Evaluation of students' practical performance in science.

Christoph Gut
University of Applied Sciences Northwestern Switzerland, School of Teacher Education, Zentrum für Naturwissenschafts- und Technikdidaktik, CH-4058 Basel, Switzerland
christoph.gut@phbern.ch (till 29th of February 2008)
christoph.gut@fhnw.ch (from 1st of March 2008)

Supervisor: Prof. Dr. Peter Labudde, University of Applied Sciences Northwestern Switzerland, CH-4058 Basel

Background

In 2004, the Swiss Conference of Cantonal Ministers of Education (EDK, Erziehungsdirektorenkonferenz) initiated the project HarmoS (Harmonisierung obligatorische Schule: Harmonization of compulsory school). HarmoS intends to harmonize and monitor the educational landscape in Switzerland (EDK 2004a, 2004b), which is divided into 26 cantons, each having its own educational system until now. The project relies on performance standards, i.e. achievement standards for students of grades 2, 6 and 9 in four subjects, namely mathematics, science, first language, and second language. The standards are handled as minimal or so-called basic standards and are built on competencies as defined by Weinert (2001): Competencies are "cognitive abilities and skills possessed by or able to be learned by individuals that enable them to solve particular problems, as well as the motivational, volitional and social readiness and capacity to utilise the solutions successfully and responsibly in variable situations".

The Swiss consortium HarmoS Science – the sub-group committed to science education – is working on a competency model that includes three dimensions: competencies („Handlungsaspekte"), domains of contents („Themenfelder"), and levels of understanding („Niveaus"). „The competencies include skills such as identifying and questioning scientific problems, developing preliminary ideas and plans, carrying out a plan of action, manipulating materials, as well as processing, interpreting and communicating data. The domains of contents are within physical or living systems, e.g. structure and properties of matter, energy and its transformation, the human body/health/well-being, ecosystems, plants and animals. These

domains are rather interdisciplinary, more like a STS-approach than driven by the systematics of biology, chemistry, and physics." (Labudde 2007). The level dimension remains yet a raw construction prior to assessment.

The competency model for science performance will be evaluated in a nationwide assessment, including paper-and-pencil-tests and experimental tests for the 2nd, 6th, and 9th year level in the German and French part of Switzerland. All tests took place in 2007 already, except for the experimental tests for the 6 and 9 grade, which are planned for spring 2008. Until the end of autumn 2008, the HarmoS consortium will propose basic standards based on the results of the assessment to the political authorities, i.e. the Swiss Conference of Cantonal Ministers of Education (EDK), who will then discuss, approve, and implement the standards by law.

An essential feature of the HarmoS science project is the so-call *science performance test*, i.e. tests that require actual performance in lab-work and hands-on activities. Such testing is necessary for assessing and monitoring students' practical skills. The results of the performance assessment will serve two goals. Firstly, the assessment evaluates the HarmoS competency model that includes practical competencies in science laboratory. A fine-tuning of the model will be done after the statistical processing and interpretation of the collected data. Secondly, on the base of these results, standards for skills in science education will be formulated.

In the Swiss educational systems, there is no tradition in assessing performance by practical tests. Much effort is made by the consortium HarmoS science at the moment to build up know-how and practical test materials that can serve to install regular performance assessments in Switzerland.

Context of the doctoral thesis

My thesis project started in August 2007 and will presumably be finished in 2010. As a member of the consortium HarmoS science, I am involved in the development of experimental problems, the organisation and implementation of the performance test for the 9th graders in spring 2008 and the analysis of the results by October 2008. After than, research and development on the base of the collected data by the HarmoS project are planed.

Research questions and their value for the scientific community

Standards have become a necessary tool of political authorities and educational researchers in

developing and monitoring education systems. In order to represent the reality of the schools, standards must be formulated on the base of empirically evaluated competency models. Therefore, the development and evaluation of appropriate competency models should be in the focus of today's educational research. Although different competency models for students' practical performance in science have been developed, little research has been done so far in evaluating these models by performance tests. For Switzerland, the first step will be done with the completion of the HarmoS project. Further development both of the competency model and the related basic standards is of great political and scientific interest. Therefore, the following research questions are adressed by my doctoral thesis:

1. Which competencies – adapted to the challenges of the 21st century – could and should be defined in the field of experimental work in science?
2. In which way should the given HarmoS competency model be refined or re-defined (with regard to special competencies, such as practical laboratory skills or competencies concerning exploring tasks) on the base of the data collected in the HarmoS science project?
3. How can these re-tuned competencies be evaluated?
4. To which degree may competency models different to the HarmoS science competency model be evaluated by the data collected in the HarmoS science project?

Methodology

Ad 1.: The development of the HarmoS science competency model takes account of different normative and empirical inputs. The model is based on an analysis of all 26 cantonal science curricula (Szlovák 2005), and is strongly influenced by results of science education research, by international curricula, by frameworks and concepts such as the PISA notion of scientific literacy (OECD 2007), the Canadian Common Framework of Science Learning Outcomes (CMEC 1997), the US standards (National Research Council 1996) and the new German standards in science education (KMK 2005) among others.

Ad 2.: The competency model for science performance is being evaluated in a nationwide assessment. In spring 2007, a sample of over 9000 students of grades 6 and 9 from the German and the French speaking part of Switzerland participated in a paper-and-pencil-test. A P&P-test together with an experimental test for the 2nd grade level followed in autumn 2007. An experimental test with about 1200 students of the 6th and 9th grade will take place in spring

2008. In the 9th grade (and similarly in grade 6), each student will be assessed in two consecutive 60-minutes-periods. In order to connect the results of the experimental test to the national P&P-test in 2007, the students are solving six P&P-tasks in the first period. During the second period they accomplish two experimental tasks. For the evaluation, the difficulty of all task items is analyzed by Rasch-models. Further statistical analysis will allow to describe the performance of the students in detail. On the base of theses results, the competency model will be re-tuned in such a way, that the redefined competencies become statistically discriminatory.

Ad 3.: For further research, I plan to focus on competencies paradigmatic to practical laboratory work in physics education, in particular to ones that little have been investigated so far, such as skills concerning exploring, inventing, and developing experiments. A refined model of theses competencies is to be developed and evaluated by an appropriate experimental test, which I plan to develop for the 9th grade level.

Ad 4.: Using the results of the Harmos experimental tests an evaluation of different science competency models (see Waddington et al. (2007), Hamann (2004), Bybee (1997, 2002) among others), may be accomplished by re-interpreting the test items according to the different competency systems. As a result, a comparison of different models on statistical grounds may be achieved.

References

Bybee, R. W. (1997), "Toward an Understanding of Scientific Literacy", In: Gräber, W.; Bolte, C. (Eds.), *Scientific Literacy*, Institut für Pädagogik der Naturwissenschaften, Kiel

Bybee, R. W. (2002), "Scientific Literacy – Mythos oder Realität", In: Gräber, W.; Nentwig, P.; Koballa, T.; Evans, R., *Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*, Leske + Budrich, Opladen

CMEC (1997), *Common Framework of Science Outcomes*, Counsil of Ministers of Education, Canada (CMEC): http://www.cmec.ca/science/framework/

EDK (2004a), *HarmoS – Zielsetzungen und Konzeption*. Erziehungsdirketorenkonferenz (EDK), Bern: http://www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html

EDK (2004b), *Projekt HarmoS – Offertenausschreibung für Entwicklung von Konpetenzmodellen,* Erziehungsdirketorenkonferenz (EDK), Bern: http://www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html

Hamann, M. (2004), *Kompetenzentwicklungsmodelle*, MNU, **57**(4), 196-203

KMK (2005), *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*, Kultusminsterkonferenz (KMK), Kluwer, München

Labudde, P. (2007), "How to Develop, Implement and Assess Standards in Science Education? 12 Challenges from a Swiss Perspective", In: Waddington, D.; Nentwig, P.; Schanze, S. (Eds.), *Making it comparable - Standards in Science Education*, Waxmann, Münster 2007

National Research Counsil (1996). *National Science Education Standards*, National Academic Press, Washington D.C.

OECD (2007), *PISA 2006 Science Competencies for Tomorrow's Wolrd*, Vol. I and II, Organisation for Economic Co-operation and Development (OECD), Paris

Szlovák, B. (2005), *HarmoS – Lehrplanvergleich Naturwissenschaften*, EDK, Bern

Waddington, D.; Nentwig, P.; Schanze, S. (Eds.) (2007), *Making it comparable - Standards in Science Education*, Waxmann, Münster

Weinert, F.E. (2001), "Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit", In: Weinert, F.E. (Ed.), *Leistungsmessungen in Schulen*, Beltz Verlag, Weinheim, 17-31

Synopsis: ESERA Summerschool 2008, York, UK

31.01.08/ Christoph Gut