DEVELOPING STUDENTS' COMPUTER-SUPPORTED SIMULATION AND MODELLING COMPETENCIES BY MEANS OF CAREFULLY DESIGNED WORKING ENVIRONMENTS

Rolf Biehler[1] and Andreas Prömmel[2]
[1]Universität Paderborn, Germany
[2]Universität Kassel, Germany
biehler@math.upb.de

The GESIM material, developed in our research group, contains learning units with teacher guides for the first 4 weeks of a probability and statistics course at upper secondary school (age 17-18). The learning of subject matter is linked to the acquirement of simulation competencies and the acquisition of skills in handling the dynamic software FATHOM*. We emphasize simulation, law of large numbers and the role of sample size and develop knowledge about distribution from the beginning. We use a "simulation scheme" for guiding students' modelling and simulation activities. We videotaped the work of about 20 pairs of students working on simulation problems and we will present results of our analyses concerning the solution process and the resulting knowledge. Moreover, we will present data on students' competencies, which were measured by means of tests and by analyzing their written and their computer work.*

INTRODUCTION

In most German federal states students in grade 12 have to do a half-year course in *stochastics*. The course content typically comprises elementary discrete probability, Bayes' theorem, binomial distribution, hypothesis testing, and sometimes confidence intervals. Until recent students have not had much probability and statistics before grade 12, but things are changing due to the new national standards (KMK, 2004), which make the fundamental idea "data and chance" obligatory for younger students. Our working group has developed a German localization of FATHOM (Biehler et al., 2006) and is designing teaching material and teaching experiments for German schools for several levels. We do classrooms research for finding out how we can best use these materials in order to support the development of students' competencies. Our students use FATHOM for doing data analyses, constructing simulation and work with probability distributions.

THE CONTEXT OF THE PROJECT

In one of our projects, we developed a whole half-year stochastics course with the continuous use of FATHOM (Meyfarth, 2006). The design challenge was how to simultaneously develop three strands of competencies: the competencies in probability and statistics, the competence of doing simulations, and the competence of using FATHOM for simulating random experiments. Meyfarth's solution was to start the course with a so-called *simulation pre-course*, where in the first 15 lessons the students concentrated on learning to simulate a series of interesting probability problems with FATHOM. The function of the pre-course was to build and update probabilistic intuitions by simulation, to learn the typical steps of a simulation, and to acquire basic and stable FATHOM competencies the whole later course can build on. We redesigned the simulation pre-course using his results, developed our new GESIM concept (Biehler, Hofmann & Prömmel, 2008), and implemented and studied the use of the GESIM material in three courses. We will report some results of these studies in this paper.

DESIGN PRINCIPLES FOR THE SIMULATION COURSE WITH FATHOM

Among the problems for further improvements that were identified by the Meyfarth study are the following:

1. How can we make the learning of FATHOM more efficient?
2. How can we better support students in learning the typical basic steps of a simulation?
3. How can we improve the relation of students working in pairs with the computer and the classroom discussions where the new knowledge has to be institutionalized?

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

4. How can we improve worksheets in order that the students take the subtasks that require higher order thinking more serious and do not only work on the technical aspects of realizing a simulation in FATHOM?
5. How can we develop students' differentiated understanding about the relation of simulation results to theoretical probability?

Learning to use FATHOM. Obviously, classroom time is not enough for becoming fluent with the software. The usual teaching approach "the teacher shows how to use and the students imitate this use" isn't optimal and doesn't take care of the heterogeneity of the learning group. We designed eFATHOM (Hofmann, 2008) a multimedia e-learning tool for learning the basics of FATHOM. This tool was successfully used in students' homework time for self-regulated learning to use of FATHOM for data analysis and for simulation. The 4 modules of eFATHOM require about 4 hours learning time.

Learning to simulate. Many simulation problems can be solved by means of the same steps. Maxara (2009) and Maxara and Biehler (2006, 2007) made a theoretical analysis of FATHOM'S simulation capabilities and showed how these problems can easily be realized in FATHOM. We invented a "simulation scheme" by which the simulation activities of the students are guided. Students are shown worked-out examples (Renkl, 1997) of simulation problems and the simulation scheme is used to help students abstract the general procedure from the examples and work on new simulation problems.

Classroom discussions and work in pairs. We think of a typical lesson as consisting of the 4 phases we abbreviate with ISPD: *i*ntroduction, *s*tudents' work in small groups, *p*resentation of small group work, *d*iscussion. A complete design of instruction has to take care of all 4 phases. The teacher's role is very important in the I- and the D-phase. In the S-phase (s)he is a supporter of small group work, which is a different role. The P- and D-phases are most important for developing shared knowledge in the classroom, for merging individual approaches that emerge in the small group work. We designed a teacher's manual for this purpose.

Worksheets for guiding students' work. The design of activities that are not too open and not too closed is a difficult problem. In the Meyfarth study (2008), we found that students tend to skip or superficially work on those subtasks of a worksheet that do not require activities with the software, but activities concerning the planning of the work, about expressing expectations and envisioning results of the simulation to be done, and about interpreting results of the simulation. Whereas in the Meyfarth study these aspects often were not done in the S-phase, but rather in the P- and D- phases, we redesigned the worksheets to include content-related learning in the S-phase in a much more substantial way.

Estimation of probabilities by simulation and theoretical probability. In the previous studies, we found that students too easily equated probability with relative frequency of an event with high sample size *n* (Maxara, 2009). For improvements, we included tasks from the beginning, where combinatorial analysis could be combined with the simulation approach, We communicated rules of thumb for the precision of the simulation method for selected n = 50, 100, 1000, 5000, and 10.000, and we added learning environments at the end of the first part of the course, where students were supposed to learn that the standard deviation of the sampling distribution of the proportion of successes decreases with sample size n by 1/square root(n) as does the 95%-prediction interval for the proportion of successes. By these means the preliminary rules of thumb for the precision of results were upgraded as prediction intervals with 95% certainty.

STRUCTURE AND CONTENT OF THE INTRO-COURSE

We call the first 15 lessons (see Figure 1) of our half-year course the *intro-course.* In part 6 we posed the maternity ward problem (Kahneman & Tversky, 1973; Sedlmeier, 1998) and the tea-tasting-lady problem in several contexts. Part 6 consists of the 3 double lessons U8-9, U10-11 and U12-13. In U8-9 the simulation scheme was re-introduced based on experiences with module 4 of EFATHOM. In U10-11 the simulation scheme was used to solve two tasks in the context of the sound-quality problem. In U12-13 the simulation scheme was used for the 10-20-test problem. We will come to details of this scheme and the mentioned problems in the next chapter.

In part 7 we provided activities around the sampling distribution of the proportion of successes. Students were supposed to discover that the spread of the sampling distribution decreases with sample size *n*, namely as 1/square root (n). In sum, the innovative elements in block 6 and 7 were to treat problems around the binomial distribution only with simulation, give students' tasks around the relevance of sample size, discover the 1/square root (n)-law by simulation and introduce a hypothesis test problem early and informally, and calculate P-values by simulation.

	Topic	Quantity of lessons
1.	EFATHOM modules 1 and 2: Learning data analysis with FATHOM	Homework
2.	Doing data analysis with FATHOM: Empirical distributions and measures of spread and location (mean, median, quartiles, IQR)	2
3.	EFATHOM module 3: Learning to do simple simulations with FATHOM without collecting measures	Homework
4.	Laplace experiments and expected value in game situations (empirical and theoretical views, using simple simulations in FATHOM) Empirical law of large numbers, rules of thumb for precision of simulation depending on sample size	5
5.	EFATHOM module 4: simulation with sampling from boxes and with collecting measures in FATHOM	Homework
6.	Modelling of stochastic situations by box models, discovering the empirical counterpart of the binomial distribution	6
7.	Law of the large numbers, sampling distribution of proportion, 1/square root(n)-law)	2

Figure 1. *Intro-course* consisting of 15 lessons

THE RESEARCH DESIGN AND RESEARCH QUESTIONS

The GESIM concept was implemented in two courses at the Jacob-Grimm-School in Kassel and in on e course at the Albert-Schweitzer-School in Kassel. We collected data with several instruments: pre- and post-test on stochastic competence, observer protocols of all lessons, recordings (screen videos plus audio) from all small group working phases, final test of the course, in particular one simulation task.

We had research questions on several levels:

1. Is the introductory course teachable in the envisioned sense?
2. What are strength and weaknesses of the global structure of the course?
3. What are strength and weaknesses of the instructional design of the small group work?
4. How did the simulation scheme affect the work of the students on the simulation problems?
5. How can students' competencies with regard to the use of FATHOM, the construction of simulations, and the knowledge in stochastics be assessed and described?

SOME SELECTED RESULTS

Feasibility and global course structure. The use of EFATHOM in the homework time was successful and accepted by teachers and students. It contributed to much less variation in FATHOM competence and a more efficient use of classroom time than in previous teaching experiments. A problem was part 7 of the course. The assigned time was too short for a deep exposition and understanding of the topics of this block. Posing the question of the precision of the simulation depending on *n* was less motivating for the students than we expected, because the use of the software made it easy to choose very high numbers of repetition without any "cost". Block 6 was intended to be preparatory for the whole course. We have no clear evidence how the teachers referred to these experiences later in the half-year course. We got some feedback that a separate block 6 (and 7), which is to be placed later in the course together with the theoretical treatment of the binomial distribution, is preferred over a simulation rich intro-course at the beginning. This is more consistent with teachers' thinking about a reasonable structure of a stochastics course. They are not used to teach a topic twice, first on an informal level using simulation and then on a more formal level.

Students' competencies. To measure the students' learning outcome we administered different assessment instruments:

- pre-post-test to measure the statistical reasoning (items n = 7, students n = 39)
- final exam to measure simulation and modelling competencies (items n = 3, students n= 46).

We used the same test as pre-test and as post-test. The items referred to elementary stochastic questions from the following domains: frequency, independence, sample space and symmetry, distribution. In the pre-post-test the students had to make first a choice and then they had to give reasons for their choice. For the correct choice they got 1 score point and for the reasons 2 score points at maximum. The maximum score the students could get was 19 points in each test. The students got 9.41 score points (49.5%) on average in the pre-test and 13.12 score points (69,1%) in the post-test. The increase in learning outcome was best in item 6, the maternity ward problem. We used the following formulation of this problem: "In a large hospital each week about 90 children are born on average. In a small hospital each week about 40 children are born on average. At which hospital is it more probable that in one week more than 65% of the born children are boys: at the large hospital, at the small hospital or the same probability at both?" In this item, the students got an average score of 0.72 of in the pre-test and 2.08 (maximum = 3) in the post-test.

The final examination was made three weeks after the end of the intro-course. The students had to deal with the maternity ward problem like the one above. They had (a) to fill out a simulation scheme, (b) sketch a distribution for assessing their distributional knowledge and give reasons for their decisions. Moreover, the students should (c) calculate the middle 95% interval of this distribution by means of the 1/square root (n) – law. The students got 6.14 (55.8% of maximum) score points for this task.

In Figure 2 we see a student's answer of the subtask (b), which is representative for the majority of solutions: The bars for the large hospital are drawn higher in the center and lower in the tails. The number of the bars remains all about the same. The relative frequency of the center is underestimated.

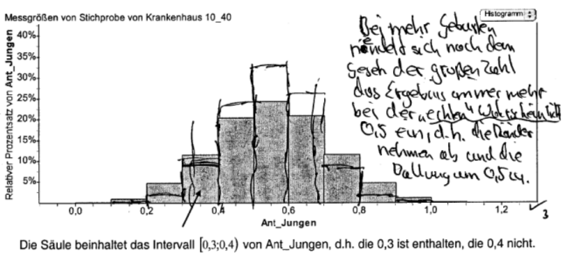

Figure 2. From a student's work

THE USE OF THE SIMULATION SCHEME

In this paragraph, we will focus on the use of our *simulation scheme* or *simulation plan*. Many simulation problems can be solved with the same structure of steps and each step corresponds to specific operations to be done in FATHOM (Maxara, 2009). Based on this analysis we presented worked-out examples showing these steps in EFATHOM (Hofmann, 2008) and developed a form to be filled out for the practical work of constructing simulations in FATHOM. The idea of developing such a form was born in reaction to the observation that some students had difficulties abstracting the scheme and its steps and to remember them well enough in the simulation process.

In Figure 3 a filled-out form is shown for the following problem that was posed in lesson U10-11: "A test person gets 12 pieces of music, has to identify whether it has CD-quality or MP3-quality and has won a prize, if (s)he gets at least 8 correct. With which probability does someone get a prize by just guessing the sound quality (CD or MP3)?". Students had to solve this problem with the complex simulation method "sampling from a box".

The structure of the simulation scheme follows the stages of the general simulation plan (Biehler & Maxara, 2007). The table is split into two columns. The left grey column shows the individual steps for the simulation method "sampling of a box" like a guideline. The right column provides a predefined structure. When filling out the simulation scheme, students have to consider specific properties of FATHOM as well as theoretical aspects. Especially step [3] - „description" and „possible values of the measure" - requires theoretical considerations. Such content-related considerations should help to prevent a schematic handling and documentation of FATHOM terms.

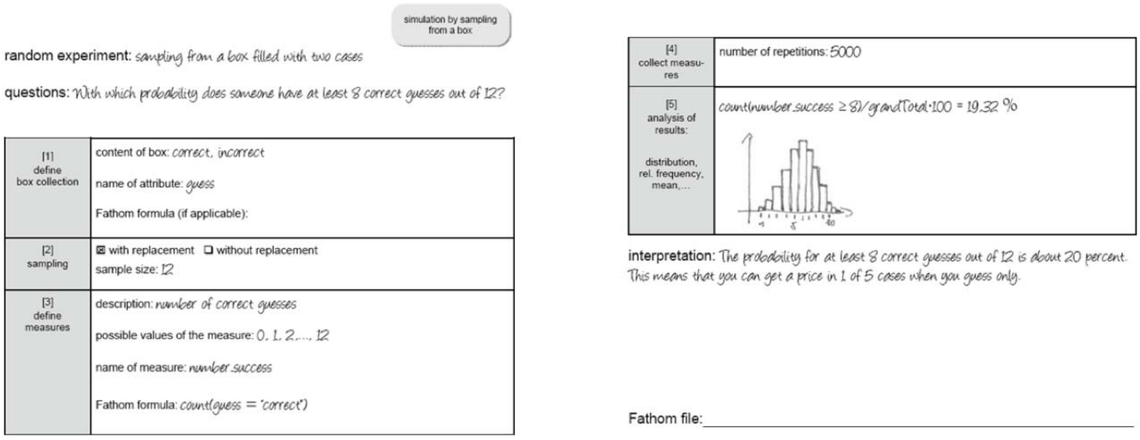

Figure 3. Filled-out simulation scheme (split up in two parts)

The scheme can have several functions for the work of the students. We distinguish three basic functions: *planning, guiding, documentation*, which broadly corresponds to a use before, during and after the work with the software. Use for documentation means that the entries of the FATHOM actions are inserted into the form after their implementation in FATHOM. Maxara's (2009) study showed that the time students needed for documentation of a simulation was in general significantly higher than for doing the simulation. We expected that our form reduces cognitive load and lets the students concentrate on the steps instead of generating a reporting scheme of their own.

The simulation scheme has a guiding function by its structure and its individual simulation steps. All essential steps are given and must successively be processed by students. Therefore it serves as a guide and supports students' acquisition of this complex simulation method.

The scheme can also be used in a preparatory *planning* phase before working with the computer: one or several steps are planned first before they are realized in FATHOM. The separation between planning and implementation is often recommended when teaching computer programming. Intertwining technical issues with conceptual ones may be much less efficient than separating these two stages. Moreover, a separation could support a more reflective collaborative approach for a pair of students as compared to working predominantly on technical aspects. The study by Meyfarth (2008) and studies on programming in mathematics education (Krummheuer, 1989) show this danger of computer work.

For research purposes, we implemented two ways of using the simulation scheme: *consecutive* use–scheme is filled out in collaborative work before the working at the computer starts and *integrative* use–scheme is filled out while the work at the computer is carried out. In the consecutive use the pairs of students fill out the simulation scheme so far as possible without computer use. After that they implement their simulation plan in FATHOM. In the integrative use there is no separation from planning phase and implementation phase.

We evaluated the FATHOM-files and the filled simulation scheme as products that were produced in the lesson U10-11. To examine the solution quality we analyzed all steps in the FATHOM files and also the according simulation schemes the students provided on the base of a master solution. Every step got 1 score point if it was correct, else 0 score points. The maximum score is 18 points in each task and 14 points in each simulation scheme. The *consecutive pairs* got 16.75 score points on average for the FATHOM-files and 11.325 score points on average for the simulation schemes. The *integrative pairs* got 16.30 score points on average for the FATHOM-files and 11.25 score points on average for the simulation schemes. This result shows that the work on the examined tasks was done in both groups of pairs very well. The simulation scheme was accepted and it helped students in solving the simulation tasks well.

At first it therefore seemed that the separation in planning phase and computer working phase has no effects on the learning success and the product quality. However, some differences arose in the solution *process*, which we wanted to analyze in more detail. The data support our hypothesis that the pairs of students with a consecutive use of the scheme show a more uniform solution process in the computer work and need less assistance when they had a planning phase. For this analysis, we had coded the FATHOM activities in the transcripts with a colour code along the steps of the simulation scheme. We got so-called text portraits that make the different solution processes of the pairs in both the groups evident. Then we coded *problematic phases* and *requests of assistance* in the computer working phase of the pairs and counted them. The pairs of the *consecutive group* had 27 problematic phases and 29 requests of assistance in total. The pairs of the *integrative group* had 59 problematic phases and 61 requests of assistance in total. In order to better understand the learning processes of the individual pairs, we are currently also analyzing the dialogues of the pairs during the solution process. First results show significant differences in style of collaboration. Trial and error strategies seem to be more frequent in the integrative group.

CONCLUSION

The GESIM concept required specific learning efforts from the students, because it includes the acquisition of tool competencies, simulation competencies and stochastic competencies. The students very well accepted the simulation scheme. For novices this is a helpful instrument for planning, guiding and documentation. We have found some evidence that separating the planning phase from the computer implementation has some advantages. The gain in competence in probability and statistics was satisfactory, but the retrospective analysis pointed also to some deficiencies of our instructional design.

REFERENCES

Biehler, R., Hofmann, T., Maxara, C., & Prömmel, A. (2006). *Fathom 2. Eine Einführung.* Berlin: Springer

Biehler, R., & Hofmann, T., & Prömmel, A. (2008). *GESIM: Ganzheitlicher Einstieg in die Stochastik der Sekundarstufe II mit Hilfe von Simulationen unter Einsatz der Software* FATHOM. *Didaktisch kommentierte Materialien für den Unterricht.* Universität Kassel.

Hofmann, T. (2008). EFATHOM. eLearning-Umgebung für den Einstieg in FATHOM. Online: www.mathematik.uni-kassel.de/~luf/.

Kahneman, D. & Tversky, A. (1973). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology 3*, 430-454.

KMK (Ed.) (2004). Bildungsstandards für das Fach Mathematik für den Mittleren Bildungsabschluss: Beschluss der Kultusministerkonferenz vom 4.12.2003. München: Wolters Kluwer.

Krummheuer, G. (1989). *Die menschliche Seite am Computer Studien zum gewohnheitsmäßigen Umgang mit Computern im Unterricht*. Weinheim: Dt. Studien-Verl.

Maxara, C., & Biehler, R. (2006). Students' probabilistic simulation and modeling competence after a computer intensive elementary course in statistics and probability. In A. Rossman & B. Chance (Ed.), *Working Cooperatively in Statistics Education. 7th International Conference on Teaching Statistics.* Salvador da Bahia, Brazil, 2006 (pp. 1–6). Online: www.stat.auckland.ac.nz/~iase/publications/17/7C1_MAXA.pdf.

Maxara, C., & Biehler, R. (2007). Constructing Stochastic Simulations with a Computer Tool – Students' Competencies and Difficulties. In *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education, 2007 (pp. 1–13).* Online: ermeweb.free.fr/CERME%205/WG5/5_Maxara.pdf.

Maxara, C. (2009). *Stochastische Simulation von Zufallsexperimenten mit* FATHOM*–eine theoretische Werkzeuganalyse und explorative Fallstudie.* Dissertation. Universität Kassel.

Meyfarth, T. (2006). *Ein computergestütztes Kurskonzept für den Stochastik-Leistungskurs mit kontinuierlicher Verwendung der Software* FATHOM. Kassel: Universität Kassel. Online: http://nbn-resolving.org/urn:nbn:de:hebis:34-2006092214683.

Meyfarth, T. (2008). *Die Konzeption, Durchführung und Analyse eines simulationsintensiven Einstiegs in das Kurshalbjahr Stochastik der gymnasialen Oberstufe – Eine explorative Entwicklungsstudie* Hildesheim, Berlin: Franzbecker. Online: http://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2006100414792.

Renkl, A. (1997). Learning from Worked-Out Examples: A Study on Individual Differences. *Cognitive Science, 21(I),* 1-29.

Sedlmeier, P. (1998). The Distribution Matters: Two Types of Sample-Size Tasks. *Journal of Behavioral Decision Making, 11,* 281-301.