ASPECTS OF STATISTICAL LITERACY BETWEEN COMPETENCY MEASURES AND INDICATORS FOR CONCEPTUAL KNOWLEDGE: EMPIRICAL RESEARCH IN THE PROJECT "RIKO-STAT"

Sebastian Kuntze, Joachim Engel, Laura Martignon and Martin Gundlach
Ludwigsburg University of Education, Germany
kuntze@ph-ludwigsburg.de

Competencies in the area of statistical literacy have been the subject of several studies which have succeeded in establishing competency measure instruments. However, the approaches differ from each other. Moreover, even though there is a body of research on conceptual knowledge potentially linked with components of statistical literacy, there is a need of collecting empirical evidence on the role of conceptual knowledge for the development of statistical literacy.
Consequently, in the research work of the project RIKO-STAT, we focus on the interrelatedness of conceptual knowledge in different areas and the competency of using representations and models in statistical contexts. In the first phase of the project, more than 450 primary students, 600 secondary students and 350 university students were assessed. The paper reports on goals and on the design of the first phase of RIKO-STAT as well as on preliminary empirical results.

INTRODUCTION

Prerequisites for participation (in the sense of participation as an informed citizen) related to statistical thinking, to decisions under risk and to data-specific communication are included in the notion of statistical literacy. Even though there are empirically successful approaches to describing areas of statistical literacy in corresponding competency models, there is a need for research concerning influencing factors on competencies in the domain of statistical literacy. Such research could provide a basis for focused interventions developing these competencies.

Accordingly, research activities of the project "RIKO-STAT" aim at exploring relationships between the learners' competency in using representations including modeling in statistical contexts and influencing factors, especially located in conceptual knowledge. A corresponding model of influencing factors has been implemented in a test instrument which was administered to primary, secondary and university students. The preliminary results presented here are limited to the sample of university students and suggest that there are moderate interdependencies between the competency score and the considered areas of conceptual knowledge.

In the first section of this paper, we will introduce the theoretical background. The second section will present the project RIKO-STAT, some of its key research questions and the design of the study. Preliminary results will be reported in the third section and interpreted in the concluding fourth section.

THEORETICAL BACKGROUND

Instructional goals linked to modelling and dealing with data have seen a rise in emphasis in many countries (Batanero, Godino, Valecillos, Green & Holmes, 1995), a development which has been accelerated by the development of standards (e.g., for the case of Germany: KMK, 2004) and international assessment studies like PISA (OECD, 2003). Correspondingly, competencies in the area of "statistical literacy" (Wallman, 1993), which link dealing with data and modelling, are at the centre of interest when focusing on such new developments. The increase in emphasis on modelling and dealing with data should be seen against the background of the goal of providing responsible citizens with competencies of evaluating data and representations, decision making under uncertainty and risk communication, which includes abilities of explaining situational contexts by statistical means and drawing conclusions on the basis of statistical evidence (cf. Wild & Pfannkuch, 1999).

Two overarching content aspects appear crucial for building up knowledge in this field: *manipulation of data by reduction* and *dealing with statistical variation*. Firstly, manipulation of data by reduction is an idea described by Kröpfl, Peschek and Schneider (2000) and concerns the spectrum between an overview on data and detail information contained in data. In this spectrum, the reduction of information is an important strategy statistical of reasoning for gaining overviews.

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE]

Secondly, Watson and Callingham (2003, cf. also Watson, 1997) emphasise that dealing with statistical variation represents a central idea of statistical literacy. In the study of Watson and Callingham, competencies were measured by means of encodings of a test designed according to this approach. These encodings conformed to the one-dimensional Rasch model.

Against the background of the differences between these approaches and the existence of other points of view, e.g., by Reading (2002) and Curcio (1987), the different approaches were integrated in a four-level competency model of "using representations and models in statistical contexts" by Kuntze, Lindmeier and Reiss (2008). This model uses the notion of data-related "reading" (cf. Curcio, 1987), which helps to establish links between the two overarching ideas introduced above. This means that the model of Curcio is enlarged by the thought that data-related "reading" sometimes requires the modeling of statistical variation, even and in particular when being presented with diagram-like representations of data (Kuntze, Lindmeier & Reiss, 2008).

The evaluation of a corresponding competency test showed that the competency could be described with a one-dimensional Rasch model (Kuntze, Lindmeier & Reiss, 2008; cf. also Lindmeier, Kuntze & Reiss, 2007).

Even though the research of Watson and Callingham (2003) and of most other researchers in this field was based on prior theoretical considerations of conceptual knowledge in the area of statistics, probability and representations, there is still a need of empirical evidence concerning influencing factors on this competency in areas of conceptual knowledge, especially as far as the areas of probability, functional reasoning and determinist views of learners (cf. e.g. Engel & Sedlmeier, 2005) are concerned. Moreover, motivational variables and epistemological beliefs about mathematics and statistics may have an impact on competency development, as they can have a filtering function in the process of competency development. These influencing factors are represented in the model in Figure 1. This model is not meant to be complete, there may be additional influencing variables, e.g. in the domain of self-regulation or meta-cognition. Moreover, other conceptual knowledge more closely related to diagrams and statistics probably play a role, too, but from a theoretical point of view, they appear almost inseparable from the competency of using representations and models in statistical contexts itself.

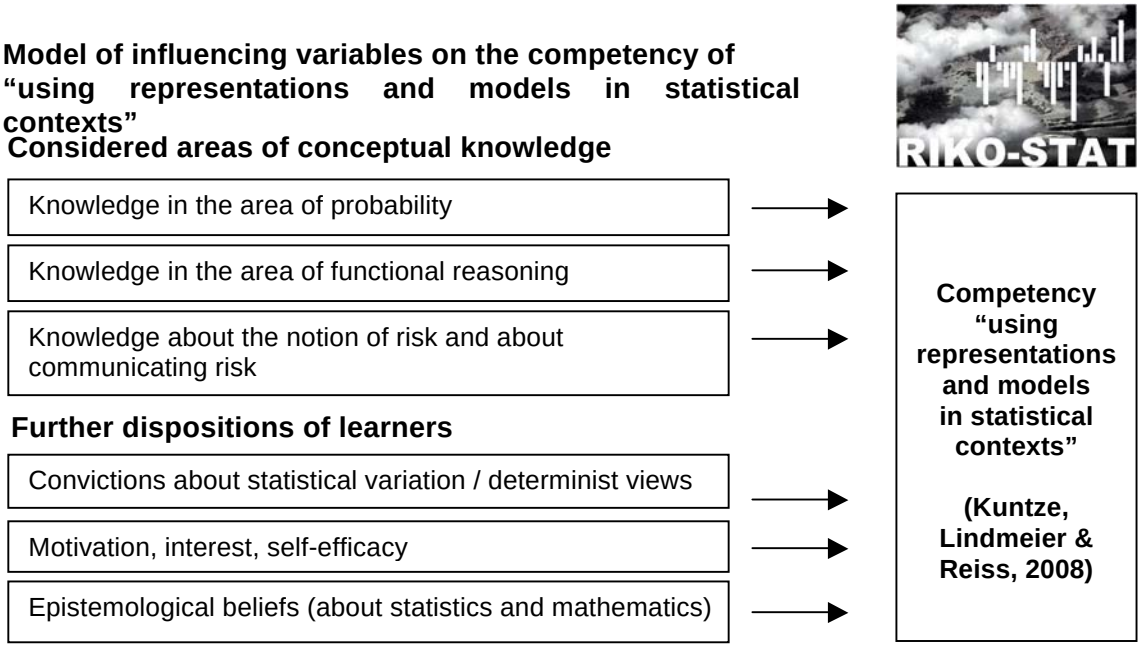

Figure 1. Overview on the model of influencing factors used in the study "RIKO-STAT"

THE PROJECT "RIKO-STAT"

The model shown in Figure 1 was used for framing the research aims of the project RIKO-STAT. This project focuses on finding out about the interrelatedness of the variables shown in Figure 1. For this reason, we will give in the following an outline of this project, present two key research questions and aspects of its research design.

RIKO-STAT focuses on the competency of using representations and models in statistical contexts and influencing variables for primary (year 4), secondary (year 9) and university students. In the first project phase, corresponding test and questionnaire instruments were developed and administered to more than 450 primary students, 600 secondary students and 350 university students. In the second project phase, learning environments for school students will be evaluated with the help of the instruments of the first phase, in order to find out about the extent, to which these learning environments can foster components of statistical literacy and related conceptual knowledge.

Selected research questions

Corresponding to the theoretical considerations presented above, two central research questions in RIKO-STAT are the following:

- What relationships are there between the areas of conceptual knowledge presented above and the competency of using representations and models in statistical contexts?
- Are there relationships between this competency and other potential influencing variables e.g. in the area of motivation and self-efficacy, determinist convictions or epistemological beliefs related to statistics and mathematics?

As a result of the length limits of this paper, we will concentrate on the first research question.

Design of the study

In the first phase of RIKO-STAT, learners in the different sub-samples were assessed by a test and questionnaire instrument. The competency test is based on an enlarged version of an instrument developed by Kuntze, Lindmeier and Reiss (2008). This test instrument allows the measure of competencies of learners in a one-dimensional Rasch model and it consists of subsets of tasks on different levels of competency (see Kuntze, Lindmeier & Reiss, 2008, also for sample test items). An additional subtest contains indicators for conceptual knowledge in the areas of probability, functional reasoning, and dealing with risk with sets of tasks focusing on each of these areas. Moreover, questionnaire items focus on determinist views, motivational dispositions of the learners and their epistemological beliefs concerning mathematics and statistics. The mathematics-related scales were based on earlier studies (e.g., Grigutsch, Raatz & Törner, 1995), whereas the scales specific for statistics were new developments.

The results reported in section 3 are based on a sample of 360 university students (304 female, 56 male), who were asked to fill in the test at the beginning of university courses.

PRELIMINARY RESULTS

The analyses of the data still being in progress, we will give a brief overview on preliminary results. Given the length limits of this paper, we concentrate on the first research question. More comprehensive and detailed presentations of results based on the data of all three sub-samples are forthcoming.

The ratios of successfully answered tasks in the competency sub-test indicate that the theoretical expectation of competency levels is supported by the data. Hence, prior studies (e.g., Kuntze, Lindmeier & Reiss, 2008) have been replicated as far as the properties of the test instrument are concerned. A new fifth competency level introduced in RIKO-STAT fits to the model of four competency levels of Kuntze, Lindmeier and Reiss (2008) and expands this model.

The new scales for statistics-specific epistemological beliefs and for motivational characteristics related to statistics have shown to be reliable (e.g., Gundlach et al., in this volume). There are positive evaluation results of the test and questionnaire instrument.

The first research question focuses on a key research interest of RIKO-STAT, namely on relationships between conceptual knowledge and the competency of using representations and models in statistical contexts. On the base of sum scores of all subtests, we observe moderate correlations (Pearson) between the competency score and the conceptual knowledge scores of functional reasoning ($r=0.30$), probability ($r=0.36$) and risk ($r=0.31$, all correlations significant with $p<0.001$). As the conceptual knowledge scores are also interrelated among themselves (r between 0.34 and 0.37), a regression analysis was performed including these three variables (adjusted

r^2=0.17, regression coefficients for functional reasoning: β=0.13, probability: β=0.25 and risk: β=0.17). The analysis reveals a tendency that the regression coefficient of conceptual knowledge in the area of probability is higher than the other ones. This tendency is also reflected in partial correlations, showing that especially the correlation between the competency score and conceptual knowledge in the area of functions becomes smaller, if we control e.g. for the conceptual knowledge score related to probability.

CONCLUSIONS

Our preliminary results suggest that even though representations and diagrams are linked to the conceptual knowledge about functions, the conceptual knowledge in the area of probability shows a relatively stronger interdependency with the competency of using representations and models in statistical contexts. An interpretation of these findings is that conceptual knowledge about probability supports the ability of dealing with statistical variation, which is an important component of statistical literacy, and that this contribution appears to be more substantial than the contribution of knowledge about functions. However, at this stage, the results should be interpreted with care, as general cognitive abilities could have had an impact on the observed correlations.

ACKNOWLEDGEMENT

This research is supported by research grants of Ludwigsburg University of Education.

REFERENCES

Batanero, C., Godino, J., Vallecillos, A., Green, D., & Holmes, P. (1995). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, *25*(4), 527-547.

Curcio, F. R. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, *18*(5), 382-393.

Engel, J., & Sedlmeier, P. (2005). On middle-school students' comprehension of randomness and chance variability in data. *Zentralblatt für Didaktik der Mathematik (ZDM)*, *37*(3), 168-177.

Grigutsch, S., Raatz, U., & Törner, G. (1995). *Mathematische Weltbilder bei Lehrern*. [Teachers' mathematical world views]. Gerhard-Mercator-Universität Duisburg Gesamthochschule, Schriftenreihe des Fachbereichs Mathematik, Preprint Nr. 296.

KMK (Kultusministerkonferenz). (2004). *Bildungsstandards im Fach Mathematik für den mittleren Schulabschluss*. München: Wolters Kluwer.

Kröpfl, B., Peschek, W., & Schneider, E. (2000). Stochastik in der Schule: Globale Ideen, lokale Bedeutungen, zentrale Tätigkeiten. *mathematica didactica*, *23*(2), 25-57.

Kuntze, S., Lindmeier, A., & Reiss, K. (2008). "Using models and representations in statistical contexts" as a sub-competency of statistical literacy–Results from three empirical studies. *11th Int. Congr. on Math. Education (ICME 11)*. Online: http://tsg.icme11.org/document/get/474.

Lindmeier, A., Kuntze, S., & Reiss, K. (2007). Representations of data and manipulations through reduction–competencies of German secondary students. In B. Philips & L. Weldon (Ed.), *Proc. of the IASE/ISI Sat. Conf. on Stat. Educ.*, Guimarães, Portugal. Voorburg, NL: ISI.

OECD. (2003). The PISA 2003 Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills. Online: www.pisa.oecd.org/dataoecd/46/14/ 33694881.pdf.

Reading, C. (2002). Profile for statistical understanding, In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Online: www.stat.auckland.ac.nz/~iase/publications/1/1a4_read.pdf.

Wallman, K. (1993). Enhancing Statistical Literacy: Enriching our Society. *Journal of the American Statistical Association*, *88*(421), 1-8.

Watson, J.M. (1997). Assessing Statistical Thinking Using the Media. In I. Gal, & J.B. Garfield, (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-121). IOS Press.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3-46.

Wild, C., & Pfannkuch, M. (1999): Statistical thinking in empirical enquiry. *International Statistical Review*, *3*, 223-266.