



HAL
open science

The VEMINT-Test: Underlying Design Principles and Empirical Validation

Reinhard Hochmuth, Marcel Schaub, Andreas Seifert, Regina Bruder, Rolf Biehler

► **To cite this version:**

Reinhard Hochmuth, Marcel Schaub, Andreas Seifert, Regina Bruder, Rolf Biehler. The VEMINT-Test: Underlying Design Principles and Empirical Validation. Eleventh Congress of the European Society for Research in Mathematics Education, Utrecht University, Feb 2019, Utrecht, Netherlands. hal-02422641

HAL Id: hal-02422641

<https://hal.science/hal-02422641>

Submitted on 22 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The VEMINT-Test: Underlying Design Principles and Empirical Validation

Reinhard Hochmuth¹, Marcel Schaub², Andreas Seifert³, Regina Bruder² and Rolf Biehler³

¹Leibniz University Hannover, Germany ²Technical University of Darmstadt, Germany ³University of Paderborn, Germany; hochmuth@idmp.uni-hannover.de

VEMINT (Virtual Entrance Tutorial for the STEM subjects), a joint project by researchers from the universities of Darmstadt, Kassel, Paderborn and Hannover, has worked on material and course-designs for bridging courses for more than 15 years. In recent years, VEMINT has focused on the development of a psychometrically validated test to provide students with appropriate hints for the use of material and to evaluate the impact of teaching and learning activities. In this paper, we present elements of the test, which are based on ideas from Feldt-Caesar (2017), who combined cognitive activity and action control theoretical approaches in the development of a three-dimensional competence structure model, and on Anthropological Theory of Didactics, to explain content-related differences between dimensions of the model. Validating Rasch-analyses show that, in accordance with the design ideas, a three-dimensional Rasch-model fits with empirical data.

Keywords: Bridging course, diagnosis, transition, calculus, Rasch-Analysis

Introduction

The average dropout rates of mathematics and engineering programs in universities in Germany are 47% and 36%, respectively (Bildungsberichterstattung, 2014). These rates are higher than the overall average drop-out rate which is 33% for all study programs at German universities. In addition to the lack of acquisition of key competences such as the ability and willingness for self-reflection and self-motivation (Hilgert, 2016), the literature (e.g., Cramer & Walcher, 2010) highlights that the lack of prior mathematical school knowledge is one of the major reasons for the high proportion of dropouts.

As a measure to reduce the number of dropouts, most universities and higher education institutes have offered mathematical remedial activities in the form of bridging courses to support students' learning of university mathematics in recent years. The idea is to reduce the gap between school and university mathematics while maintaining the standards of the programs. One of the major challenges is to design remedial bridging courses that are sensitive to individual needs. For example, while some of the students need more support in simplifications of algebraic expressions, other students may need support in elementary calculus. It is also important for the students to know what prior knowledge is needed to take the courses at the university. To support students in deciding which of the modules of the bridging course to take, it is desirable to have an appropriate diagnostic tool available. Such tool should also measure the impact of bridging courses. A substantial content analysis is essential in order to provide detailed feedback to individual learners and advice to lecturers in choosing topics and in deciding how intensely topics should be taught in transition courses (Biehler & Hochmuth, 2017; Hochmuth, 2018). For more than 15 years, the universities of Darmstadt, and recently also Hannover, Kassel, and Paderborn, have been offering bridging courses in the cooperative project Virtual Entrance Tutorial for STEM subjects (VEMINT,

STEM in German is MINT) for students who wish to pursue mathematics-related programs (Bausch et al., 2014; Fischer, 2014). In recent years, the authors have also worked on developing a diagnostic test, the VEMINT-test, which is the focus of this paper.

The structure of the paper is as follows. After sketching the main goals of the VEMINT-test, we outline basic ideas of the competence structure model that underlies our test design and provide a briefpraxeological content analysis of tasks assigned to the dimensions of the model. Finally, we present results from a Rasch-Analysis empirically validating the three dimensions of the competence structure model.

The Development of the VEMINT-Test

The development of the VEMINT-test follows two partly opposing, goals: While working as a diagnostic test to understand better students' prior knowledge and gaps, it should also help to evaluate the impact of bridging courses. In addition, the pre-test needs to give information in a short time so students can decide what modules they need to take in the bridging course. In recent years, the diagnostic test has been refined several times based on both empirical and theoretical findings. Thereby, our main aim was to develop a standardized diagnostic instrument in two essentially equivalent versions (Form A and Form B), which can be used every year at the beginning and at the end of a bridging course. An expert rating determined the particular content of the test, which focused on analysis and elementary algebra and took into account the German minimum standard requirements (see, Cooperation Schule Hochschule, 2014). The items were adjusted according to quality criteria for tests and technical possibilities (e.g., the use of the Moodle STACK plug-in, Sangwin, 2007).

The final test in two forms (A and B) was designed for 60 minutes and comprised 22 tasks from secondary levels I and II in German schools. The examined topics are algebraic terms, equations, functions and calculus (differentiation and integration). The test was developed using a three-dimensional competence structure model that will be described next. Afterwards we present results from a Rasch-Analysis, considering data that has been collected in the winter semester of 2017/2018. These results empirically reproduce the three-dimensional structure, thus validating our test-design in this respect.

A Competence Structure Model and its Dimensions

The competence structure model is based on the idea of "Basic Mathematical Knowledge" (BMK), which has been introduced by Feldt-Caesar (2017) as follows:

Basic mathematical knowledge and skills refer to the mathematical knowledge, skills and abilities which are to be retrievable to all pupils at the end of the two secondary levels in the form of mathematical terms, theorems and procedures in the long term and independently of the situation, i.e. in particular without the use of aids. (Our translation, p. 182)

BMK serves "as a prerequisite for successful further learning, especially in a course of study" (our translation, p. 180), which requires that knowledge be sustainable and connected in the sense of "intelligent knowledge" (Feldt-Caesar, 2017; Weinert, 2000) and, furthermore, transferable and connectable with new content (Feldt-Caesar, 2017; Nitsch Bruder & Kelava, 2016). For related

diagnostic and feedback goals these ideas imply that items of a test should cover not only calculus-oriented tasks but should also aim at “intelligent knowledge” and “understanding oriented knowledge” where the latter is not the same as “understanding.” We agree with Skemp (1976) that no “understanding,” neither instrumental nor relational understanding, is validly testable without additional intersubjective interactions. On the other hand, “understanding oriented knowledge” including knowledge necessary for “relational understanding” beyond knowledge necessary for “instrumental knowledge” is an important basis for further learning (Bruder, Feldt-Caesar, Pallack, Pinkernell & Wynands, 2015; Skemp, 1976). These general underlying goals call for a sample of tasks which cover various qualitative facets of knowledge and associated mental activities. Moreover the sample should be helpful both in regards to diagnostic and evaluation goals as well as for providing appropriate feedback. To cover and operationalize the idea of such different qualitative facets we have developed a three-dimensional competence structure model on the basis of ideas from Feldt-Caesar (2017) and applied Anthropological Theory of Didactics (ATD, Chevallard, 1992)) for the content analysis of tasks.

In Feldt-Caesar (2017), cognitive, activity and action control theoretical approaches are combined to describe and analyze mental activities with respect to five conceptual levels. Within our context, the levels “elementary actions,” including identification and realization, and “basic actions,” including recognizing, describing, linking, applying and justifying, are important. In Feldt-Caesar (2017), both types of actions have been related to the so-called “content elements” that are conceptualized in terms, theorems, and procedures. In view of our focus on analyzing test tasks within the institutional context of the transition from school to university mathematics, it turned out that a praxeological view on “content elements” was probably more appropriate for the current project.

Accordingly, in the following we apply a few basic praxeological notions from ATD: The simplified [P, L]-model with P representing the praxeological block (i.e., the technical aspect(s) connected to a type of task and) with L representing the technological-theoretical block. The latter covers in particular: describing a technique, validating (e.g., proving or justifying how and that a technique works), knowledge about the efficiency and motives of a technique as well as knowledge about variants and simplifications, and finally, knowledge required or useful for controlling the application of a technique, which is particularly relevant for connecting practices and praxeologies. Our starting point for the identification of P and L lies in the single tasks of the test, curricula school books and in the VEMINT group consented experiences concerning the actual implementation and institutionalization of praxeologies and corresponding technical and technological aspects of school mathematics. There are, of course, further relations between above mentioned cognitive notions like “instrumental” and “relational understanding”: From the viewpoint of knowledge and meaning structures, they can be referred to qualitative differences in relationships between practical and technological-theoretical blocks of one or several praxeologies. A more detailed analysis of such issues, however, lies beyond the scope of this paper. Here, the praxeological notions mainly serve as a tool for describing and analyzing qualitative aspects of “content elements” with regard to the three dimensions of the competence structure model.

Dimension I (which is associated to instrumental understanding) is characterized by elementary

actions and cognitive operations such as identification and realization, and usually refer to a content element that does not have to be transferred to other content elements. The actions to be taken consist of a single step and are related to a “first elementary level” of mental activities in Feldt-Caesar (2017). Content-wise and in praxeological terms, Dimension I covers tasks from school that can be solved by techniques without especially taking into account technological aspects like supporting argumentations, proofs, or contextual embedding, etc. The techniques in Dimension I tasks are more or less extensively institutionalized in school mathematics, represent typical tasks of school examinations, and, at least partly, also appear in final exams (e.g., baccalaureates). A characteristic task in the test is the calculation of the derivative of $(3 - x^2)^6$. It can be calculated by algebraic operations that can be executed without referring to basic rules and notions. This kind of tasks, focusing on techniques without links to technologies and supporting mathematical organizations, seems to be dominant in school internationally (e.g., Barbé, Bosch, Espinoza & Gascón, 2005).

In Dimension II, the one-step approach is maintained but a linking to content elements of the same

A polynomial function f fulfils the following properties: f'' is negative for $x < 4$ and positive for $x > 4$. Which of the graphs belongs to f ?

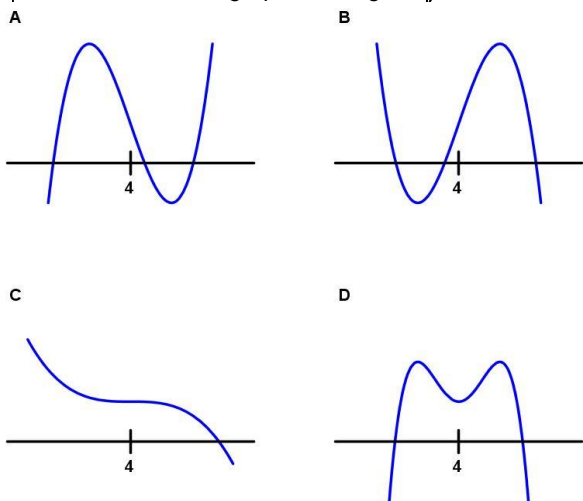


Figure 1: Example for a task of dimension II

Moreover, in the school context, this type of tasks typically demands verbal descriptions or even verbal justifications. A further common aspect of Dimension II tasks is that the technological aspects need no additional validation, for example by linking to technological aspects of praxeologies living in other domains; hence they are of a local nature. Figure 1 presents an example for a Dimension II task. It is expected that to solve this task students use relationships between verbal, graphical, and symbolic representations and switch across those representations. In order to solve the task, the second derivative is to be interpreted as the curvature of the function f . What is unusual about this task is that the values of the second derivative are given and, conversely, a function f must be reconstructed based on the given information.

Dimension III tasks are multi-step tasks, and characterized by several elementary actions and cognitive operations (associated to “intelligent knowledge”). The tasks refer to content elements

topic is necessary to solve the associated tasks. Such tasks are characterized by the main reference to one content element and possibly to other content elements, directly related to the main content element. Besides elementary actions, cognitive operations such as describing, linking, applying, and justifying are necessary in most cases. These tasks cannot be solved by simple calculations and are often unfamiliar to students (so that they need probably relational understanding). Content-wise, Dimension II covers tasks that relate to technological-theoretical blocks of praxeologies implying validation and justification. Students are not sufficiently trained in techniques of Dimension II tasks such that such technological aspects are superfluous for solving

from at least two different content areas. Thus, solving of Dimension III tasks combines techniques from different praxeologies; the techniques involved are often more complex, or the technological aspects relevant for the combination of the techniques are either weakly established or combine aspects of different praxeologies. A task of this type, that combines techniques and is well-established in school, looks as follows (Feldt-Caesar, 2017): “The function f is given by $f(x) = x^2 - 7$. Determine the equation of the tangent t , which touches the graph of the function f at the point $P(-2,1)$.” In this task, content elements of differential calculus must be linked and the results of the differential calculus must be used to establish a linear function (equation of a tangent). Thus, both content from secondary level I and secondary level II must be applied. Another task (item 13) of Dimension III is the following: “Simplify the following expression and collect the variables $\frac{x^{k-n}}{y^{2n}} : \frac{x^{2k-n}}{y^{n-1}}$ ($x, y \neq 0$).” In this task, the required combination of techniques from division of fractions and power calculations is nested; in addition, related technological aspects have to be combined in an appropriate way. This means that it is not enough to apply simple and local techniques from different domains. Content-wise, the main characteristic of Dimension III tasks is the link of different techniques and technologies, which is as such not well-institutionalized in school mathematics.¹

Rasch-Analysis

Roughly one half of 362 students from three German universities who participated in bridging-courses (29% female) filled out test-form A before and test-form B after the courses. The other half took the opposite order. No sequence effects were found. Each participant received each test-form only once, and, although pre- and posttest samples were related (both samples consisted of the same participants), the occurrence of test-effects is not plausible. Using pre- and posttest samples in one analysis reduces the probability of the occurrence (respectively the effect) of a possible limitation of variance, which can lead to a possible underestimation of all correlational measures and possibly limiting the information value of assertions regarding validity. Participants were tested before and after a bridging course and therefore were not on the same performance level at both times of measurement.

To scale the data, we first estimated two one-dimensional Rasch-models, one for each measurement point. A Rasch-model is a probabilistic psychometrical model for generating measurements from categorical data (e.g., right/wrong-answers in a performance test) as a function

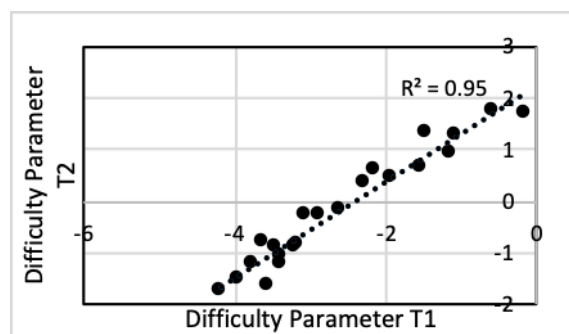


Figure 2: Scatterplot of the rank of difficulty

¹ It should be noted that assigning tasks to a single of the dimensions is not always possible due to different possible solutions of a task. For the classification, several possible solutions were anticipated and the most probable one was used to determine the dimension. Analyses of some of our test tasks can be found in Feldt-Caesar (2017), which investigates validation issues.

of the relation between the ability of the respondents and the item difficulty (Rasch, 1980). EAP/PV-reliabilities were suitable, $R = .868$ (resp. $.865$ for Model 2), as well as the variances, $V = 2.256$ (resp. 1.967 for Model 2). Figure 2 shows the

scatterplot of the rank of the difficulty parameters. Most importantly, the correlation of the difficulty parameters is high, $R^2 = .95$. For this reason, measurement invariance can be assumed.

Because of the high correlation between both measurement points, the approach of virtual cases could be applied (see Hartig & Kühnbach, 2005; Rost, 1996). If a participant took part at both measurement points, he or she represents two cases in the data. This approach has two advantages: Firstly, resulting scores are in a consistent metric and secondly, the appearance of the problem of constrained variance is less likely (see above). With this approach, a data set of 724 cases resulted. Again, a one-dimensional model was calculated but also a three-dimensional model consisting of the three performance levels described before. Results are shown in Table 1. Both models show good variances and EAP/PV-reliabilities as well as discrimination indices. The three-dimensional model shows significantly less deviance than the one-dimensional model ($p < .001$), so it fits the data better.² Figure 3 shows the wright-map of the three-dimensional model. Thus, the Rasch-analysis empirically confirmed the three-dimensional structure of the test.

Dimensionality	One-dimensional	Three-dimensional		
		I	II	III
Performance Level				
Number of tasks	22	8	10	4
Variance	2.216	2.346	1.963	3.661
EAP-PV-Reliability	.872	.868	.856	.842
Weighted MNSQ	.86-1.32	.92-1.12	.83-1.24	.93-1.06
Discrimination Indices	.36-.69	.52-.68	.39-.71	.69-.73
Deviance	16293.17612	16258.04687		
Estimated Parameters	44	31		
Difference Deviance	35.12925			
Difference Parameters	5			
Sig. (χ^2 -Difference-Test)	< .001			

Note: EAP/PV-Reliability – Expected-A-Priori / Plausible-Values-Reliability (see Rost, 1996)

Table 1: Results of the Rasch-Analysis

The order of tasks within the dimensions could be understood as a consequence of how intensively and how often (regarding different classes, schools etc.) the tasks are practiced; the order could hardly be deduced by theoretical considerations alone. In accordance with our content analysis, Dimension II in fact includes tasks which need conversions between different representations of a function, even if these tasks are not unfamiliar. The task in Figure 1 is even rated as a rather difficult one within Dimension II in the Rasch model. Within Dimension III, besides the degree of institutionalization and the time of treatment of indicated techniques in school,

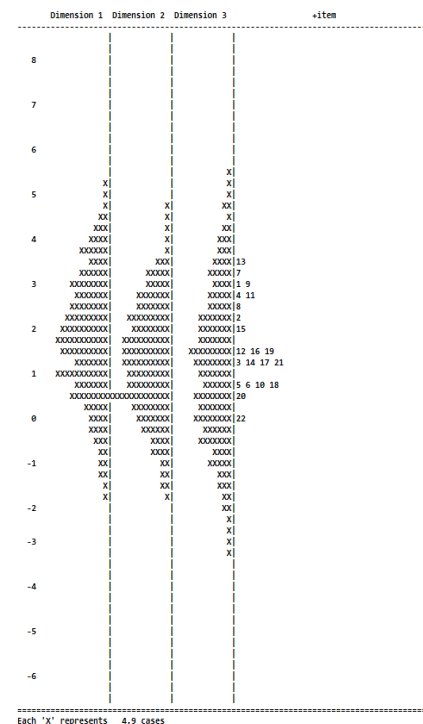


Figure 3: Wright-map of the 3D

² One item (Item 3) shows a weighted MNSQ larger than 1.2 (values should lie in the range of $0.8 < \text{MNSQ} < 1.2$). This is not too problematic for all other indices perform very well.

the praxeological interconnectedness seems to be reflected in the order of tasks (items 3, 7, 10, 11, 12, 13, 15, 18, 20, 22). It is perhaps a little bit surprising that the complex term transformation task (item 13) turned out to be the most challenging task of Dimension III.

Discussion

The main aim of bridging courses is to support students by providing learning materials and samples of tasks for their self-evaluation. From the design of material and courses as well as the point of view of teaching, specific test instruments are required for a possibly objective evaluation of the impact of support measures. The test instruments should be most sensitive for specific mathematical knowledge domains and should allow relating pre-post-comparisons to specifics of bridging courses in order to adopt courses in view of actual pre-test results. Our results show that our test takes into account qualitatively different knowledge dimensions and might be a valuable tool for reaching these goals, which, of course, needs further investigations.

References

- Barbé, J., Bosch, M., Espinoza, L., & Gascón, J. (2005). Didactic restrictions on the teacher's practice: The case of limits of functions in Spanish High Schools. *Educational Studies in Mathematics*, 59 (1/3), 235–268.
- Bausch, I., Biehler, R., Bruder, R., Hochmuth, R., Koepf, W., Fischer, P.R., & Wassong, T. (2014). VEMINT– Interaktives Lernmaterial für mathematische Vor- und Brückenkurse [Interactive learning material for mathematics pre- and bridging-courses]. In I. Bausch, R. Biehler, R. Bruder, P.R. Fischer, R. Hochmuth, W. Koepf, S. Schreiber, & T. Wassong (Eds.), *Mathematische Vor- und Brückenkurse* [Mathematics pre- and bridging-courses] (pp. 261-276). Wiesbaden: Springer Spektrum.
- Biehler, R., & Hochmuth, R. (2017). Relating different mathematical praxeologies as a challenge for designing mathematical content for bridging courses. In R. Göller, R. Biehler, R. Hochmuth, H.-G. Rück (Eds.), *Didactics of mathematics in higher education as a scientific discipline Conference proceedings. Khdm-Report 17-05* (pp. 14-20). Kassel: Universität Kassel.
- Bruder, R., Feldt-Caesar, N., Pallack, A., Pinkernell, G., & Wynands, A. (2015). Mathematisches Grundwissen und Grundkönnen in der Sekundarstufe II [Basics in mathematics in secondary education]. In W. Blum (Eds.), *Bildungsstandards aktuell: Mathematik in der Sekundarstufe II* [Education standards: Mathematics in secondary education] (pp. 108–124). Braunschweig: Schrödel.
- Bildungsberichterstattung (2014). *Bildung in Deutschland 2014* [Education in Germany 2014]. Bielefeld: W. Bertelsmann Verlag.
- Chevallard, Y. (1992). Fundamental concepts in didactics: Perspectives provided by an anthropological approach. In R. Douady & A. Mercier (Eds.), *Research in didactique of mathematics, selected papers* (pp. 131–167). Grenoble: La Pensée Sauvage.
- Cooperation Schule Hochschule (2014). *Mindestanforderungskatalog Mathematik* [Minimum requirements of mathematics] (version 2.0). Retrieved from: https://www.hs-karlsruhe.de/fileadmin/hska/SCSL/Lehre/mak V2.0B_ohne_Leerseiten.pdf.
- Cramer, E., & Walcher, S. (2010). Schulmathematik und Studierfähigkeit [School mathematics and

- academic achievement]. *Mitteilungen der DMV*, 18(2), 110-114.
- Feldt-Caesar, N. (2017). *Konzeptualisierung und Diagnose von mathematischem Grundwissen und Grundkönnen. Eine theoretische Betrachtung und exemplarische Konkretisierung am Ende der Sekundarstufe II* [Concepts and diagnosis of basics in mathematics. A theoretical perspective and specification at the end of secondary education]. Wiesbaden: Springer.
- Fischer, P. R. (2014). *Mathematische Vorkurse im Blended-Learning-Format: Konstruktion, Implementation und wissenschaftliche Evaluation* [Mathematics preparation courses in blended-learning-format: construction, implementation and scientific evaluation]. Wiesbaden: Springer.
- Hartig, J., & Kühnbach, O. (2006). Schätzung von Veränderung mit „plausible values“ in mehrdimensionalen Rasch-Modellen [Estimation of variables with “plausible values“ in multidimensional Rasch-models]. In A. Ittel & H. Merkens (Eds.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* [Measurement of change and longitudinal studies in empirical educational studies] (pp. 27-44). VS Verlag für Sozialwissenschaften.
- Hilgert, J. (2016). Schwierigkeiten beim Übergang von Schule zu Hochschule im zeitlichen Vergleich—Ein Blick auf Defizite beim Erwerb von Schlüsselkompetenzen [Difficulties in the transition from school to high school in a time comparison—a view to deficits of acquisition of key-competences]. In A. Hoppenbrock, R. Biehler, R. Hochmuth, & H.-G. Rück (Eds.), *Lehren und Lernen von Mathematik in der Studieneingangsphase* [Teaching and learning of mathematics during the introductory phase] (pp. 695-709). Wiesbaden: Springer.
- Hochmuth, R. (2018). *A general scheme for a heterogeneous manifold of transitions*. Paper presented at the 6th International Congress of Anthropological Theory of Didactics, Grenoble (22nd -26th Jan. 2018).
- Neumann, I., Pigge, C., & Heinze, A. (2017). *Welche mathematischen Lernvoraussetzungen erwarten Hochschullehrende für ein MINT-Studium* [Which mathematical learning requirements do university lecturers expect for STEM-courses]? Institut für die Pädagogik der Naturwissenschaften Kiel. Retrieved from https://www.telekom-stiftung.de/sites/default/files/malemint_broschure_langfassung2.pdf
- Nitsch, R., Bruder, R., & Kelava, A. (2016). Schülerhandlungen als Elemente fachdidaktisch motivierter Kompetenzmodellierungen [Student's performance as elements of didactically motivated competence modelling]. *Journal für Mathematik-Didaktik*, 37(2), 289-317.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (expanded edition). Chicago: The University of Chicago Press.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion* [Textbook test theory, test construction]. Bern: H. Huber.
- Sangwin, C. (2007). *Computer aided assessment of mathematics*. Oxford: Oxford University Press.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77 (1), 20-26.
- Weinert, F. E. (2000). Lehren und Lernen für die Zukunft-Ansprüche an das Lernen in der Schule [Teaching and learning for future requirements to learning in school]. *Rheinland-Pfälzische Schule*, 12, 240-246.