for communication competence (Kulgemayer & Schecker 2009) and for judgment competence (Hostenbach et al. 2011). While these models describe structures of competence in a rather static way, attempts have also been made towards dynamic models of competence development (Hammann 2004, Aufschnaiter & Rogge 2010). All of these more sophisticated models have yet to be taken account of in the development of future versions of the science education standards.

2.2 Basic concepts

Syllabi for the science subjects used to represent what is seen in textbooks as the structure of the discipline. Following a brief preamble with rather general goals of teaching the particular subject, traditional syllabi would be catalogues of subject knowledge from simple facts to more sophisticated concepts. They specified what teachers were required to teach grade by grade, subject by subject.

For the new standards, a different approach was chosen. A limited number of basic concepts was agreed upon in each expert group that were seen as essential for the subject. For chemistry, for example, they were:

- **The matter – particle relationship**
  For chemistry it is assumed that all matter is built of sub-microscopically small particles. They do not usually exist isolated, but rather aggregate to compounds by forming chemical bonds. They form aggregates with specific properties (e.g. forming metals or salt crystals). The variety of materials results from the multitude of possible combinations of a limited number of different atoms.
- **The structure – property relationship**
  The properties of matter result from the kind of particles they are built of and from the structure of the aggregate. Combination and structure of the aggregate determine the properties more than the character of the individual atoms.
- **The character of chemical reactions**
  In chemical reactions substances are transformed. Particles and aggregates interact with each other in processes of attraction and repulsion, thus forming and breaking chemical bonds.
- **The energetic aspects of chemical transformations**
  Energy is stored in all substances. The amount of energy stored is a characteristic property. In chemical reactions, the energy amount within the reacting system changes by energy-exchange with the environment.

The authors of the standards for physics education named *matter*, *interaction*, *system* and *energy* as fundamental for their domain, while the biologists identified *system*, *structure and function*, and *development* as their three basic concepts.

With these basic concepts, science phenomena and processes that are relevant at school level can be organized and interpreted. They are the fundament for students to systematically and cumulatively construct their science knowledge and understanding. In the German science education discussion, this process is called *vertical* crosslinking (cf. Glemnitz 2007, Wadouh 2007). At the same time, an understanding of these basic concepts supports *horizontal* crosslinking (cf. BLK 1997, Nentwig 2009), i.e. connecting knowledge from one subject domain with experience from other contexts. Students encounter the structure - property relationship in chemistry as well as in biology, or they learn to regard physical as well as biologic phenomena as systems.

Traditional syllabi tend to be overloaded with content. Basic concepts lead to comprehensible structures and help to narrow the bulk down to reasonable amounts of indispensable content.

2.3 Attainment levels

Ultimately, the development of standards should lead to empirically validated models of attainment levels. These should describe different levels of proficiency of students in the different competence areas. This task has been accomplished for German language, First Foreign Language (French or English), and Mathematics (IQB 2012a-f). The empirical validation was docked to the data acquisition of recent PISA rounds. As science was last in the PISA sequence, data were just recently collected, and their evaluation is currently not yet available.

Therefore only normatively expected attainment levels can be reported at this time. Three levels of cognitive complexity were proposed for all four areas as a rather rough measure:

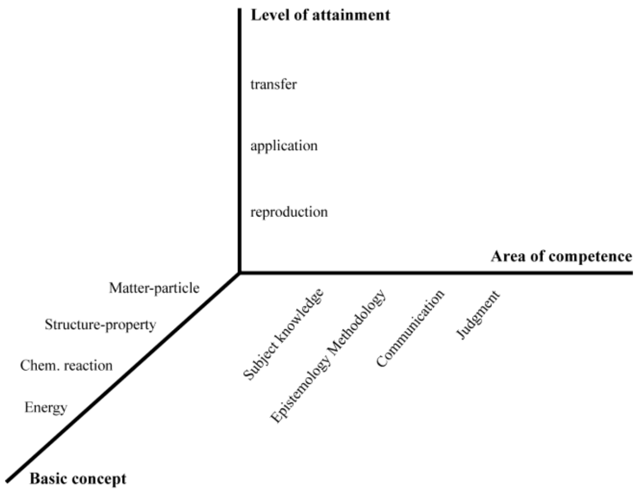
I *reproduction* of facts, methods and skills
II *application* of facts, methods and skills in new contexts
III *reflection* of necessary facts, methods and skills for self-dependent problem solving

This rather cursory taxonomy was then applied to the four areas of competence (table 1).

Table 1 Levels of attainment in four competence areas

		Level of attainment		
		I	II	III
Area of competence	Subject knowledge	Reproduce facts and concepts	Choose appropriate facts and concepts to solve given problems	Solve complex problems constructively using reflected choice of facts and concepts
	Epistemology Methodology	Describe methods and conduct experiments by instruction	Choose suitable method to solve given problem	Solve complex problems constructively using reflected choice of method
	Communication	Grasp information from and reproduce in subject related modes of representation	Search information and present in addressee related modes	Search and use information to argue in a reflected way
	Judgment	Understand and reproduce science content within an argument	Choose and use science content for an argument to judge a debated issue	Choose arguments from different perspectives to reflect science related decisions

Fig. 2 Three-dimensional model of three categories defining chemistry standards



Eventually, a tiered model of competence levels will be produced post-hoc from the empirical data, following the procedures of PISA and other large-scale assessments. An example of what it might look like will be shown below from the evaluation of student attainment of standards in Mathematics education.

Competence, basic concepts and attainment levels are the three categories, by which the science standards are defined in Germany. This can best be illustrated in a simplifying three-dimensional model (figure 2):

3 Science standards examples

Table 2 Standards for the competence area *subject knowledge* in chemistry (cited from KMK 2004c, p. 11 in this author's translation)

To know chemical phenomena, facts, terms and concepts	
F1	**The matter – particle relationship**
	Students …
F 1.1	name and describe basic chemical substances and their typical properties.
F 1.2	describe the submicroscopic composition of select substances using models.
F 1.3	describe the composition of atoms using an appropriate atomic model.
F 1.4	use models of chemical bonding to interpret particle aggregates, steric structures and intermolecular interactions.
F 1.5	explain the diversity of matter based on varying combinations and arrangements of particles.
F2	**The structure – property relationship**
	Students …
F 2.1	describe principles of classification for substances, e.g. by their typical properties or composition of particles.
F 2.2	use an appropriate particle model to explain properties of substances
F 2.3	deduce potential use of substances from their properties and reflect advantages and risks
F3	**The character of chemical reactions**
	Students …
F 3.1	describe phenomena of matter- and energy-transformation.
F 3.2	explain matter- and energy-transformation with rearrangement of particles.
F 3.3	identify the transferred particles in select donor-acceptor-reactions and name the type of reaction.
F 3.4	write reaction equations using their knowledge of conservation of atoms and of constant atomic proportions.
F 3.5	describe the reversibility of chemical reactions.
F 3.6	describe exemplary cycles of matter in nature and technology as systems of chemical reactions.
F 3.7	describe ways of controlling chemical reactions by variation of parameters.
F 4	**The energetic aspects of chemical transformations**
	Students …
F 4.1	declare that in chemical reactions the energy content of a system changes through energy exchange with the environment
F 4.2	explain energetic phenomena in chemical reactions with the transformation of one energy form into another
F 4.3	describe ways of controlling chemical reactions by use of catalysts.

Above, the science standards have been defined in rather broad terms. To fill them with life, basic concepts and competence areas needed to be more explicit. Table 2 gives an example, how they are set in relation to each other. The standards for the competence area subject knowledge are arranged by *basic concepts:*

In comparable ways, the standards for the other competence areas and for the other science subjects were formulated as expected performance of students. Table 3 gives an example from the physics standards:

Table 3 Standards for the competence area *communication* in Physics (cited from KMK 2004d, 12 in this author's translation)

Research sources of information and present results	
	Students …
K1	communicate physics related insights and their possible application, using appropriate terms and modes of representation.
K2	distinguish everyday language from technical language in describing physical phenomena.
K3	research various sources for information.
K4	describe the set-up of simple technical appliances and their function.
K5	document their research results.
K6	present their research results in addressee-related modes.
K7	discuss their research results from physical points of view.

For each of the sciences biology, chemistry and physics and for each of the four areas of competence, such lists of expected competences were written by the groups of science education experts on behalf of the Ministers of Education.

In the next step, the intentions of the standards were illustrated with exemplary tasks. These tasks demonstrated, how competences should be brought to use. In the style of PISA tasks, the questions and problems were embedded in commonplace contexts. An explanation was added to each task showing in detail, which competence was addressed, and which level of attainment the solution of the task would demonstrate. Table 4 gives an example from the chemistry standards.

4 Implementation of standards

Germany is a federal republic with 16 states. With a minimum of uniformity, stipulated in interstate agreements, the states have always maintained their authority in all cultural and educational affairs. School systems, teacher training and syllabi vary from state to state to an often deplored extent. Hence, the new standards had to be implemented in each state separately, albeit the general structure had been accepted by the Conference of Ministers of Education (KMK). The implementation processes varied, and so did the results. Therefore only a general overview can be given here.

Table 4 Exemplary task from the standards for chemistry (cited from KMK 2004c, 26-28 in this author's translation

ALCOHOLIC FERMENTATION

Stem:

In the alcoholic fermentation, yeast decomposes glucose in aqueous solution in the absence of oxygen to carbondioxide and ethanol. During a fermentation process at room temperature a group of students found the following measurements:

Time in min	60	120	180	240	300	360
Volume of carbondioxide in mL	2,3	4,6	9,2	18,4	36,8	73,6

The material depicted below plus several glass pipes, corks, tubes, and support material were available to the students, also an aqueous solution of glucose, an aqueous suspension of yeast and water.

Tasks:

1 Draw a graph of an experimental set-up, with which the students might have found these results.
2 Write the reaction equation for the alcoholic fermentation in words using molecular formulas. Name the reaction conditions for the formation of alcohol.
3 Demonstrate the formation of carbondioxide as a function of time in a graph. Use your own words to describe the quantitative correlation between time and carbondioxide formation.
4 Industry produces ethanol from sugary solutions for various purposes. Choose and argue for a separation process with which you could gain pure ethanol from the reaction mixture. Describe the process with a graph.

Task	Expected performance	level	areas of competence			
			K	E.	C	J
1	**Experimental set-up** Production of a labeled graph of the experimental set-up for the quantitative pneumatic capture of carbondioxide	I	1.1	2	2	
2	**Reaction equation** word equation: glucose -> ethanol + carbondioxide formula equation: $C_6H_{12}O_6$ -> 2 C_2H_5OH + 2 CO_2 **Reaktionsbedingung** aqueous solution of glucose, exclusion of oxygen (anaerobic), room temperature, yeast	I/II	3.1 3.4		2 4	3
3	**Evaluation** graphic representation: (labels, suitable scaling), data entered correctly verbal description: the more time, the more carbondioxide doubling of volume per time unit	I/II		6 7	2 4 8	2
4	**Separation procedure** e.g. distillation, reverse osmosis; justified with physical/chemical properties such as boiling points, concentration gradient	III	1.1 2.2	7	4 8	3

Thematic context	Competence areas			
	Subject knowledge	Epistemology methodology	communication	judgment
	Students …	Students …	Students …	Students …
• When air is thick pollutants in the air • London, Los Angeles, Bejing summer smog and winter smog • ozone – filter for life ozone and radiation • Nothing goes without water water as solvent, transport medium, raw material • Dirty sludge or crystal clear water purification und sewage treatment • Water – an element? anlaysis and synthesis of water • The lightweight among gases Hydrogen	MP[1].. explain separation processes with their knowledge of properties. MP … distinguish materials by measurable properties. MP … describe the submicroscopic structure of matter with appropriate models. SP … describe aggregate state on particle level. CR … describe that chemical reactions are generally reversible. CR … describe examples of stoffkreisläufe in nature and technology as systems of chemical reactions. EN … describe that in chemical reactions systems exchange energy with their environment. EN … describe the effect of katalysts on activation energy.	… plan and conduct experiment, attending safety considerations. … use test reactions for identification of substances. … accept the importance of accurate protocols of their observations … describe aberrations of measurements and interpret them appropriately. … show examples of links between laboratory experiments and everyday phenomena.	… use various sources of information. … choose appropriate form of protocol (text, table, graph) for their findings. … distinguish between everyday language and technical language in describing their findings. … present their findings in various modes of representation … communicate their findings properly addressee related.	… realize that chemical reactions take place all around them. … describe the relevance of chemical reactions in natural and technical processes. … explain environmental issues from the point of view of mass conservation. … judge reports in the media regarding their technical correctness. … find connections between their findings and knowledge acquired in biology and physics. … show examples of energy transfer processes in their everyday environment.

[1] Reference to the basic concepts: MP = matter - particle, SP = structure – properties, CR = chemical reaction, EN = energy

Table 5 Exemplary page from school curriculum (cited from Tausch & Wachtendonk 2011) in this author's translation)

The standards as issued by the KMK are written in rather general terms, notwithstanding the exemplary tasks that came along with them. To bring the new approach into schools, the most important step therefore was to develop curricula, which showed what content might be used to develop an understanding of the basic concepts, and how teaching should be organized to support a cumulative development of competences in the four areas.

In the core curriculum for the state of Lower Saxony, for example, the four areas of competence are assigned to the basic concepts. Within this frame, the subject teachers of a school would then design the school curriculum on the basis of which the students are finally taught. Obligatory content is defined that is used to deal with basic concepts and to develop competences. The subject content is assigned to contexts of everyday, technological, environmental or societal relevance in order to help students to better understand why learning of science might be meaningful. (For the relevance of context-based learning see Nentwig & Waddington 2005.) A page from a school curriculum might look like shown in Table 5. Following such lines, teachers are now expected to teach.

5 Evaluation of standards

The standards as issued by the KMK were based on a normative understanding of learning outcomes. The authors - expert teachers and science educators - thought that these were the competences average students would have acquired by the end of compulsory school at age 15 to 16. As the overarching goal of setting standards was the quality assurance of the educational endeavor in the school system, this expectation would, of course, have to be evaluated. Turning away from syllabus driven input-orientation the question was no longer “What was taught in school by teachers?” but rather “What are students’ learning outcomes?” (cf. Bernholt et al. 2012)

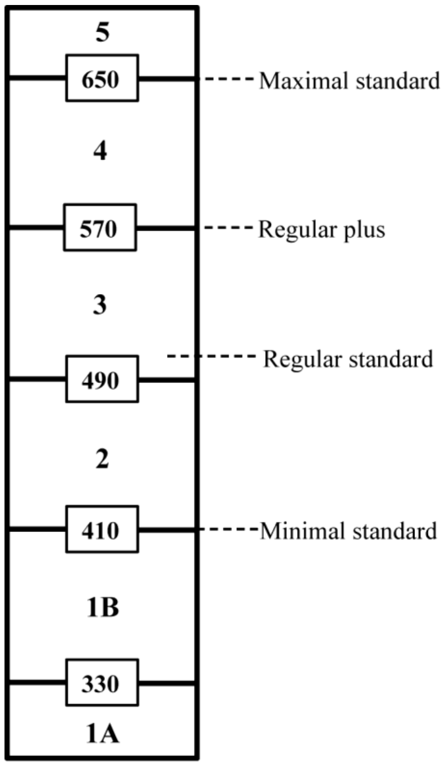


Figure 3 Tiered model of competence levels for mathematics (cited from IQB 2008 in this author’s adaptation)

In order to measure these outcomes, the standards had to be operationalized in far more detail than the original documents provided. The role of evaluator was given to the newly founded Institute for Educational Progress (IQB), who in turn engaged other agents in the respective subject areas for the development of suitable tests. The evaluation was linked to the PISA sequence, i.e. language first, then mathematics and finally science. For language, the first round of evaluation has already been completed and reported (Koeller et al. 2010).

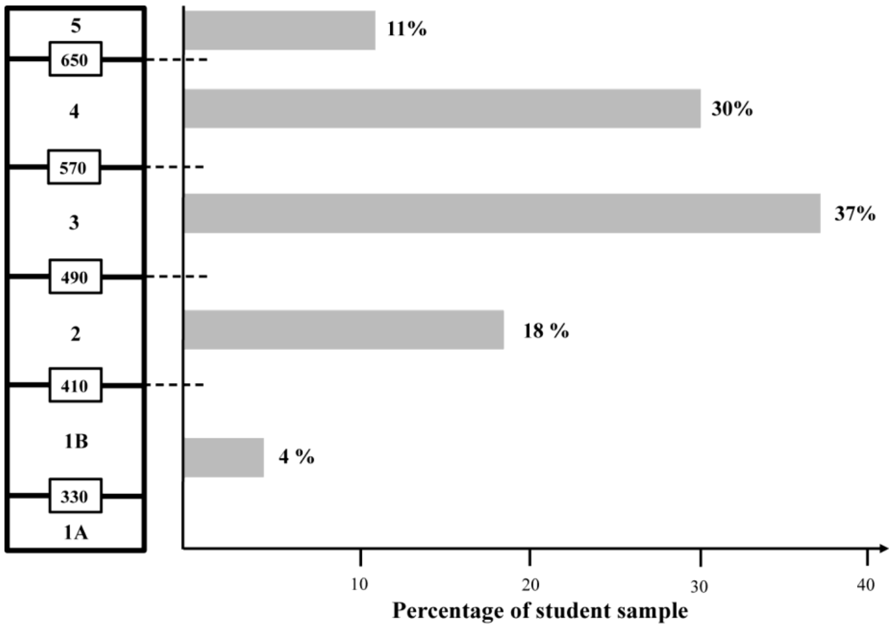
As of now, the process of evaluating and reporting the science data is not yet completed. The report is expected around the turn of the year 2012/13. In that report, the data for mathematics will be included. There have, however, been *scaling reports for math* that indicate which direction the evaluation takes.

Following largely the PISA procedure, a representative sample of well over 10000 students in year 9 and 10 were tested. Their performance with the mathematics items was entered into a scale with a defined mean of 500 points. While the initial documents proposed three levels of attainment, now five levels of equal width were defined post-hoc from the data (see figure 3). On this scale, easy to solve items have values below 400 scale points, difficult ones score above 600.

Performances below 410 scale points are considered insufficient. This level is defined as a minimal standard. It is seen as a crucial threshold below which a person would not be enabled to reasonably participate in societal and professional life. Students below this level would need special support.

The regular standard implies that not everyone is expected to reach this level (above 490 scale points), while some others will overachieve. "Regular plus" is a level which ambitious teachers should strive for with as many students as possible, at the same time keeping in mind the minority of talented students with potential for the highest level, the maximum standard. Figure 4 shows the distribution of the student sample on the scale of competence levels.

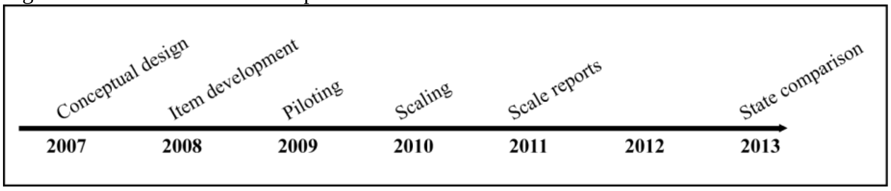
Figure 4 Competence distribution of students in math (cited from IQB 2008 in this author's adaptation)



In similar ways, a tiered competence model can be expected for science. The evaluation for the science subjects is in the hands of a group of science education researchers (ESNaS) in cooperation with the Institute for Educational Progress (IQB) (cf. Kauertz et al. 2010). Figure 5 shows the timeline for the evaluation of the science standards.

Several groups of experienced teachers assisted by science education experts and advised by experts for psychometry devised several hundred science items, basically following the PISA scheme: one item stem introducing the context and

Figure 5 Time line for the development of tests for the evaluation of science standards



giving some basic information followed by a number of questions about the context.
The peculiarity of these items was that they were aiming at competences as described above rather than at traditional content knowledge. These items were piloted, partially revised and scaled under the guidance of the IQB. To make them meaningful from a science education perspective, valid from a curricular viewpoint and reliable and sufficiently selective was an obvious challenge. For this round of evaluation the emphasis is on the competence areas *subject knowledge* and *epistemology/methodology*. The areas of communication and judgment have been postponed until PISA's next science focus in 2018. As of today, the report is yet to be published.

6 Concluding considerations

The introduction of standards to the educational system in Germany mainly served two purposes: the enhancement of *educational quality* and the *harmonization* between states.

When international comparative studies such as TIMSS and PISA proved that education in Germany was not as effective as had been believed by most, politicians felt the country's competitiveness endangered. Science and technology in particular were seen as problematic with little popularity among students and dissatisfying enrollments for further education in that field. Hence the setting of standards and the close monitoring of learning outcomes are expected to improve the situation.

When PISA results were scrutinized for the different states in the federal republic, alarming differences were detected. While some states could compete with the top in the PISA ranks, others performed far below the average. Standards combined with national assessment should therefore help to level the differences. Certainly this problem is far too complex to be solved with standards alone. Issues of teacher training, curriculum, school structure, facilities and support need to be considered. This is still work in progress.

Not only were standards and the focus on educational outcome rather than input new to the educational system in Germany, a whole new category was introduced with them: *competence*. Although the term had been widely in use with a multitude of connotations, it was now specified for the results of students' learning.

While "subject knowledge" and "epistemology/methodology" had been familiar as knowledge of and knowledge about science, the inclusion of communication and judgment as expected competences was new to most science teachers. The requested changes were an enormous challenge. Documents were produced by the states' ministries of education, and in-service training courses were offered in various degrees of quantity and quality. Research has yet to show to what extent the change has reached classrooms.

Germany was not alone in introducing competence as a category of learning outcome. Around the same time as in Germany, the concept appeared in educational documents in many European countries (cf. Waddington et al. 2007). The German speaking countries Austria and Switzerland, in particular, followed the German notion of the concept, basing their considerations on the same report of Klieme et al. (2003).

Nationwide *assessment* was a controversial issue. It was argued that teachers' fear of external scrutiny might stifle educational creativity and innovation. In this case, the reform was the educational innovation. Looking at the standards should compel responsible teachers to adjust their classroom practice in order to enhance competence development. Thus, external evaluation might be seen as a chance rather than a threat.

"Teaching-to-the-test", in other cases the opponents' battlecry against centralized assessment, was advisable in this context. If the assessment is looking at hitherto uncommon competences and at content newly structured by basic concepts, then teachers are expected to teach to the test. Teaching and assessing need to be aligned in order to make the change valuable. How far this process has grown is yet to be seen

Certainly, assessing competences is a challenge. Valid, reliable, and feasible instruments need to be developed to assess students' science competencies rather than just their content knowledge. Especially "soft" competences like communication and judgment are difficult to grasp in feasible large scale procedures. Not without reason has their assessment been postponed until PISA comes around to science the next time.

In some countries with a tradition of setting standards, nationwide assessment has led to what some see as debatable use of data (e.g. league tables, high-stakes examination). In Germany the authorities have confirmed that with each round of evaluation only a representative sample will be tested. Thus, the results allow no judgment below the system level. States can be compared (for the purpose of harmonization), as they provide the student samples, but not schools and absolutely not individual students.

One strongly debated issue in the implementation of standards was the curriculum or rather its absence. The standards were issued first, and the alignment of the curriculum was left to the states, most of whom passed the task on to schools and eventually to teachers. While teachers indispensably have to be involved in the curriculum reform, not few felt overburdened with this task and not sufficiently valued as partners in this process. Among parts of the teaching force there was a tendency to sit and wait for this reform to pass – as others had in the past.

This might be the most critical issue in the implementation of standards to the educational system in Germany. They were served from the top down. Within a relatively short time they were made compulsory. Curricula still had to be developed, and not everyone was enthusiastic about the concept of competence as a goal for teaching. Whether the standards, as defined in the initial documents, will be reached by the average student will be seen when the evaluation report has been published.

In Switzerland, although the same understanding of standards was applied based on the so called Klieme report, the process was reversed. The concept was elaborated, tasks were tested, a competence model was validated (cf. Labudde 2007, Labudde 2008, Ramseier et al. 2011). Teachers were involved all along. The result was finally subjected to political hearings, and only then did the Swiss ministers of education release the standards as a framework for curriculum development and national monitoring (EDK 2011). This lack of haste might have suited the development in Germany as well.

References

Aufschnaiter v., C. & Rogge, C. (2010). Wie lassen sich Verläufe der Entwicklung von Kompetenz modellieren? (How can the development of competences be described?). Zeitschrift für Didaktik der Naturwissenschaften, 16, 95–114.

Beaton, A. E., Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1996). Science Achievement in the Middle School Years: IEA's third International Mathematics and Science Study (TIMSS). Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College

Bernholt, S., Neumann, K. & Nentwig, P. (Eds.) (2012). Making it tangible. Learning outcomes in science education. Münster: Waxmann

Bernholt, S., Parchmann, I. & Commons, M. (2009). Kompetenzmodellierung zwischen Forschung und Unterrichtspraxis (Modelling Scientific Competence between Research and Teaching Practice). Zeitschrift für Didaktik der Naturwissenschaften, 15, 219-245

BLK (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung) (Ed.) (1997). Gutachten zur Vorbereitung des Programms "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts". Materialien zur Bildungsplanung und Forschungsförderung, Heft 60. Bonn: BLK. http://www.blk-bonn.de/papers/heft60.pdf. Accessed 20 September 2012

DeBoer, G. (2011). The Globalization of Science Education. Journal of Research in Science Teaching 48 (6) 567-591

EDK (2011). Grundkompetenzen für die Naturwissenschaften. Nationale Bildungsstandards. Freigegeben von der EDK-Plenarversammlung am 16.Juni 2011. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK). http://edudoc.ch/record/96787/files/grundkomp_nawi_d.pdf. Accessed 20 September 2012.

Glemnitz, I. (2007). Vertikale Vernetzung im Chemieunterricht: Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext. Berlin: Logos

Hammann, M. (2004). Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung (Competence development models: characteristics and relevance). Der mathematische und naturwissenschaftliche Unterricht (MNU) 57(4), 196-203.

Hostenbach, J., Fischer, H., Kauertz, A., Mayer, J., Sumfleth, E. & Walpuski, M. (2011). Modellierung der Bewertungskompetenz in den Naturwissenschaften zur Evaluation der Ntionalen Bildungsstandards (Modeling the evaluation and judgment competence in science to evaluate national educational standards). Zeitschrift für Didaktik der Naturwissenschaften, 17, 261–288.

IQB (2012a). Kompetenzstufenmodell zu den Bildungsstandards im Kompetenzbereich Sprechen und Zuhören – hier Zuhören – für den Mittleren Schulabschluss. http://www.iqb.hu-berlin.de/bista/dateien/Deutsch_KSM_Hre_2.pdf. Accessed 20 September 2012.

IQB (2012b). Kompetenzstufenmodell zu den Bildungsstandards im Kompetenzbereich Lesen für den Mittleren Schulabschluss. http://www.iqb.hu-berlin.de/bista/dateien/Deutsch_KSM_Les_2.pdf. Accessed 20 September 2012.

IQB (2012c). Kompetenzstufenmodell zu den Bildungsstandards im Kompetenzbereich Rechtschreiben. www.iqb.hu-berlin.de/bista/dateien/Deutsch_KSM_Ort.pdf. Accessed 20 September 2012.

IQB (2012d). Kompetenzstufenmodelle zu den Bildungsstandards im Fach Französisch für den Mittleren Schulabschluss - Hörverstehen und Leseverstehen. http://www.iqb.hu-berlin.de/bista/dateien/Kompetenzen_Fra_v2.pdf. Accessed 20 September 2012.

IQB (2012e). Kompetenzstufenmodelle zu den Bildungsstandards im Fach Englisch für den Mittleren Schulabschluss - Hörverstehen und Leseverstehen. http://www.iqb.hu-berlin.de/bista/dateien/Kompetenzen_Englisch_MSA.pdf. Accessed 20 September 2012.

IQB (2012f). Kompetenzstufenmodell zu den Bildungsstandards für den Hauptschulabschluss und den Mittleren Schulabschluss im Fach Mathematik. www.iqb.hu-berlin.de/bista/dateien/Kompetenzstufen_1.pdf. Accessed 20 September 2012.

Kauertz, A., Fischer, H. E., Mayer, J., Sumfleth, E., & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe I. (Standard-related modelling of competence in science for lower secondary school). Zeitschrift für Didaktik der Naturwissenschaften, 16, 135–153.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., et al. (2003). Zur Entwicklung nationaler Bildungsstandards [About the development of National Education Standards]. Berlin: Bundesministerium für Bildung und Forschung.

Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG [Models of Competences for Assessment of Individual Learning Outcomes and the Evaluation of Educational Processes. Description of a new DFG priority research program]. Zeitschrift für Pädagogik, 52(6), 876-903.

Kulgemeyer, C. & Schecker, H. (2009). Kommunikationskompetenz in der Physik: Zur Entwicklung eines domänenspezifischen Kommunikationsbegriffs (Physics Communication Competence: on the development of a domain-specific concept of communication). Zeitschrift für Didaktik der Naturwissenschaften, 15, 131–153.

KMK (2003a). Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mittleren-SA.pdf. Accessed 20 September 2012

KMK (2003b). Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-Deutsch-MS.pdf. Accessed 20 September 2012

KMK (2003c). Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mathe-Mittleren-SA.pdf. Accessed 20 September 2012

KMK (2003d). Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-BS-erste-Fremdsprache.pdf. Accessed 20 September 2012

KMK (2004a). Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss für die Fächer Biologie, Chemie, Physik. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Mittleren-SA-Bio-Che-Phy.pdf. Accessed 20 September 2012

KMK (2004b). Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf. Accessed 20 September 2012

KMK (2004c). Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Chemie.pdf. Accessed 20 September 2012

KMK (2004d). Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Physik-Mittleren-SA.pdf. Accessed 20 September 2012

Köller, O., Knigge, M. & Tesch, B. (Hrsg.) (2010). Sprachliche Kompetenzen im Ländervergleich (Language competence in state comparison). Münster: Waxmann

Labudde, P. (2007). How to Develop, Implement and Assess Standards in Science Education? 12 Challenges from a Swiss Perspective. In D. Waddington, P. Nentwig, & S. Schanze (Hrsg.*)*, Making it Comparable: Standards in Science Education. Münster, New York, München, Berlin: Waxmann, S. 277–301.

Labudde, P. (2008). Developing and Implementing New National Standards in Science Education: The Role of Science Educators. In Ralle, Bernd & Eilks, Ingo (Hrsg.), Didaktik der Naturwissenschaften: Promoting Successful Science Education - The Worth of Science Education Research. Aachen: Shaker Verlag, S. 63–74.

McClelland, D. C. (1973). Testing for competence rather than for intelligence. American Psychologist, 28, 1-14.

Nentwig, P. (2009). Damit es nicht Stückwerk bleibt. Horizontale und vertikale Vernetzung am Beispiel von Chemie im Kontext. In A. Feindt, V. Elsenbast, P. Schreiner, A. Schöll (Hrsg.): *Kompetenzorientierung im Religionsunterricht*. Münster: Waxmann, S. 197-209.

Nentwig, P. & Waddington, D. (Eds.) (2005). Making it relevant. Context based learning of science. Münster: Waxmann.

Nentwig, P. & Schanze, S. (2007). Making it comparable - Standards in science education. In D. Waddington, P. Nentwig & S. Schanze (Eds.), Making it comparable - Standards in science education (pp. 11-19). Münster: Waxmann.

Neumann, I. (2011). Beyond Physics Content Knowledge. Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge. Berlin: Logos.

Neumann, K., Viering, T. & Fischer, H. (2010). Die Entwicklung physikalischer Kompetenz am Beispiel des Energiekonzepts (The development of physical competence with the example of the energy concept). Zeitschrift für Didaktik der Naturwissenschaften, 16, 285–298.

Niedersächsisches Kultusministerium (Hrg.) (2007). Kerncurriculum für das Gymnasium Schuljahrgänge 5 bis 10 – Naturwissenschaften. http://db2.nibis.de/1db/cuvo/datei/kc_gym_nws_07_nib.pdf. Accessed 20 September 2012.

OECD (2001). Knowledge and skills for life - First results from the OECD programme for international student assessment (PISA) 2000. Paris: OECD.

OECD (2004). Learning for Tomorrow's World. First Results from PISA 2003. Paris: OECD.

OECD (2006). Assessing Scientific, Reading and Mathematical Literacy. Paris: OECD

Ramseier, E., Labudde, P. & Adamina, M. (2011). Validierung des Kompetenzmodells HarmoS Naturwissenschaften: Fazite und Defizit. (Validation of the HarmoS Science Competeny Model: Results and Deficiencies). Zeitschrift für Didaktik der Naturwissenschaften 17, 7-33

Schanze, S. & Nentwig, P. (2008). Standards im Naturwissenschaftlichen Unterricht – ein internationaler Vergleich (Standards in Science Education – an International Comparison). Zeitschrift für Didaktik der Naturwissenschaften 14, 125-143.
Tausch, M. & von Wachtendonk M. (Hrsg.) (2011) Synopse zum Kerncurriculum. http://www.ccbuchner.de/downloads/308/Chemie_Synopse_NI_1.pdf. Accessed 20 September 2012
TIMSS International Study Center (1998). TIMSS Highlights. http://timssandpirls.bc.edu/timss1995i/TIMSSPDF/C_Hilite.pdf. Accessed 20 September 2012
Waddington, D., Nentwig, P., & Schanze, S. (2007). Standards in science education. Münster: Waxmann.
Wadouh, J. (2007). Vertikale Vernetzung und kumulatives Lernen im Biologieunterricht. Dissertation. Essen: Universität Duisburg-Essen. http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-20492/DISSERTATION_komplett_5_BIB_end.pdf. Accessed 20 September 2012.
Weinert, F. E. (1999). Concepts of Competence. Contribution within the OECD-Project Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo). Neuchâtel: Bundesamt für Statistik.
Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), Defining and selecting key competencies (pp. pp. 45-65). Seattle, WA: Hogrefe & Huber.