ANALYSIS OF COMPETENCE STRUCTURE IN THE AREA OF COMPETENCE *USE OF CONTENT KNOWLEDGE* IN CHEMISTRY

Roxana Timofte
"Babes-Bolyai" University, Cluj-Napoca, Romania
Elke Sumfleth
University of Duisburg-Essen, Germany
Maik Walpuski
University of Duisburg-Essen, Germany

Abstract. In 2004 were proposed the national standards for science in Germany. For the evaluation of the proposed standards, the project Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level, with the acronym ESNaS, was started in 2007 (Kremer *et al.*, 2012; Kauertz *et al.*, 2010). A three-dimensional competence model was developed, with the three axes: competence area, complexity and cognitive processes. The focus of this study was the development of an instrument to assess pupils' competencies in the area of competence *use of content knowledge* in Chemistry. A number of 101 items were developed by using the ESNaS model, in order to evaluate the competence in Chemistry of High School German pupils. Items were developed on the theoretically-defined aspects of the topic *Chemical Reactions*: acid-base reactions, redox reactions, fundamentals of reactions, reactions used in analysis, organic chemistry reactions. The Rasch model was used for the analysis of data. A number of N=568 German pupils (9th, 10th and 11th class) were tested, and it was revealed that the instrument is unidimensional and has good psychometric properties.

Keywords: Rasch model, competence analysis, assessment in Chemistry, chemical reactions, dimensionality

As a consequence of the unsatisfactory results of the German students at the TIMSS tests in 1990s and 2000s, the educational standards were introduced in German education system in 2004 (Köller *at al.*, 2012). In order to evaluate the national standards, an interdisciplinary project "Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level" (ESNaS) was started in 2007 (Kremer *et al.*, 2012, Kauertz *et al.*, 2010). A three-dimensional competence model was proposed, and the dimensions were: area of competence, cognitive processes and complexity level (Figure 1).

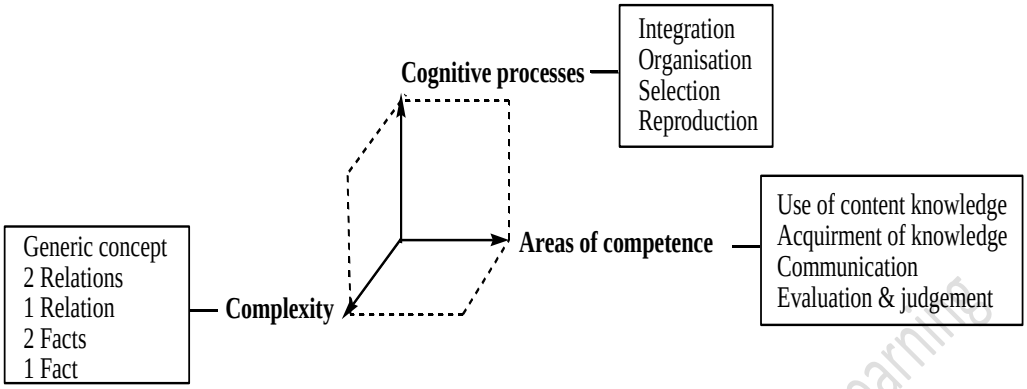

Figure 1. *Three-dimensional ESNaS competency model*

The areas of competence were: use of content knowledge, acquirement of knowledge, communication, evaluation and judgment. The cognitive processes specified in the model refer to the processes necessary to solve a task. The complexity dimension reflects the structure of the information used in the item. For the area of competence *use of content knowledge*, four basic concepts were defined by the national education standards in Germany (Kremer *et al.*, 2012): chemical reactions, energy, particulate nature of matter and structure-properties relationship. For the basic concept *chemical reactions*, five aspects were defined theoretically (Figure 2) and the corresponding indicators were identified in curricula from the states in Germany (Timofte, 2015).

Area of competence	Use of content knowledge
Basic concept	Chemical Reactions
Aspects	Redox Reactions Acid-Base Reactions Organic Reactions Reactions used in Quantitative and Qualitative Analysis Fundamentals of Chemical Reactions

Figure 2. *Aspects of the basic concept* "*Chemical Reactions"*

In Germany there is a curriculum for each county and it was necessary to pursue a curricula analysis, to identify the most frequent indicators across the curricula in Germany. The most frequent indicators were: redox reactions, oxidation, acid-base reaction, pH, substitution, addition, reactions used in quantitative analysis (general information), reactions used in qualitative analysis (general information), detection of organic compounds, definition of reaction, law of mass conservation, word equation, formula equation (Timofte, 2015). The identification of the most frequent indicators across curricula was required, in order to follow the principle of alignment between curricula and assessment (Webb, 1997).

Item Response Theory (IRT) models are used to analyse and interpret empirical data in the area of assessment. Rasch model is one of the IRT models. When using Rasch models for analysis, data is converted in ratio scaled data and person parameters and item parameters are generated (Bond *et al.*, 2001). A number of parameters could be obtained during the Rasch analysis of data: Wright map, item difficulty, item fit statistics, local independency of items, item separation reliability, person ability, person separation reliability, different item functioning, dimensionality of the instrument (Boone *et al.*, 2006). Item fit statistics is a way to test assumptions of assessment. MNSQ values ranging from 0.75 to 1.33 are regarded as being acceptable. It is considered that an item is a misfit if the output MNSQ>1.33 (Wilson, 2005). The acceptable values for items measure in Rasch model are ranging from -3.5 to +3.5 logits (Rauch & Hartig, 2008). The acceptable value for person reliability (similar to Cronbach's alpha) is minimum 0.75. Items with correlation values higher than 0.7 are considered to be locally dependent (Linacre, 2011).

Aim of the study

The aim of this study was the development of an instrument to measure German pupils' abilities in the area of competence *use of content knowledge*, on the topic *chemical reactions* and to determine the psychometric parameters of the instrument. Furthermore, it was aimed to establish if the aspects defined theoretically were belonging to the same empirical scale.

Methods

A number of 101 multiple choice single select and open-ended items were tested, on topics related to the five aspects depicted in Figure 2. The items were developed by using the ESNaS model. A multi-matrix design was used for the development of the test booklets and a number of N=568 pupils (9^{th}, 10^{th} and 11^{th} class) from North Rhine-Westphalia county in Germany were tested. Data was analysed with the Winsteps programme version 4.

Results and discussion - analysis of data

Assessment of item fit: From 101 items tested, 11 (10.9%) were eliminated, since they exhibited values for outfit MNSQ > 1.3. The analysis was carried out further without these items.

Items measure: Items' difficulty measures were in the range of -2.67 and +3.39 logits (Figure 3).

Reliability: Both person and item and reliability separation values were within accepted ranges: person reliability: 0.75, person separation: 1.71, item reliability: 0.95, item separation: 4.60.

Local independence of items: All correlations are below 0.7 and hence, it can be considered that the items are locally independent.

Dimensionality: The unidimensionality of the instrument was tested by the principal component analysis of the residual error.

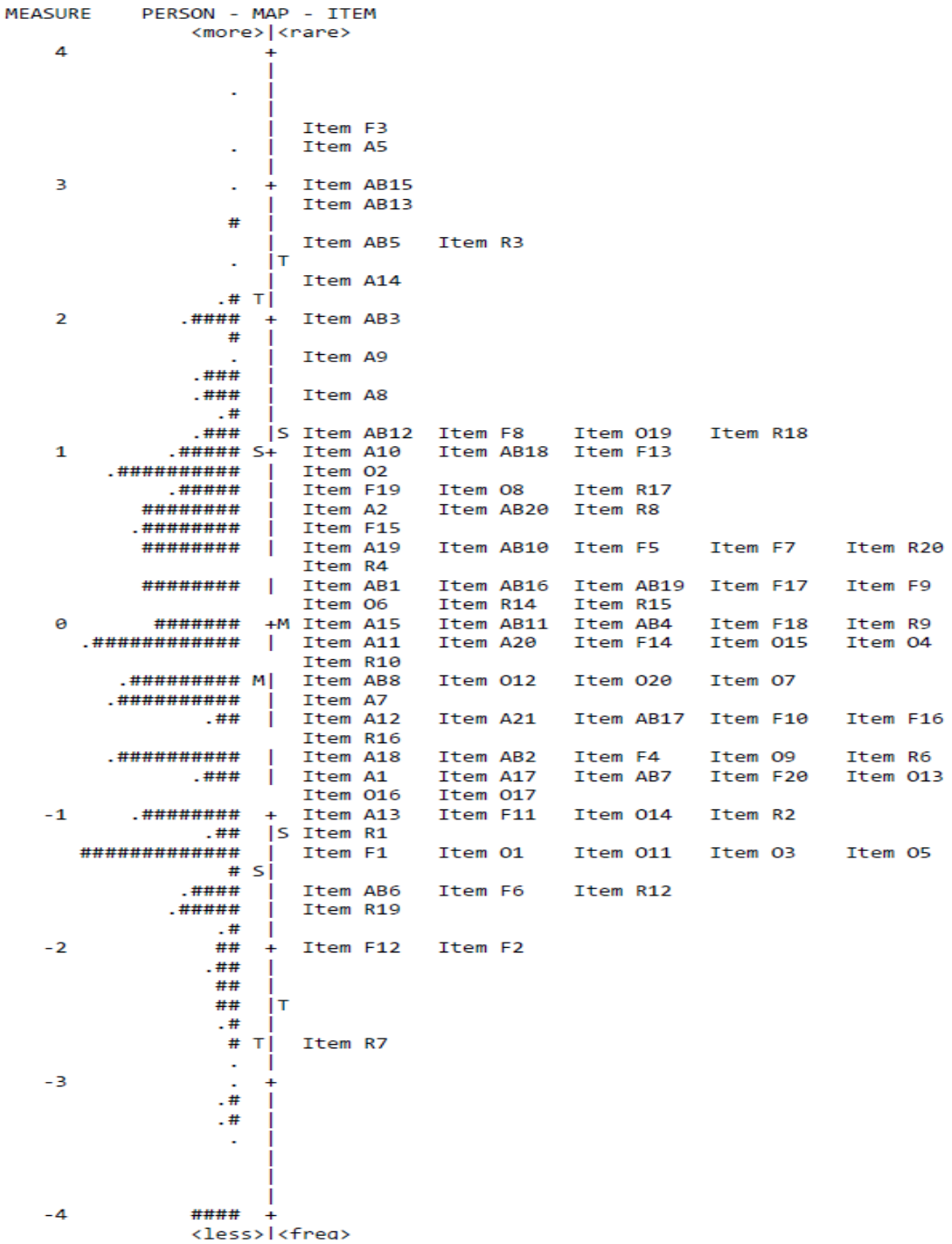

Figure 3. *Wright map*

In Figure 4 is depicted the item loading plot. Analysis has shown the presence of three contrasts, with majority of items related to one contrast. The value of disattenuated correlation for each of the three contrasts (item clusters 1-3, 1-2, 2-3) was 1. Consequently, it was concluded that the instrument is unidimensional, presenting some strands.

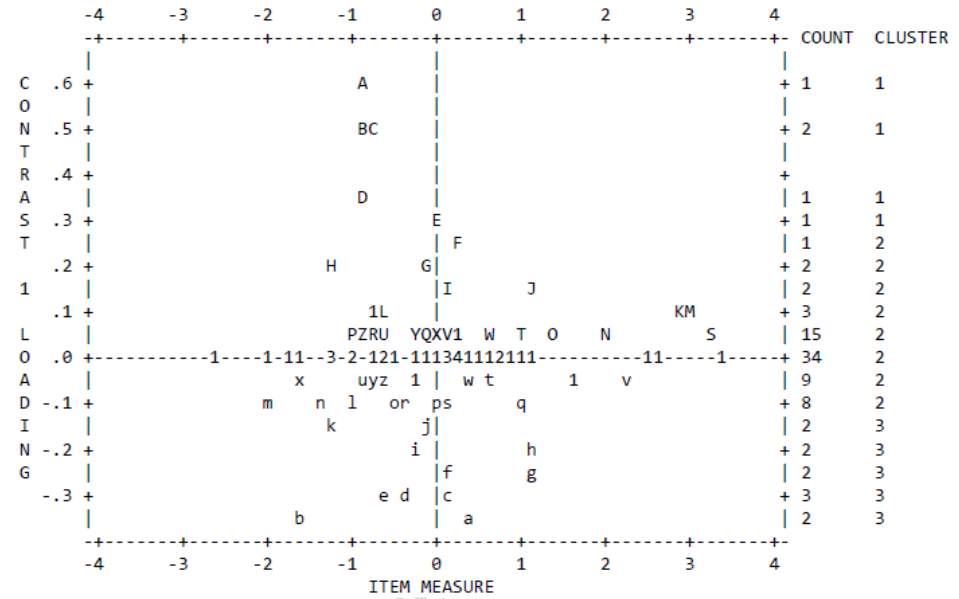

Figure 4. *Item loading plot*

Conclusion

An instrument to measure pupils' competencies on the topic *Chemical Reactions* was developed. The instrument comprises 101 items having different levels of difficulty and was tested on 568 pupils from Germany. Systematic data analysis has proven that the instrument has good reliability and validity and that it is unidimensional.

References:

Bond, T. G., Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, NJ: Lawrence Erlbaum.

Boone, W. J., Scantlebury, K. (2006). The Role of Rasch Analysis When Conducting Science Education Research Utilizing Multiple-Choice Tests. *Inc. Sci Ed*, *90***,** 253– 269.

Kauertz, A., Fischer, H., Mayer, J., Sumfleth, E., & Walpuski, M. (2010). Standardbezogene Kompetenzmodellierung in den Naturwissenschaften der Sekundarstufe. *Zeitschrift für Didaktik der Naturwissenschaften, 16*, 135-153.

Köller, O., Parchmann, I. (2012), Competencies: The German Notion of Learning Outcomes. In S. Bernholt, K. Neumann, P. Nentwig (Eds), *Making it tangible. Learning outcomes in science education* (pp. 151-173). Münster: Waxmann Verlag GmbH.

Kremer, K., Fischer, H. E., Kauertz, A., Mayer, J., Sumfleth, E., & Walpuski, M. (2012). Assessment of standards-based learning outcomes in science education: Perspectives from the German project ESNaS. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Making it tangible. Learning outcomes in science education* (pp. 201-218). Münster, New York: Waxmann.

Linacre J. M., (2011). *A User's Guide to Winsteps: Rasch-Model Computer Programs*, winsteps.com.

Rauch, D., & Hartig, J. (2008). Interpretation von Testwerten in der IRT. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (pp. 240–250). Berlin: Springer.

Timofte, R. S. (2015). Content analysis of Chemistry curricula in Germany. Case study: Chemical reactions. *Acta Didactica Napocensia*, *8*(4), 25-31.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Council of Chief State School Officers and National Institute for Science Education Research Monograph No. 6). Madison, WI: University of Wisconsin -Madison, Wisconsin Center for Education Research.

Wilson, M. (2005). Constructing measures: An item response modeling approach. Mahwah, NJ: Erlbaum Associates.

Acknowledgments

RT is grateful to Dr. Carina Gehlen for fruitful discussions on the topic of competence assessment.