Durand-Guerrier, V., Hochmuth, R., Goodchild, S., & Hogstad, N. M. (Eds.). (2018). PROCEEDINGS of INDRUM 2018
Second conference of the International Network for Didactic Research in University Mathematics. University of Agder and INDRUM: Kristiansand, Norway.

Evaluating Innovative Measures in University Mathematics – The Case of Affective Outcomes in a Lecture focused on Problem-Solving

Christiane Kuklinski[1], Elena Leis[1], Michael Liebendörfer[1], Reinhard Hochmuth[1], Rolf Biehler[2], Elisa Lankeit[2], Silke Neuhaus[2], Niclas Schaper[2], and Mirko Schürmann[2]

[1]Leibniz Universität Hannover, Institut der Didaktik für Mathematik und Physik, Germany, kuklinski@idmp.uni-hannover.de ; [2]Universität Paderborn, Germany

The transition from high school to university mathematics has proven to be difficult for many students but especially for pre-service secondary teachers. To support these students at mastering this transition, various universities have introduced support measures of various kinds. The WiGeMath project developed a taxonomy that makes it possible to describe and compare these measures concerning their goals as well as their frame characteristics. We exemplify the use of the taxonomy in the description of one specific innovative measure that was part of the WiGeMath evaluations. Moreover, we present first results concerning the goal-fulfilment of this measure concerning affective characteristics of the student cohort and their predominant beliefs.

Keywords: Transition to and across university mathematics, Novel approaches to teaching, Teacher education, Motivational developments, Beliefs.

BACKGROUND

In German mathematics teacher education, pre-service teachers first study at university before they enter a practical training. In this first phase, there is a strong focus on mathematical content, in particular in higher secondary teacher education where students mostly attend the same courses as mathematics major students. In these shared lectures, many internationally known problems of the secondary-tertiary transition arise (Gueudet, 2008), in particular, motivational problems and drop-out are often reported. There is a substantial decline in students' mathematical interest in the first semester with Cohen's d around 0.4 (Rach und Heinze, 2013, 2016), a decline in their mathematical self-concept with Cohen's d ranging from 0.5 to 0.7 (Rach und Heinze, 2013, 2016), and a strong dominance of controlled motivation over autonomous motivation (Liebendörfer, in press) in the terminology of Ryan and Deci's (2017) self-determination theory. Consequently, many pre-service students experience their university courses as a necessary evil rather than a helpful qualification towards their aspired job (Kalesse, 1997; Liebendörfer, in press).

THE WIGEMATH PROJECT

To counteract the negative effects which for many students seem to occur at the transition between school and university mathematics, many universities have introduced support measures of various kinds. Even though university internal evaluations of these measures mostly exist, a framework that helps to facilitate the comparison of design and outcomes of different measures has until recently been lacking. The WiGeMath project (Wirkung und Gelingensbedingungen von Unterstützungsmaßnahmen für mathematikbezogenes Lernen in der Studieneingangsphase; Effects and success conditions of mathematics learning support in the introductory study phase) [1], which is a joint research project of the Universities of Hannover and Paderborn (Colberg et al., 2016) led by Biehler, Hochmuth and Schaper, has developed a framework for goal dimensions and frame conditions of mathematics learning support in universities (Liebendörfer et al., in press) that aims at such a comparison. Moreover, the project has used the framework in first evaluations of various support measures at different universities in Germany. Some exemplary results for one representative of the category of redesigned lectures, which is one type of support measure that was evaluated in the project, is presented below.

THE WIGEMATH TAXONOMY

The aim of the WiGeMath project is to develop and exemplify in use a taxonomy that categorizes features and goals of Projects of Mathematics Learning Support (PMLS) and to use this taxonomy to evaluate different support measures at German universities. All projects that fall under the category of PMLS have in common that they try to support students at the beginning of their university studies in mastering the critical transition to university mathematics. They are innovative insofar as they deviate from the standard format of lectures and tutorials that is encountered in university mathematics even though the way how they do this differs. In the WiGeMath project, different PMLS are subsumed under one of four categories, namely bridging courses, mathematics support centres, support measures that parallel courses and redesigned lectures. Due to space limits, this text focuses only on redesigned lectures; a description of the other types of PMLS is given in (Liebendörfer et al., in press).

Redesigned lectures are lectures that offer particular support to students that have been shown to have higher risks at failing mathematics courses or focus on very specific learning goals in a non-traditional way. We examined both redesigned lectures that address preservice secondary teachers, who often show the greatest problems with the transition from school mathematics to more abstract mathematical content, and redesigned lectures which address engineering students who had already failed a compulsory mathematics test of some kind. All redesigned lectures have in common that new mathematical content is not the focus of teaching.

Different PMLS have different aims some of which are explicit but some of which stay hidden even to the teaching staff until they are inquired about them by an outsider. WiGeMath aimed at evaluating different PMLS based on their own assumptions following a program evaluation approach (Chen, 1990) as well as comparing them on grounds of an encompassing taxonomy that should categorize descriptive (non-normative) goals in the sense of criterions that the PMLSs set out to meet in their conception, features and conditions of PMLSs.

The taxonomy was constructed in a two-fold process. A first draft was developed by the project members by means of a document analysis, taking into account documents provided by project partners. The WiGeMath project collaborates with 11 partner universities in Germany at which PMLS have been implemented. The draft for the taxonomy was then used as a guiding thread for guided interviews with teaching staff of eight PMLS, two of each category. The interviews were taped and transcribed and afterwards the draft for the taxonomy was tested by trying to fit mentioned goals into the draft's categories. This led to minor refinements and reformulations of categories and yielded the final WiGeMath taxonomy.

This final WiGeMath taxonomy consists of three main categories, namely frame conditions, measure categories/ characteristics and goal categories. The frame conditions include various sub-categories, which help to characterize the student cohort addressed by a PMLS, the way it came about and developed, its embeddedness in the university course system, organisational characteristics that may pertain to it, characteristics of the room where it is held, financial and staff conditions and lastly characteristics of the learning culture. Measure categories/ characteristics serve to describe certain elements that characterize the PMLS in its structure, its didactical elements and its teaching staff. Finally, goal categories encompass various sub-categories of goals that either regard the individual learner or goals that the university may have in implementing the PMLS as a broader organization as well as a sub-category that allows to describe the quality of the goals. Not every sub-category has to be relevant in the description of every PMLS and some aspects of a PMLS may pertain to more than one category but the use of these categories aims to give an all-encompassing description of a PMLS's characteristics.

In the following, the use of the WiGeMath taxonomy shall be exemplified by applying it to one of the redesigned lectures.

CHARACTERIZATION OF REDESIGNED LECTURES

All in all, six redesigned lectures at five German universities were evaluated in the WiGeMath project. Out of these, four addressed preservice secondary teachers and two addressed engineering students. To reach a characterization of each lecture, an interview guided by the WiGeMath taxonomy was held with the teacher before the start of the semester. It was audiotaped and transcribed and the transcript was then used to name the measure's characteristics and sort them into the right categories of the WiGeMath taxonomy. What follows is the exemplary description of one of the evaluated lectures along the lines of the taxonomy's categories and sub-categories.

Frame conditions

Concerning the **characteristics of the student cohort**, those students that attended the given redesigned lecture were preservice secondary teachers in their first semester meaning that most of them were less than 20 years old and had just graduated from high school. The lecture has had round about 200 participants in each of its turns.

The **development of the lecture** officially started in 2011. Since then, there has not been a strict script, which is followed each year, but the different lecturers who have been responsible focused on different aspects. Nonetheless, the basis of the lecture always is the book by Grieser (2013) which deals with problem-solving strategies and proofs.

As to the **embeddedness of the lecture** in the wider system of university lectures, it is compulsory for preservice secondary teachers and voluntary for mathematics majors in their third semester. For preservice secondary teachers the lecture has substituted Linear Algebra as a first semester lecture though they still have to attend Linear Algebra in their second semester.

Staff conditions have been marked recently by problems to find qualified tutors to give the tutorials that support the lecture. As noted before, though there is only one lecturer per semester, it is not always the same one.

Finally, the **learning culture** is characterized by a strong focus on the students being active in their learning. They are supposed to try new methods and solve tasks during the lectures as well as during the tutorials. The concept of a “thinking pause” is very much enforced in the problem solving process. The lecturer gives some new input at the beginning of each class and collects and discusses results after the students have worked on problems or proofs.

Measure categories/characteristics

As to the **structural characteristics**, the measure consists of a lecture of two times 90 minutes per week with a tutorial of 90 minutes per week. One cycle of the measure starts at the beginning of each winter semester and finishes at its end (October through January).

The **didactical elements** include weekly homework and tutorial work of three to four exercises. All exercises may be worked on in groups. The solutions are discussed in the tutorials but no exemplary solutions are handed out. During the lecture there are phases of teacher talk, partner work and individual work. The book by Grieser (2013) is named as a reference text and can be accessed online on campus. At the end of the semester, a written exam concludes the course.

Concerning the **characteristics of the teaching staff**, the lecturer has his PhD in mathematics and is responsible for the contents of the lecture as well as the tutorials, the exercises and the final exam. The tutorials are given by six tutors who are students in higher semesters. These same tutors also have to correct the exercises which are handed in by students. As mentioned before, the selection of tutors proved difficult due to a small number of qualified applicants.

Goal categories

In the category of **individual learning goals** the measure focuses on activity-oriented rather than on knowledge-related goals: Both mathematical working strategies, like problem-solving strategies or use of examples and counter-examples, and learning strategies shall be improved. Attitudinal goals play a major role, as well: The measure aims at strengthening process beliefs and weakening toolbox beliefs in the sense of Grigutsch and Törner (1998) and wants to introduce the students into the mathematical professional community. Affective characteristics are to be influenced insofar as anxieties shall be lessened, interest and motivation shall be strengthened and the students shall gain a higher mathematical self-efficacy. Moreover, the measure wishes to let students recognize the relevance of its contents for further university studies.

The **system-related goals** include the preparation of the participants for their further university studies and the decrease of the number of dropouts. Besides, the measure wants to increase the quality of the feedback that students receive during their studies.

The **quality of the goals** as understood in the WiGeMath taxonomy is not to be understood in a normative sense but rather as a description of their substantiality. With this aim in mind, goals are examined concerning how specific, measurable, accepted, realistic and time-phased they are. Such a description of the quality of a goal would be done for every goal individually in a thorough analysis but due to space restrictions we will only focus on one specific goal in this paper to illustrate the point: One of the measure's goals is to improve affective characteristics of the participants, i.e. to lessen maths anxiety, increase motivation and interest and improve the participants' mathematical self-efficacy, in the course of the semester. This goal is specific to the point that it explicates what shall be achieved when by whom. It remains unspecific in naming why the goal is important, who holds the responsibility to reach it and which preconditions or limitations possibly exist. The goal is indirectly measurable through a survey directed at the students and it is accepted as it was named by the lecturer as a goal he wants to achieve rather than a goal he has to achieve due to orders given from above. The goal seems to be realistic to the point that the affective characteristics of the students seem to change for worse quite fast at the beginning of their university studies so it seems plausible that they may be changed for better even within a single semester. Still, this has to be checked as will be shown below. Finally, the goal is time-phased as it shall be reached within a limited time, namely the duration of the measure.

In this paper we will evaluate the following research question: To what extent was the described redesigned lecture successful in achieving the last-mentioned goal of influencing affective characteristics of the student cohort and in how far was the lecture successful in changing beliefs away from toolbox beliefs towards process beliefs?

METHODS

To measure the extent to which the above mentioned goals were reached, two questionnaire surveys were conducted with the participants of the described lecture. The first survey (t1) took place in the second week of the winter semester 2016 and the second one (t2) was conducted in the second to last week of the same semester. For each survey a questionnaire was developed, each one laid out to take about thirty minutes to complete. These questionnaires were handed out at the end of a lecture so that only those who were present that day could participate and participation was voluntary which the students were informed about. Moreover participation was anonymous but students used an individually constructed code so that the results of each participant in the first survey could be compared and contrasted to the results in the second survey. 163 students participated in the first survey and 103 in the second. We analyze the data of the 76 participants who answered both questionnaires.

We used adopted versions of the scale of Schiefele, Krapp, Wild and Winteler (1993) to measure interest, the scale of Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, (2002) to measure mathematical self-concept, a translation of the instrument of Longo, Gunz, Curtis and Farsides (2014) to measure the experience of competence, autonomy and social relatedness, an adopted version of the PISA 2000 instrument for self-efficacy (Kunter et al., 2002), a shortened version of the scales by Grigutsch and Törner (1998) to measure beliefs (application, process, system and toolbox) and a translated version of the academic motivation scale (Vallerand, Pelletier, Blais, Briere, Senecal and Vallieres, 1992) to measure different types of motivational regulation (intrinsic, identified, introjected and extrinsic). We used Likert scales ranging from 1 to 4 for self-concept, self-efficacy and beliefs, from 1 to 5 for motivational regulation, from 1 to 6 for interest and from 1 to 7 for the experience of competence, autonomy and social relatedness. All reliability coefficients (Cronbach's alpha) are acceptable or better, compare Table 1.

For each scale and each survey, a descriptive data analysis was conducted in order to get an overview of the results. Though the entirety of scales included more than the ones mentioned above, we concentrate on these only as our focus is to check to what extent affective characteristics of the student cohort, their experience of competence and their attitude towards different beliefs was changed in the course of the semester.

RESULTS

Table 1 shows the changes in mean values during the semester and effect sizes (Cohen's d) as well as p-values of paired t-tests.

Scale	Number of items	Cronbach's α		Mean value		Cohen's d	p-value
		t1	t2	t1	t2		
Interest for mathematics	9	.83	.83	4.23	3.96	0.32	**.001**
Mathematical self-concept	3	.81	.81	3.03	2.96	0.12	.149

Experience of competence	6	.80	.81	4.63	4.25	0.40	< **.001**
Experience of social relatedness	6	.85	.89	5.39	5.40	0.01	.900
Experience of autonomy	6	.73	.77	4.81	4.61	0.20	.132
Mathematical self-efficacy	4	.83	.87	2.72	2.66	0.10	.353
Application beliefs	4	.80	.88	3.01	3.02	0.02	.889
Process beliefs	4	.67	.85	3.26	3.18	0.12	.306
System beliefs	7	.79	.84	2.97	2.93	0.07	.534
Toolbox beliefs	5	.66	.74	2.75	2.56	0.34	**.002**
Intrinsic regulation	5	.88	.88	3.82	3.55	0.33	**.001**
Identified regulation	4	.72	.78	4.01	3.81	0.24	**.026**
Introjected regulation	4	.73	.78	2.04	2.19	0.18	.097
Extrinsic regulation	4	.64	.72	1.78	1.88	0.12	.278

Table 1: Scales and their Cronbach's alphas, means, effect sizes of changes between the two surveys and p-values for a significant change.

We see a substantial decline in interest and in the experience of competence, whereas students' mathematical self-concept and self-efficacy did not change significantly. We can also see that the objective of reducing toolbox beliefs was clearly achieved, but not the objective of strengthening process beliefs. The mean values of motivational regulation show that intrinsic and identified regulation are dominating although they are decreasing in the course of the semester.

DISCUSSION

The observation of a decline in interest is similar to the results showing a decline in traditional courses; however, student's mathematical self-concept does not change significantly, which is a major difference (Rach und Heinze, 2013, 2016). Although in our tests we were not able to show that the course could raise student's self-efficacy, it did not reduce it significantly either, which may still be an achievement. The dominance of intrinsic and identified motivation is a positive result as a study in traditional courses found extrinsic and introjected regulations to be dominant (Liebendörfer, in press). Thus, although students' interest in university mathematics and their intrinsic motivation may reduce, they do not seem to develop a stronger feeling of being inadequate for studying mathematics in the newly designed lecture. The decline in interest as well as intrinsic and identified regulation may be explained by a decline in the experience of competence. The change in students' toolbox beliefs is remarkable as beliefs are rather stable by definition and toolbox beliefs did not change in other studies in the first year of lower secondary or primary mathematics teacher education (Kolter, Liebendörfer & Schukajlow, 2016; Liebendörfer & Schukajlow, 2017).

These results show that a specifically designed lecture may reduce problems of the secondary-tertiary transition in mathematics. Nonetheless, the question remains whether such lectures prepare the students for their further studies just as well as traditional teaching does, considering that the course covered fewer mathematical topics.

Moreover, the analysis that has been done to this point cannot ensure that the results obtained were produced by the innovative measure alone. First of all, the lecturer's personality has an influence on the measure's outcomes that could not be separated from the outcomes of the measure itself in our study. A possible further effect may be caused by the change in order of other lectures, in explicit the postponement of Linear Algebra to a later semester: Whereas students usually experience their low competency in both Analysis and Linear Algebra in the first semester, in this case it is only one lecture.

In order to test this hypothesis, a next step in the WiGeMath project will be to distribute the same questionnaire that was used in the investigation described above to a different innovative measure at a university where different courses are attended simultaneously, as well as to a traditional lecture. This will make comparisons more explicit.

As to the taxonomy that was developed by the WiGeMath project, this in part resembles the objectives of other taxonomies (Krathwohl, 2002) though with a different focus. Whereas other taxonomies are mostly concerned with individual learning outcomes, the WiGeMath taxonomy targets a description and ensuing comparison of innovative measures as a whole. Though other taxonomies exist which classify systems of higher education institutions (for example the Carnegie Classification of Institutions of Higher Education, described in Bartelse & Vught, 2009), the perspective taken by WiGeMath to interpret such characteristics as goals is a new one.

So far, the taxonomy is only laid out to serve innovative measures and even in this area will have to be adapted as measures develop and improve. Traditional lectures have not been taken into consideration so far but we propose that these would also benefit from a similar taxonomy in terms of communicating frame conditions and learning goals. In our interviews with lecturers, we found that often even to them goals remained implicit until they were asked about them specifically. This might even more be the case in traditional lectures that have "worked" for a long time.

As mentioned above, in many cases goals stay hidden until a framework like the one developed by the WiGeMath project provides a common language to talk about them. Even though lecturers have specific intentions when they design a course with specific learning goals that a student cohort with certain characteristics shall achieve in a setting framed by staff conditions, learning culture, etc., they often lack guidelines to arrange these in a way that is comprehensible for others. Yet, only if they can explicate their ideas, can an evaluation be successful and show strengths as well as possible weaknesses of the designed course. In our example, the lecturer had his PhD in mathematics and had hardly been in contact with didactical theories and frameworks until the point of the WiGeMath evaluation. Hence, he had would not speak in terms of mathematical beliefs, for example. When the concept was explained to him, though, he clearly saw that one intention of the lecture was to change students' beliefs but to that point he simply lacked the vocabulary to explain this intention.

Our taxonomy will help to communicate goals between universities, staff and students as it provides a frame of reference and a common language as has been shown for one example in this text.

NOTES

1. The WiGeMath project is funded by the German Federal Ministry of Education and Research (BMBF), grant identifiers 01PB14015A and 01PB14015B.

REFERENCES

Bartelse, J., & Vught, F. van. (2009). The European Higher Education Classification: Objectives and Concepts. In F. van Vught (Ed.), *Mapping the Higher Education Landscape* (pp. 57–69). Springer Netherlands.

Chen, H.-T. (1990). *Theory-driven evaluations*. Newbury Park, California: Sage.

Colberg, C., Biehler, R., Hochmuth, R., Schaper, N., Liebendörfer, M., & Schürmann, M. (2016). Wirkung und Gelingensbedingungen von Unterstützungsmaßnahmen für mathematikbezogenes Lernen in der Studieneingangsphase. In *Beiträge zum Mathematikunterricht 2016* (pp. 213–216). Heidelberg: WTM-Verlag für wissenschaftliche Texte und Medien.

Grieser, D. (2013). *Mathematisches Problemlösen und Beweisen*. Wiesbaden: Springer.

Grigutsch, S., & Törner, G. (1998). *World views of mathematics held by university teachers of mathematics science* (Preprint Series No. 420). Duisburg: University of Duisburg, Department of Mathematics. Retrieved from http://duepublico.uni-duisburg-essen.de/go/duett-05272002-102811

Gueudet, G. (2008). Investigating the secondary–tertiary transition. *Educational Studies in Mathematics, 67*(3), 237–254.

Kalesse, I. (1997). Change in Mathematical Views of First Year University Students II. In M. Hannula (Ed.), *Current State of Research on Mathematical Beliefs V* (pp. 30–35). Helsinki: University of Helsinki, Dept. of Teacher Education.

Kolter, J., Liebendörfer, M., & Schukajlow, S. (2016). Mathe – nein danke? Interesse, Beliefs und Lernstrategien im Mathematikstudium bei

Grundschullehramtsstudierenden mit Pflichtfach. In A. Hoppenbrock, R. Biehler, R. Hochmuth, & H.-G. Rück (Eds.), *Lehren und Lernen von Mathematik in der Studieneingangsphase* (pp. 567–583). Wiesbaden: Springer Fachmedien Wiesbaden.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice, 41(4)*, 212–218.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., … Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Inst. für Bildungsforschung.

Liebendörfer, M., Hochmuth, R., Biehler, R., Schaper, N., Kuklinski, C., Khellaf, S., … Rothe, L. (2017). A framework for goal dimensions of mathematics learning support in universities. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10, February 1 – 5, 2017)*. Dublin: DCU Institute of Education & ERME.

Liebendörfer, M. (in press). *Motivationsentwicklung im Mathematikstudium.* Wiesbaden: Springer.

Liebendörfer, M., & Schukajlow, S. (2017). Interest development during the first year at university: do mathematical beliefs predict interest in mathematics? *ZDM, 49*(3), 355–366.

Longo, Y., Gunz, A., Curtis, G. J., & Farsides, T. (2014). Measuring Need Satisfaction and Frustration in Educational and Work Contexts: The Need Satisfaction and Frustration Scale (NSFS). *Journal of Happiness Studies*, *17*(1), 295–317.

Rach, S., & Heinze, A. (2013). Welche Studierenden sind im ersten Semester erfolgreich? *Journal Für Mathematik-Didaktik*, *34*(1), 121–147.

Rach, S., & Heinze, A. (2016). The Transition from School to University in Mathematics: Which Influence Do School-Related Variables Have? *International Journal of Science and Mathematics Education*, 1–21.

Ryan, R. M., & Deci, E. L. (2017). Identity Development, Self-Esteem, and Authenticity. In *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York: Guilford Publications.

Schiefele, U., Krapp, A., Wild, K. P., & Winteler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI). *Diagnostica 39*(4), 335–351.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts: SESSKO*. Göttingen: Hogrefe.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*(4), 1003–1017.