Designing and evaluating an e-learning environment for supporting students' problem-oriented use of statistical tool software

Biehler, Rolf
Universitaet Paderborn
Warburger Str. 100
Paderborn, 33098, Germany
E-mail: biehler@math.upb.de

Hofmann, Tobias
Universitaet Paderborn
Warburger Str. 100
Paderborn, 33098, Germany
E-mail: toho@math.upb.de

Introduction

This paper discusses technology use in introductory statistics education. We use the multi-purpose tool software FATHOM for supporting students' learning in various aspects. We created an eLearning environment EFATHOM, by which students are supported in their initial learning to use FATHOM. After having briefly introduced FATHOM, we will describe the empirical studies and the theoretical background and assumptions from which EFATHOM was designed. The structure of the content and the underlying theory of instructional design are presented as well as results from various evaluation and acceptance studies.

Technology support in introductory statistics education

When we browse through existing courses and related textbooks, we can distinguish several types of using technology in introductory statistics. Some courses come with a system of stand-alone applets providing interactive experiments, visualizations and simulations. Others integrate a data analysis tool and provide exemplary data sets. Sometimes data sets are provided in several formats. Students are supposed to learn a tool alongside with the course and use it for applying statistical methods on data as part of the goals of such a course. A third option is the integration of only one tool that also plays the role of "hosting" experiments, visualizations and simulations. Many tools such as traditional statistical packages such as SPSS do not support this very well. Minitab, EXCEL and R are examples of tools that are often used for creating and using interactive experiments in addition to their use for data analysis. The adaptability and programmability of these tools alongside with their easy-to-use simulation capabilities is the basis for creating interactive experiments similar in function to applets. The advantage of such "embedded experiments" is that teachers and students can adapt these experiments using their knowledge of the software. There is no need of learning to program applets.

However, Minitab, EXCEL and R have several disadvantages to serve as a tool that optimally supports learning and doing statistics, especially at school but even at university level. Based on the belief that statistics educators should not just live with and adapt to tools that are provided by the software industry, the first author developed requirements for a tool that is able to support learning and doing statistics in introductory statistics courses at school and university level (Biehler, 1997). Since several years, the software FATHOM (www.keypress.com) is available. It fulfills many of the requirements and surpasses them in various respects, based on the most recent technology.

We have used FATHOM in various courses at university and school level. We created a German

localization of FATHOM and developed materials to support the use of the software (Biehler et al, 2006; 2011). We aim at designing adequate learning trajectories where learning statistics progresses hand in hand with learning this software to practice statistics and to do interactive experiments for exploring statistical methods and graphs. From the literature in mathematics education it is clear that such a process is very complex even if we use a tool that is relatively easy to learn. Researchers speak of the process of instrumental genesis: The "tool" has to be acquired as an "instrument" by every student for working on certain problems (Trouche, 2004). The classrooms teachers and educational designers have to think of the complex process of orchestration of integrating technology. We have done empirical research where we observed and evaluated the use of FATHOM in classrooms, in small group work and in individual work of students. We used pre-and post tests to analyze achievements and we used the Camtasia screen recording program together with video cameras to record and analyze students' working on statistical problems with FATHOM (Meyfarth, 2008; Maxara & Biehler, 2006). The findings showed, among other things, that and how the process of instrumental genesis, i.e. the students' learning process to use FATHOM for statistical problem solving can be better supported. That was the reason for creating the multimedia learning environment EFATHOM, whereby students can learn the basic uses of FATHOM. We think that the design principles we used are of general interest for those who intend to design, create and evaluate a learning environment for supporting the process of instrumental genesis of any statistical tool. A more detailed account of the theory-based design principles and the evaluation studies will be published as part the Ph.D. thesis of the second author (Hofmann, 2011).

The software FATHOM as a tool for learning and doing statistics

One feature of FATHOM that makes it easy to learn and become productive is the ability to *drag-and-drop* variables into a graph (Fig. 1, left).
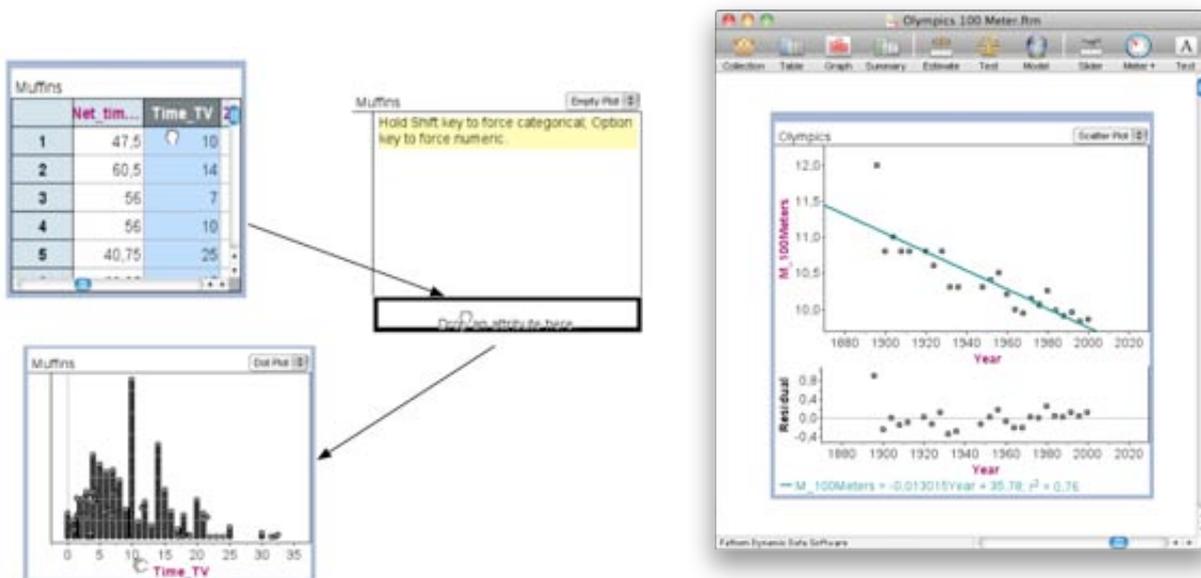
Figure 1. left: Dragging an attribute into a graph in FATHOM; right: Dragging the point at the year 1896 changes the least-squares fit and the residuals

Multiple graphical displays can be created and arranged on the screen. In addition "result tables" can be added to the screen, where numerical statistics can be displayed for those attributes that are dragged into a result table. Interactive multiple linked representation make explorations easily possible. The dynamic changes can be evoked by dragging points or by changing sliders. In Fig. 1 (right) we display the results of Olympic games of 100m race for men. We have fitted a least-squares line and observe residuals with an outlier on a slight upward trend. We can explore how the point at the year 1896 will change the least-squares

line and the respective residuals in the residual plot that is dynamically linked to the scatterplot and the recalculated line.

FATHOM can also be easily used to create simulations. We will describe the feature on the basis of the step-wise "simulation scheme" that we also use for supporting students (Maxara & Biehler, 2006).

1. Construct the model for the (single or multiple stage) random experiment
2. Identify events and random variables of interest
3. Repeat the model experiment and collect data on events and random variables
4. Analyze data: relative frequency (events); empirical distribution (random variables)

We will use the famous hospital problem adapted from Tversky & Kahneman (1974) for demonstrating how to proceed. Question: In a small hospital A about 40 children are born per month, in a large hospital B about 100 children are born per month. In which hospital is it more likely to get a month with at least 60% boys: A, B, or both are equally likely? Most people do not recognize the effect of sample size, which makes the event less likely in large samples that is in the large hospital.
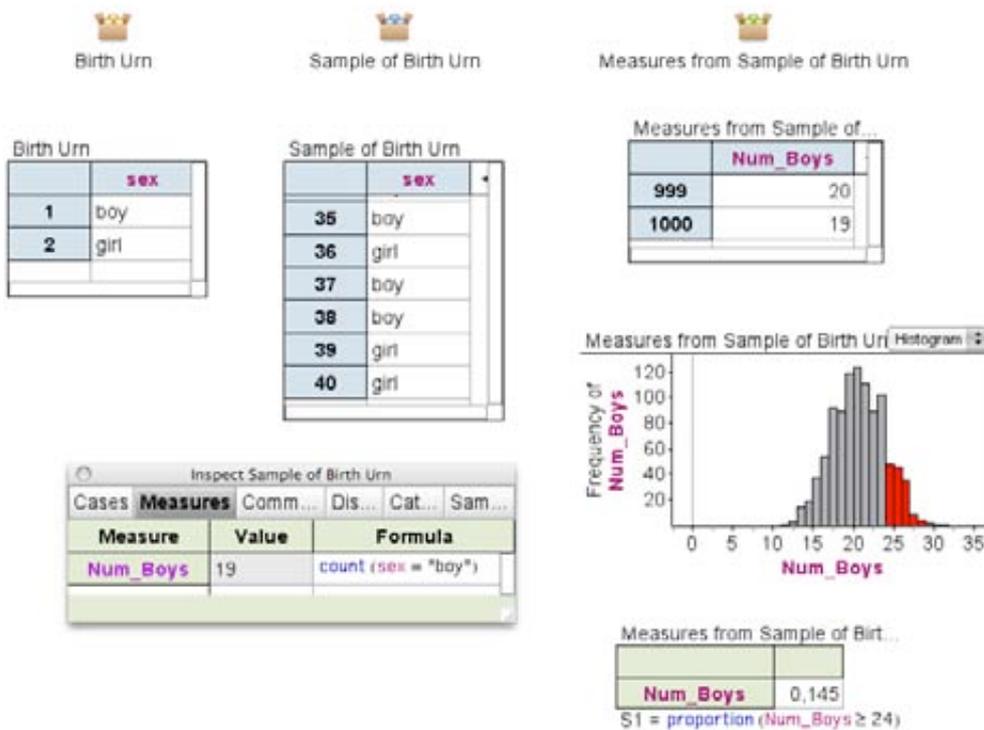
***Figure 2 Simulation of the "hospital problem" with** FATHOM*

Fig. 2 shows the simulation in FATHOM. Data are kept in "collections" whose content is shown in the table below each collection. This simulation has been created according to the following steps of our simulation scheme:

1. Model the random experiment. We create a "birth urn" with one attribute named "sex" and two cases "boy" and girl". Urn is a European synonym of box, which is often used in teaching probability and statistics (e.g. Heitele, 1975). We simulate the births of 40 children by sampling with replacement from this "birth urn". The results are put into the "sample of urn" collection.
2. Identify vents and random variables of interest. We define the number of boys as the random variables of interest: A "measure" named Num_Boys is created by the formula count (sex="boy").
3. We repeat the model experiment 1000 times (command "collect measures"). Results are collected in the collection "measures of sample of birth urn".
4. Analyze data: A histogram shows the distribution and highlights those cases (months) that have 24

or more boys. A result table calculates the respective proportion of months (14,5 %).

We can repeat the same process with the large hospital (100 boys) and have to calculate the proportion of months where Num_Boys is at least 60. This will be the case in about 0,1 % of all months. The construction of the first simulation uses the general FATHOM formula editor at several places. The modification to n = 100 is easily done and all the results are dynamically updated so that a comparison can be made. Similarly simulation can be used to generate all kinds of sampling distributions, and introduce p-value based hypothesis testing by simulation. A command called "scramble attribute values" re-randomizes the values of one attribute while keeping values of all other attributes constant. When we automatically collect measures of interest for instance the mean difference of two subgroups, we can create the reference distribution of the randomization test and judge the statistical significance of the difference.

Resources for and problems of introducing FATHOM to students

„The best way to learn the mechanics of FATHOM is to have somebody show you. You can learn the basics in just a few minutes. These basics include entering data, making graphs to plot your data points, and putting lines on the graphs. If you don't have an instructor handy, you can probably figure it out on your own." (Erickson, 2004, p. 10) .

For further reading, Erickson suggests material with a short written introduction to using FATHOM and further books that show how statistical problems can be worked on with FATHOM. The German localization of FATHOM was originally provided with an accompanying book structured according to application problems such as graphing data, comparing groups, simulation and hypothesis testing etc (Biehler et al., 2006), where all problem solution were demonstrated step-by-step. Moreover the German version of the FATHOM help system was a further resource. This problem-oriented introduction is very different from a traditional manual, because we were convinced that typical manuals are not very efficient. However, our book requires very careful and concentrated reading with parallel experimenting with the software. It was mainly aimed at teachers and instructors and proved to be not so efficient with students. In our classroom experiments we usually started with a personal introduction by an expert in front of the students with subsequent imitations by the students at the computer. Short quick reference manuals for specific purposes were created. We designed specific systems of tasks by which the students were supposed to elaborate their expertise with FATHOM with regard to certain types of data analysis and simulation problems.

The teacher-as-expert-presenter-and-student-as-imitator model of introducing a software tool has serious drawbacks. The heterogeneity of the learning group with regard to statistical competence and computer handling competence is usually large. Moreover the learning speed and learning curve of individual learners when working with the software for the first time is also widely spread out. The teacher as presenter is moreover unable to walk across the classroom and give individual help. Tutorial videos where a virtual expert takes over the role of the teacher in demonstrating steps of using the software promise to be superior if the videos can be paused, rewound and speeded up according to individual needs and when they are designed so that learners can work with the software FATHOM in parallel.

We thought that tutorial videos can be used in various educational scenarios:

Scenario 1: stand-alone introduction without any teacher,

Scenario 2: part of students' homework as preparation for the use of the software in classrooms,

Scenario 3: part of classroom teaching that enables the teacher to effectively play the role of a pedagogical facilitator, because time becomes available for providing individual feedback to students.

Applicability in scenario 1 is important for facilitating the first and second steps into the software for new users of FATHOM.

We found that a system of tutorial videos alone is not sufficiently supportive. Rather we aimed at learning units with integrated video tutorials, but also with other elements such as tasks to check ones knowledge and skills, and text elements that explain the background of software features.

Moreover, FATHOM poses several very specific challenges to a learner. Although the drag-and-drop

features are very easy to use after having been learned, this interface design is uncommon for many other software pieces running under the Microsoft Windows environment. Learners have to be introduced into these specificities. The dynamic features FATHOM can also be best demonstrated in a video as compared to just talking about them. A typical use of FATHOM is a multi-step process of interacting with data, which includes the production of various result graphs and tables and the re-organization of the desktop. For learning such multi-step procedures, the introduction has to be very different from just showing how one can get such and such a statistical display in one or two steps. This procedural complexity is particularly distinctive for the process of simulation.

Overview of the content structure of the eLearning system EFATHOM

The system is structured according to the following four modules that are based on each other:

1. Introduction to the software and first steps into data analysis
2. Tools for analysing data
3. Simple simulations
4. Simulation by sampling from a virtual urn

We pay attention to the principles of "instrumental genesis" in that each module does not only address software features in isolation from using them on statistical tasks. Moreover, we do not yet assume that learners are familiar with statistical concepts and methods but we introduce some of them in the modules. For instance, we suggest certain methods for doing a group comparison to our learners together with introducing a FATHOM feature how to easily split the distribution of an attribute into several distributions for each category of a categorical variable. Or, the method of simulation is accompanied by explaining the empirical law of large numbers to students and by presenting rules of thumb about how accurate and certain it is to estimate a probability from an observed relative frequency.

Module 1 introduces the general structure of the FATHOM desktop and how to input data, how data sets are organized in collections and tables with attributes and cases, and how they can be visualized in graphs. Module 2 concentrates on how to modify, enhance, and interact with graphs, and how to exploit their property of being linked with other graphs. Additionally, the FATHOM formula editor is introduced and used for applying and creating user-defined formulas for obtaining numerical results.

In module 3, learners can learn to simulate throwing one or two dice according to the simulation scheme we introduced above. Learners learn to play with the number of repetitions, observe variability and stabilization of the results and use relative frequencies to answer a few simple probability questions in the two dice context. In module 4, the powerful but more complex method of simulation by sampling from urns – as illustrated above - is introduced. In a first step, learners create a virtual urn, from which they draw n-times with or without replacement. In the next step, they define events or random variables of interest and repeatedly collect "measures" with N repetitions for getting estimates for the relevant probabilities.

Module 1 and 2 are essential for starting up using the software. We have used these modules with students from grade 8 up to college level. With these modules they are prepared to use the software for elementary data analysis. Module 3 provides a basic background to simple simulation (from grade 9 onwards). Although module 4 is also designed to work with 9-graders, we have used it more often with older students. The pay-off of learning, practicing and applying simulation by sampling from urns is higher in upper secondary grades or at university level, when there is more time to learn advanced concepts in probability and statistics.

Simulation scheme for sampling from an urn

Among the various innovations in EFATHOM, a new simulation scheme and simulation overview was developed for providing more orientation for the students. This development was based on studies on how students simulated and used earlier simulation plans (Maxara & Biehler, 2006; Meyfarth, 2008).
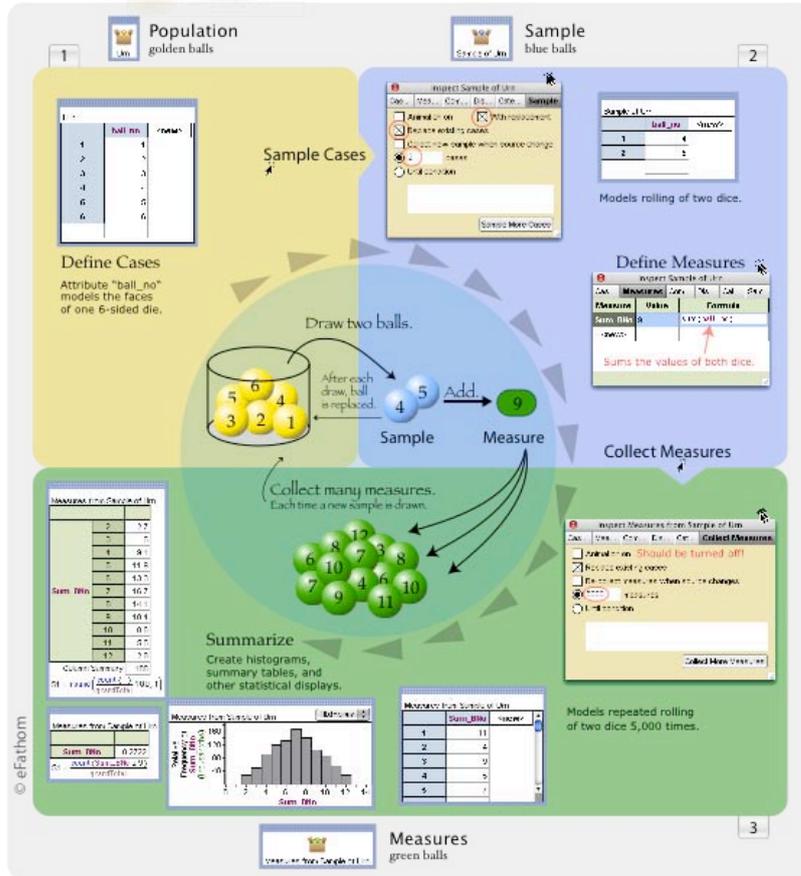
Figure 3 Simulation overview for the process of sampling from an urn

Fig. 3 represents the step-wise process of a simulation by sampling from an urn. The three relevant collections of data are distinguished by different colours. The interactions that FATHOM offers at various places are included as well as potential final and intermediate results. This picture is supposed to help students not to loose orientation with regard to a complex screen.

Instructional-theory based design of EFATHOM

A basic idea for working with EFATHOM is to work with a split screen, where EFATHOM is running in the left part and FATHOM is running in the right part of a computer screen. Fig. 4 shows such an arrangement. In the left, a tutorial video is shown on how to simulate two dice and how to estimate the probability that two dice show the same numbers at a time. The right part shows how the learner has realized this simulation in FATHOM. A learner can flexibly switch his/her attention from EFATHOM to FATHOM and work forwards and backwards in both systems while being able to relate and overview content and activities simultaneously in both systems. The tutorial videos can be paused and rewound if a learner has got stuck in implementing a FATHOM process.
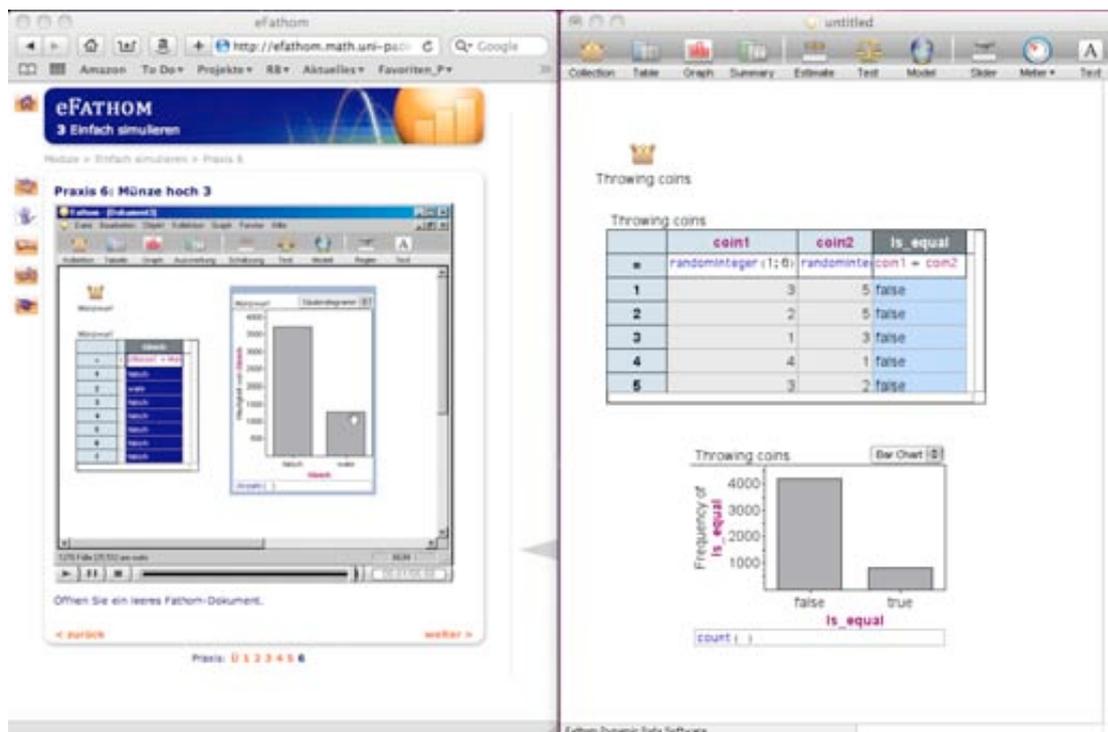
***Figure 4 Split screen arrangement of* FATHOM and EFATHOM**

Each EFATHOM module consists of the following elements: *introduction, practice, knowledge, tasks and applications, check-up,* which we call the macrostructure of a module. The microstructure of each component is based on theories of knowledge acquisition with digital media such as cognitive-load theory (Chandler & Sweller, 1991; Sweller, 2005), the cognitive theory of multimedia learning (Mayer, 2001, 2005) and the minimalist instruction principles of Carroll (1990). We combine these theory elements with results from our implementation studies with FATHOM (Meyfarth, 2008) and a theoretical tool analysis of FATHOM by Maxara (2009), which was based on the theory of instrumental genesis, which resulted in differentiated knowledge with regard to the constraints and affordances of FATHOM as a cognitive tool for doing simulation and data analysis.

Introduction: Contains explicitly stated goals in terms of software competence and in terms of statistical competence for guaranteeing "goal transparency" for the learner.

Practice: Several short tutorial videos (2-5 minutes) present exemplary cases for working with FATHOM that the learner is supposed to observe and then repeat himself/herself in the software. The aim is to build practical procedural knowledge and skills. The design is based on the "theory of worked examples" with self-explanations, which are provided by the virtual tutor in our case (Paas & van Merriënboer, 1994). The idea was that students watch a short tutorial video completely before they start to implement the process in the software, in order to get a holistic picture of the whole process first. Implementing such an observed process into the software was supposed to be the second step, where they can go back selectively to the video if they get stuck. Subsequently, EFATHOM provides further tasks to practice and to vary the process on analogous examples.

Knowledge: This component summarizes, generalizes, reflects and conceptualizes the implicit knowledge of the preceding practice component by means of further examples and static illustrations. For instance, a table of the width of the interval that covers the middle 95% in a sampling distribution of a proportion is given (adding statistical competence) or a table of useful FATHOM formulas for defining random experiments or events and random variables of interest (adding software competence).

Tasks and applications: Interactive exercises are offered, partly with feedback adapted to the intermediate answer of the learner. Some tasks require doing a complex FATHOM process and we offer a

system of hints, starting from minimal hints to more advanced hints for scaffolding the learning process. Learners can practice and routinize their practical skills and diagnose whether they have successfully learned.

Check-up: Tests are offered for diagnosis of knowledge elements.

Summary of the evaluation studies

We have done several empirical studies concerning the use and acceptance of EFATHOM. The early studies were directed towards formative evaluation. We were interested in using these results for improving EFATHOM. Students' technical and navigation problems with earlier versions disappeared in later versions of the software. Later studies were directed towards the following questions:

1. How long do students need for working on the modules?
2. How do students work with the modules in detail?
3. How is the acceptance of EFATHOM as a supportive tool for school and university students?
4. To which extent did the students acquire the FATHOM competence and the statistical competence that was intended?
5. How does EFATHOM compare to traditional ways of introducing into FATHOM?

We will selectively describe some of the results. A detailed account describing goals, methods and results of all studies can be found in Hofmann (2011).

Ad 1 Duration of use) We made a case study on the use of EFATHOM with n = 122 students, asking them how long they worked with EFATHOM. The median working time was about 45 minutes for modules 1 and 2 and about 60 minutes for modules 3 and 4. The interquartile ranges were lying between 20 and 40 minutes. This median time span is a feasible amount of time for our implementation contexts at school and university level. The spread we observed confirms the heterogeneity of students learning speed and time they intend to invest in working on a module.

Ad 2 Detailed usage) In two case studies with 19 school students in total, we recorded the interaction with EFATHOM and FATHOM using the screen recoding software Camtasia. Most students proceeded through EFATHOM linearly page by page with little going back and forth, and with little skipping of pages. We found surprising results with regard to the use of the tutorial videos. Contrary to what we expected, students very rarely waited until the end of a video before they started to implement what was being shown in FATHOM. Nearly all students worked simultaneously in FATHOM while watching the EFATHOM videos. The groups differed with regard to how often and how long they paused the video. The reason why we recommended sequential work was the following. We conjectured that listening and looking at the videos and working simultaneously in FATHOM will lead to cognitive overload and will prevent students from focusing on the overall goal and structure of a process. The tendency of just working step-by-step was conjectured to be substantial. It would be interesting for a future study to explore in more depth whether the simultaneous style of working has really the conjectured disadvantages. The current study just provides descriptive results only.

Ad 3 Acceptance) Part of the acceptance study was done with university students how participated in a course "elementary probability and statistics". The sample consisted of 63 students of Kassel University in November 2008 and 350 students at Paderborn University in April 2009. In both introductory courses students were required to learn FATHOM as part of the course. Homework assignments included tasks where FATHOM had to be used for doing data analyses or simulating random experiments. EFATHOM was used in the first two weeks. We asked the students to fill out a questionnaire on their evaluation of working with EFATHOM. We used a 5 point Likert scale and asked for the degree of agreement with a statement. The same questionnaire was also used in a sample of 39 school students from grade 12, where we made a design-based experiment introducing school students simultaneously into FATHOM and probability and statistics (Biehler & Prömmel, 2010). The results that are summarized in Table 2 show a very positive feedback.

Item eFathom ...	Mean score university students	Mean score school students
... is a good introduction to Fathom	4,44	4,25
... is efficient introduction to Fathom	4,33	4,08
... provides orientation for the user	3,96	4,30
... has intuitive navigation	3,60	3,84
... allows quick navigation	3,69	4,05
... has clearly arranged pages	4,03	4,35
... has appealing pages	3,70	3,70
... has reasonable structure	4,29	3,95
Split screen arrangement is helpful	4,32	3,89
Clearly stated module goals	4,21	3,92
Tutorial videos are understandable	4,69	4,43
Working with eFathom is fun	3,59	3,46

Table 1 Results from the acceptance study (Likert scale from 0 to 5 (no to full acceptance))

Ad 4 Resulting competence) This question was difficult to do research on, because usually the learning conditions of students cannot be controlled and we cannot isolate the effect of EFATHOM. We observed a small number of students (n = 7) working with EFATHOM and recorded their activities with Camtasia. Practically all students were able to solve the problems in the element task & application successfully. Moreover, they were able to successfully work on simple simulation problems after having used EFATHOM. In the study reported on by Biehler & Prömmel (2010) EFATHOM was used in the first week of the teaching sequence. The classroom notes provided anecdotic evidence that the heterogeneity of FATHOM competence was substantially reduced and learning time in the classroom could be devoted more to developing statistical competence of the students instead of software competence. The final test after 4 weeks, where one part of the assessment asked students to solve a simulation problem with FATHOM showed excellent results of the whole approach where the use EFATHOM was one of the components.

Ad 5 Comparing EFATHOM with printed material) We divided a subgroup of participants of our course "Elementary probability and statistics" (n = 40) randomly into two groups for one week. Group A (the EFATHOM - group) used modules 3 and 4 from EFATHOM for that purpose. Group B (the print group) had to use the chapter of Biehler et al. (2006) for self-regulated learning to use FATHOM for simulation. As these were homework assignments we were not able to completely control the learning conditions. However EFATHOM was not available to the print group.
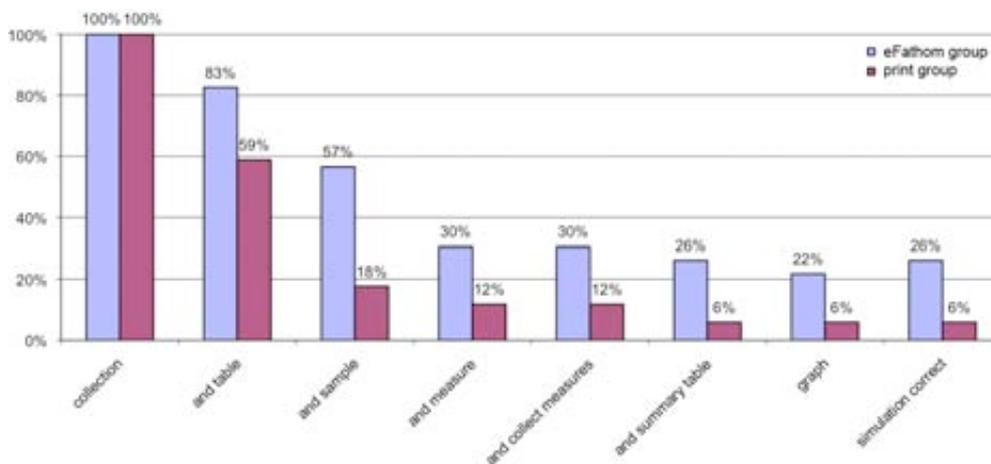
Figure 5 Results from the comparison study of using EFATHOM or printed material

When students returned to the course after having worked on the assignment, they were given a simulation task that they had to solve with FATHOM. They had to document the process by our simulation scheme. Moreover we recorded the process of working with Camtasia. We coded and reduced these data. Fig. 5 shows for each step of the simulation process the percentage of students that successfully finished this step. Due to the complexity of modeling the situation to be simulated the overall success rate was less than in tasks that required less transfer. Anyway, the EFATHOM – group performed better.

Further perspectives

The success of eFathom has lead to a new German edition of Fathom for schools that is provided with eFathom together with activities and instructional material for grade 7 to 12 (Biehler et al., 2011).

REFERENCES

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review, 65*(2), 167-189.

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

Biehler, R., & Prömmel, A. (2010). *Developing Students' Computer-Supported Simulation and Modelling Competencies by Means of Carefully Designed Working Enviroments*. Proceedings of ICOTS 8, Ljubljana, Slovenia [*www.stat.auckland.ac.nz/~iase/publications/.../ICOTS8_8D3_BIEHLER.pdf]*

Biehler, R., Hofmann, T., Maxara, C., & Prömmel, A. (2011). *Daten und Zufall mit Fathom*. Braunschweig: Schroedel.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.

Erickson, T. (2004). *The Model Shop – Using Data to Learn about Elementary Functions*. Oakland, CA: eeps media.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6*, 187 - 205.

Hofmann, T. (2011, in preparation). *eFathom: Entwicklung und Evaluation einer multimedialen Lernumgebung für einen selbstständigen Einstieg in die Werkzeugsoftware Fathom [Ph.D. thesis]*. Kassel: Kassel University.

Maxara, C., & Biehler, R. (2006). Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability. *Proceedings of the 7th International Conference on Teaching Statistics (ICoTS 7)* from http://www.stat.auckland.ac.nz/~iase/publications/17/7C1_MAXA.pdf.

Maxara, C. (2009). Stochastische Simulation von Zufallsexperimenten mit Fathom – Eine theoretische Werkzeuganalyse und explorative Fallstudie. *Kasseler Online-Schriften zur Didaktik der Stochastik (KaDiSto) Vol. 7*. Kassel: Universität Kassel [Online: http://nbn-resolving.org/urn:nbn:de:hebis:34-2006110215452].

Mayer, R. E. (2001). *Multimedia Learning*. Cambrigde: Cambridge University Press.

Mayer, R. E. (2005). *Cognitive Theory of Multimedia learning*. In: R. E. Mayer (ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31-48). Cambridge: Cambridge University Press.

Meyfarth, T. (2008). Die Konzeption, Durchführung und Analyse eines simulationsintensiven Einstiegs in das Kurshalbjahr Stochastik der gymnasialen Oberstufe eine explorative Entwicklungsstudie. *Kasseler Online-Schriften zur Didaktik der Stochastik (KaDiSto) Vol. 6* [https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2006100414792].

Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). *Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach*. Journal of Educational Psychology, 86(1), 122-133.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In: R. E. Mayer (ed.), *The Cambridge Handbook of Multimedia Learning* (pp.. 19-30). Cambridge: Cambridge University Press.

Trouche, L. (2004). Managing the Complexity of Human/Machine Interactions in Computerized Learning Environments: Guiding Students' Command Process through Instrumental Orchestrations. *International Journal of Computers for Mathematical Learning, 9*(3), 281-307.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.