Chapter 7
Learning Mathematics Through Inquiry: A Large-Scale Evaluation

Ton de Jong, Petra Hendrikse, and Hans van der Meij

Introduction

It is currently widely accepted that learning for understanding can only take place when learners adopt an approach in which they process learning material in an active way and engage in "deeper" processes such as asking questions, searching for structures and creating abstractions (e.g., Jonassen, 1991; Mayer, 2002; Novak, 1998; von Glaserfeld, 1987). In their seminal book *How People Learn*, Bransford, Brown, and Cocking (1999) showed that active learning positively affects the construction of understanding and contributes to the development of transferable knowledge. Grabinger (1996) further emphasized that this way of learning and knowledge creation also stimulates learners to connect new information to their existing, personal knowledgebase and that learners need to have new information situated in real or realistic contexts to foster transfer. This view is contrasted with former approaches in which conveying information to learners was seen as the main form of instruction and in which context-free knowledge was seen as the goal to be reached. In an overview of differences between deep and surface approaches to learning in science, Chin and Brown empirically distinguish a number of key learning processes including searching for causally coherent explanations and question asking that characterize good students (Chin & Brown, 2000). The importance of active learning has also been recognized in much earlier work. Dewey (1916), for example, already stressed the importance of "doing" science, mathematics, and history to gain understanding of these domains. "Doing" means that learners abstract, discover, and prove. In Bruner's work as well, learning is seen as an active

T. de Jong (✉) and H. van der Meij
Faculty of Behavioural Sciences, Department of Instructional Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: a.j.m.dejong@utwente.nl

P. Hendrikse
Faculty of Behavioural Science, Teacher Training Department, University of Twente, 7500 AE Enschede, The Netherlands

process in which learners develop new ideas based on prior knowledge (Bruner, 1973; Bruner, Goodnow, & Austin, 1956). Bruner's work partly had its origin in mathematics learning, which is the topic of this chapter

In this chapter, we specifically examine such approaches to learning mathematics. We have seen a shift from a more procedure-oriented view of teaching mathematics to one of helping learners to think mathematically in order to engage in meaningful activities and to understand relationships between mathematical concepts (Bransford et al., 1999; Schoenfeld, 2006). This shift is seen in a change of emphasis from traditional algorithmic problems to insight problems. According to Van Streun (1989), mathematics teachers make a distinction between "routine" problems and "thinking" problems. Routine problems can be solved with the use of algorithms and without much dependency on insight or understanding (see Schoenfeld, 1985; van Streun, 1989). As long as the learner can classify the problem in the correct class, problem solving will take place almost automatically. However, when problems become more complex or when classes of problems are not self-evident, then an algorithmic approach loses its effectiveness. In teaching, a decision should thus be made to either instruct and practice algorithms with satisfactory performance on routine problems, or to adopt a more time consuming, insightful approach to support the construction of more flexible knowledge that can be applied to thinking or transfer problems (Gravemeijer et al., 1993). Cobb and McClain (2006) make a similar distinction by pointing out that statistics education traditionally aims at teaching routines whereas a more conceptual stance that focus on "big ideas" is also needed.

Contemporary approaches in mathematics seek to design conditions that stimulate and support learners to engage in active learning processes that yield conceptual knowledge. One of these is the Realistic Mathematics Education (RME) movement based on the work of Freudenthal (1991). Another approach is inquiry learning in which learners actively investigate mathematical relationships (Pea, 1987). An example of a highly successful implementation of inquiry learning in the field of mathematics is the Jasper series (Cognition and Technology Group at Vanderbilt, 1992, 1997). More recent approaches often use ICT (information, communication, and technology) to facilitate more conceptual learning (Bottino, Artigue, & Noss, 2009; Noss & Hoyles, 2006). These approaches capitalize on the interactive and dynamic capacities of ICT (Atkinson, 2005). Among these, the use of microworlds or simulations (often in the form of applets as in the ESCOT project (Underwood et al., 2005)) has been influential (Kuhn, Hoppe, Lingnau, & Wichmann, 2006). Applets have been developed that support an inquiry approach in science learning (de Jong, 2006a) and there have been some uses of applets in conjunction with computer-supported collaborative learning (Staples, 2007). An example of a computer-based inquiry learning environment for mathematics is SimCalc (Roschelle & Kaput, 1996; Roschelle & Knudsen, this volume). In SimCalc, students can manipulate formulae and observe the consequences of their changes in a number of ways such as animations, tables, and graphs. Another research project that has explored inquiry approaches for using technology to enhance the learning of mathematics is Cabri Géomètre

(Balacheff & Sutherland, 1994; Falcade, Laborde, & Mariotti, 2007; Laborde, 2002), which is a microworld that allows learners to directly manipulate geometrical objects and to observe the effects of their manipulations (Falcade et al., 2007). Yet another example of an inquiry environment in mathematics is PIE (Probability Inquiry Environment) (Vahey, Enyedy, & Gifford, 2000) that focuses on probability theory. Students can manipulate simulations and view their effects in dynamic representations.

We observe, however, that whereas an increasing number of studies have been documenting the effectiveness of technology-enabled inquiry approaches for science learning (Linn, Lee, Tinker, Husic, & Chiu, 2006) research on the effectiveness of (technology-enabled) inquiry for mathematics education are still scarce and often anecdotal. Research on mathematics learning has tended to focus on charting inquiry processes (Linn et al., 2006) although the recent study by Rasmussen and Kwon (2007) found that learners who followed an inquiry approach based on RME scored higher on items assessing mathematical "thinking" or conceptual knowledge than a traditional control group and performed equally on the measures of "routine" mathematical knowledge. A study of learning mathematical ideas with a technology-enabled inquiry approach has been recently completed by Eysink et al. (2009). This research compared the effects of different technological learning environments for learning about probability and found that inquiry learning was the most successful in developing deeper conceptual knowledge.

A general finding in the inquiry literature is that learners need support and that an appropriate balance between guidance and freedom needs to be found (de Jong, 2006b). As Freudenthal has stated: "Guiding means striking a delicate balance between the force of teaching and the freedom of learning" (Freudenthal, 1991, p. 55). Many of the aforementioned learning environments (e.g., many applets, Cabri Géomètre) concentrate on providing students with the opportunity to simulate and manipulate objects or phenomena. Not all of these environments, however, offer the necessary instructional support for these activities. An exception can be found in recent developments in SimCalc Mathworld (Roschelle & Knudsen, this volume). As a response to this issue, in the current study we have developed, over a number of iterations, a set of software-based learning environments to support mathematical inquiry activities. These learning environments give students many opportunities for investigation and exploration, but also provide embedded instructional support for their inquiry. We discuss next a large-scale evaluation of these newly developed learning environments that were compared with a traditional form of teaching.

Basic Setup of the Inquiry Learning Environment

The study focused on students learning about functions in mathematics. More specifically, it treated topics such as linear formulas, parallel lines, domain and range, and solving equations and inequalities. The study used the "Getal and Ruimte"

(Numbers and Space) method, which is widely used in secondary mathematics education in the Netherlands.

We developed a series of four inquiry environments using SimQuest software (van Joolingen & de Jong, 2003), which is an authoring tool for creating simulations with integrated instructional support that may consist of, for example, explanations and assignments. The four learning environments provided the learner with four different concrete contexts for exploring mathematical ideas about functions: Mobile Phones, Windmills, Tsunami,[1] and Benefit Concert. These learning environments were aligned with relevant chapters of the Numbers and Space method.

Figure 7.1 displays a screenshot from Windmills showing the interactive, dynamic, and graphical components of the learning environment (see the left and middle parts of the screen). Students can manipulate values of variables and observe the consequences of these manipulations in a graphical, numerical, and pictorial ways. The right side of the screen displays assignments, such as a task description, and provides students with guidance on how to operate the interactive parts of the environment. After completing an assignment, students receive feedback on their performance.
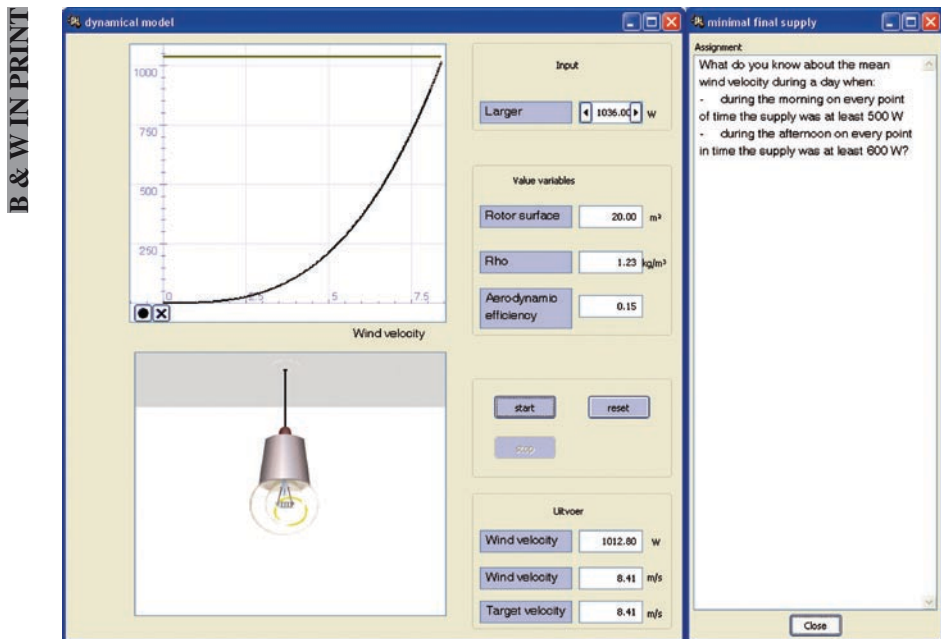

Fig. 7.1 Screenshot of SimQuest Windmill application

---

[1] This application was developed before the tsunami of December 2004 took place.

Development of the Learning Materials

The SimQuest applications serve a central role in the learning materials developed for this study. As mentioned above, each SimQuest application had a specific context to link mathematics to a real world problem (e.g., buying a mobile phone) that may be interactively explored in the simulation of the context. Learners can manipulate variables and then observe the results in various representations, such as graphs, animations, and output fields. The interactive parts are embedded in an instructional environment. SimQuest applications provided learner support through the sequencing of assignments and models as well as explanatory texts. The assignments followed a specific structure and generally started with an introductory text to provide the context and explain the variables, to pose a problem-solving question, and to introduce the interactive part. In addition, the underlying models increased in complexity by adding variables, with each model progression level having its own interface and set of assignments. Topical themes were directly visible for students; however, the degree of complexity was not.

The experimental materials and SimQuest applications were iteratively developed based on the results of a series of preliminary studies that involved 77 students, 41 girls and 36 boys, from secondary education (average age 15–16). Students had different profiles, with 36 students taking predominant courses from the cultural and social sciences (e.g., law, history and geography), and 41 students primarily taking science courses (e.g., mathematics and physics). Participants in the first two preliminary studies had already been taught the subject matter; participants in the third had not. Three mathematics teachers also participated in the third preliminary study. Data in all three preliminary studies was collected through interviews, think-aloud protocols, observations, and log data.

During this development process, we focused on the elicitation and support of learning activities such as abstracting, structuring, evaluating, interpreting, and proving. For example, structuring assignments sometimes invited students to examine differences in results between two different situations and elaborations often promoted reevaluations of how results could have been calculated. In addition, we stimulated and supported students to communicate using the language of mathematics (e.g., in presenting formulas). We found in the third preliminary study that students frequently engaged in these desired learning activities, such as in attempts to formulate a solution process in abstract, general terms or when students began to look at certain situations (structuring) in order to be able to show that a proposition is true (proving).

The basic design decisions for the four SimQuest applications were refined over the course of these studies in three main ways. First, the context of each initial application was used throughout the set of assignments the students completed. In the revised applications, a series of assignment contexts went through transitions in which concrete content gradually became more abstract. That is, students started with a familiar and realistic concrete context that then was translated into a mathematical context. The new mathematical context was further generalized and again

a mathematical view on this general information was taken. Second, the goal of the assignments shifted from constructing equations and concepts to exploring their properties. Third, assignments were extended with subassignments to provide additional support for students who failed on the original assignment. These subassignments were shaped according to the following design.

- Step 1: consider which variable(s) you are going to change and what output you are going to look at.
- Step 2: what are the different possibilities for the values of the variable(s)?
- Step 3: try out the different possibilities.
- Step 4: look back at the process. What can you conclude?

Every subassignment consisted of two components, one that asked the student for a possible approach and one that gave an exemplary elaboration. Fourth, additional support outside of the computer-based SimQuest materials was provided so that the revised materials also included classroom conversations on key topics (e.g., Cobb & McClain, 2006) and requesting students to make subject-matter overviews (e.g., Horton et al., 1993). The classroom conversations allowed special support (such as evaluating, interpreting, and reasoning) that is hard to elicit and support in an online learning environment. In these conversations, students would need to verbalize their ideas and "defend" these against others. As a result, they were encouraged to reflect and think more deeply about what they had done. Students could also be confronted with new and alternative viewpoints, possibilities, and relations to consider. One of the interlocutors in these conversations was the teacher who also can bring in the socio-cultural aspects of the mathematics profession.

The preliminary studies also made it clear that students needed a way to structure the information they received. For this reason we asked the students to make a paper-based subject-matter overview that was *not* to report the outcome of a calculation, but rather to describe what they learned from an assignment and how that knowledge related to the mathematical domain. The overview was intended to stimulate students to draw abstract and general conclusions and to help them to structure different domain elements.

Finally, the third preliminary study indicated that teachers needed support as well. Therefore, for the large-scale study we developed a teacher guide that described a scenario on how to deal with the various information sources (e.g., SimQuest environment, tools, textbook) in all the lessons on functions. The guide described how teachers could alternate between the textbook and SimQuest simulations in such a way that they would be coordinated with each other by roughly dividing each lesson into the four phases: orientation, introduction, processing, and recapitulation. We then indicated which parts of the applications could be used in each phase of each lesson. The final sequence that was developed had these components: (a) introduction by the teacher, (b) students work with the SimQuest materials alone or in groups, (c) whole class conversation (intended to foster processing), and (d) completion of a topic or subtopic, creation of a subject-matter overview consisting of a short description of important findings (intended to foster recapitulation and reflection).

Method

Subjects

In this study, the experimental condition used the inquiry materials (i.e., SimQuest applications, classroom conversations, and subject-matter overviews), whereas for the control condition, "standard" didactic lessons were given (e.g., teacher led questions and answers). Eleven schools from across the Netherlands participated in the study. The experiment started with 470 students in 20 classes. However, due to illness and absentees, the final dataset consisted of 418 students (206 male and 212 female). The students came from secondary education classes and ranged in age from 15 to 16. Of these students, 155 had an "M-profile" (cultural and social science) and 263 had an "N-profile" (science). The N-profile attracts students who are primarily interested in science and technology topics, whereas the M-profile students tend to focus on courses in culture, economics, and society. Students in the N-profile on average have stronger background knowledge in science than students in the M-profile, which is not surprising as the N-profile curriculum contains more science elements. The division of classes over conditions was not arbitrary; schools had chosen to place classes in the experimental or in the control condition, often based on practical reasons. Seven classes (140 students) were in the control condition and 13 classes (278 students) were in the experimental condition. The division of gender and of M- and N-profiles over conditions is shown in Table 7.1. Chi-square analyses showed that gender and profiles were not divided evenly across the two conditions. The control condition contained more students from the N-profile and more male student participated in the control condition, which was probably due to the historical trend that the M-profile attracts more female than male students.

Table 7.1 Posttest scores (adjusted)

Condition	Profile	Gender	Number	Mean	SE
Control	M	Male	21	23.51	1.95
		Female	19	22.35	2.06
	N	Male	55	30.30	1.24
		Female	45	32.58	1.36
	Total		140	27.19	0.86
Experimental	M	Male	42	25.19	1.40
		Female	73	23.72	1.10
	N	Male	88	32.15	0.95
		Female	75	28.78	1.03
	Total		278	27.46	0.57
Total			418	27.32	0.50

Test

The pretest used in this study assessed relevant prior mathematical knowledge. The pretest, which had a maximum score of 40, consisted of four main questions and 14 subquestions on first-degree and second-degree equations and geometry. Students were allowed a maximum of 20 min to complete the pretest, which was found to have a Cronbach's α reliability of 0.74.

The posttest consisted of six main questions that split into 15 items that covered the topics of linear functions, investigating functions, equations and inequalities, and applications (e.g., optimizing a surface). Six items from the posttest measured procedural knowledge, six other items measured conceptual knowledge. It was hypothesized that the control group would perform better on the procedural items whereas the experimental group was expected to score higher on the conceptual items. The remaining three items measured a combination of conceptual and more technical – procedural – knowledge. No predictions were given for those three items.

The maximum score on the posttest was 63. Because in some schools different classes participated in the experiment and not all classes took the test at the same time, four different versions of the posttest were developed. Items on the different versions of the posttest differed in appearance but not in content. The students' score on the posttest was counted as an actual mark for their school examination. The Cronbach's α reliability of the posttest was 0.68.

Procedure

Teachers in the experimental condition attended an introductory meeting a few months before the start of the series of lessons where they worked with the four SimQuest simulations. They also received a schema for 12 lessons that described the materials from the textbook and the simulations that should be covered in each lesson. Prior to the start of the lesson series, the teachers received the teacher guide and a software manual plus CD with the instructions and software for installing the simulations. In the first lesson, the students were introduced to the software, and a member of the research team was present in the lesson to assist with software installation and to answer students' questions. The pretest was administered in the second lesson, and the actual activities for the experimental and control conditions began in the third lesson, continuing to the end of the implementation. The lessons for the experimental condition had a general format of introductions in which the software was used for demonstration purposes, after which students worked with the software and on exercises from the textbook. There were also classroom conversations that were sometimes held after students had individually worked with the software in order to discuss their findings or to discuss the main issues of a series of problems. Each topic ended with a summary and the students completed a subject-matter overview. The posttest was taken around a week to 10 days after the last

lesson. Teachers in the participating schools were free to follow their own insights and organization in implementing the program that led to considerable differences between schools. For example, the number of lessons on subtopics could differ between two and four, and lesson length varied between 45 and 60 min. In addition, the availability of computer rooms, data projectors, and other technical facilities varied considerably between schools that influenced the students' time on task. Because the use of a computer program in mathematics lessons was new to the teachers, the lessons often took more time than anticipated, which led some teachers to cover less of the domain than originally intended.

Results

The control group scored significantly higher on the pretest ($M = 19.98$, SD $= 6.58$) than the experimental group ($M = 14.56$, SD $= 7.37$). A regression analysis, using the enter model,[2] gave a significant model for *pretest* scores with the factors of Condition, Gender, Profile, and their interactions ($F_{7, 410} = 14.14$, $p < 0.001$). The factor Condition was significant (df $= 410$, $t = -6, 57$, $p < 0.001$), as was the factor Profile (df $= 410$, $t = 4.92$, $p < 0.001$) in favor of the N-profile. Gender was not significant and there were no interactions.

On the overall (uncorrected) *posttest* score, the control group again outperformed the experimental group. However, the average relative difference (absolute difference/total number of points) between both groups decreased from 13.6% to 6.3%, which indicates that the two groups have come closer together. Given the significant difference between the two conditions on the pretest, these scores were used as a covariate for additional analyses. An enter model regression gave a significant model ($F_{8,409} = 25.361$, $p < 0.001$) for the posttest scores with the factors Condition, Profile, Gender, and the interactions between these three factors using the pretest as a covariate. The factor Condition was not significant ($n = 418$, df $= 409$, $t = 0.260$, $p > 0.05$), whereas the factor Profile was significant ($n = 418$, df $= 409$, $t = 7.01$, $p < 0.001$, Cohen's $d = 0.73$, one-sided test).[3] Gender was not found to be a significant factor ($n = 418$, df $= 409$, $t = -0.93$, $p > 0.05$). There were no significant interactions. Table 7.1 shows the posttest scores (adjusted with the pretest scores) for the control and the experimental group with a further division into gender and profile.

---

[2] In order to run a regression analysis, one of the optional models has to be chosen. The different options are: enter, stepwise, remove, backward, and forward. The enter model is also called forced entry model. All variables specified are entered into the model in a single step. This model is generally used (when there are no specific expectations).

[3] One-sided tests are performed for specified predictions (e.g., the control condition performs better on procedural items). All other tests are two-sided.

We also compared the performance of the conditions on conceptual and procedural items for which the data are displayed in Table 7.2. A regression analysis (enter model) gave a significant model for *procedural* items ($F_{8,409}=17.380$, $p<0.001$) with the factors Condition, Profile, Gender, and their interactions with pretest scores as covariate. Students in the control condition outperformed the experimental condition students ($n=418$, df$=409$, $t=-1.687$, $p=0.046$, Cohen's $d=-0.18$, (one-sided test). There was a trend for girls to score higher on the procedural items in the control condition ($n=418$, df$=409$, $t=-1.777$, $p=0.076$, Cohen's $d=-0.19$, two-sided test). For *conceptual* items, a regression analysis (enter model) also resulted in a significant model ($F_{8,409}=10.858$, $p<0.001$) with the factors Condition, Profile, Gender, and the four interactions over these factors with pretest score as covariate. There was no significant difference between conditions ($n=418$, df$=409$, $t=0.466$, $p=0.321$, one-sided test), but there was a trend for boys to outperform girls ($n=418$, df$=409$, $t=-1.835$, $p=0.067$, Cohen's $d=-0.19$). There was also a trend for an interaction between condition and profile ($n=418$, df$=409$, $t=-1.751$, $p=0.081$, Cohen's $d=-0.18$, two-sided test): M-profile students performed better in the experimental condition while the N-profile students scored higher in the control condition.

As stated earlier there were considerable differences between schools. Therefore, we also explored the data of two more or less comparable classes (both classes come from the same school, have an N-profile, one is in the control condition, the other in the experimental condition). A regression analysis (enter model) yielded a significant model for *posttest* scores with the factors Condition, Gender, and their interaction with pretest scores as covariate ($F_{4, 39}=7.11$, $p<0.001$). There was a trend for students in the experimental Condition to score higher ($n=44$, df$=39$, $t=1.90$, $p=0.065$, Cohen's $d=0.61$). Gender was not significant, nor was the interaction between gender and condition. The pretest significantly influenced the results on the posttest ($n=44$, df$=39$, $t=4.99$, $p<0.001$, Cohen's $d=1.60$, one-sided test) and there were no significant interactions.

Table 7.2 Posttest scores (adjusted percentages) and SE (of the adjusted scores) on conceptual and procedural items

Condition	Profile	Gender	Number	Procedural		Conceptual	
				Mean	SE	Mean	SE
Control	M	Male	21	0.55	0.04	0.55	0.07
		Female	19	0.59	0.04	0.41	0.07
	N	Male	55	0.63	0.03	0.73	0.04
		Female	45	0.71	0.03	0.75	0.05
	Total		140	0.62	0.02	0.61	0.03
Experimental	M	Male	42	0.54	0.03	0.59	0.05
		Female	73	0.53	0.02	0.53	0.04
	N	Male	88	0.64	0.02	0.73	0.03
		Female	75	0.62	0.02	0.66	0.04
	Total		278	0.58	0.01	0.63	0.02
Total			418	0.60	0.01	0.62	0.02

A regression analysis (enter model) didn't give a significant model for *procedural* items ($F_{4, 39} = 1.99$, $p = 0.114$) with the factors Condition, Gender, and their interaction with pretest scores as covariate. A regression analysis (enter model), gave a significant model for *conceptual* items with the factors of Condition, Gender, and their interaction with pretest as covariate ($F_{4, 39} = 3.48$, $p = 0.016$). There was a trend for students in the experimental Condition to score higher ($n = 44$, df $= 39$, $t = 1.59$, $p = 0.060$, Cohen's $d = 0.51$, one-sided test). Gender was not significant, nor was the interaction between gender and condition.

Discussion and Conclusion

In this work, we have developed a set of computer-based inquiry learning environments for mathematics. The materials were developed iteratively over a range of preliminary studies. Besides the computer materials, there was an instruction booklet and a teacher manual with guidelines for the setup of a series of lessons, for classroom discussions, and for creating subject-matter overviews. In addition, the teacher manual explicitly linked the new material to the existing textbook. The material was not confined to a single lesson or a limited part of a topical domain but covered a series of 12 weeks of lessons on all the topics of functions treated in the textbook. Once developed, we tested the materials by "letting them loose" in a larger set of schools of a divers nature and compared the results with achievements in traditional classrooms that just followed the textbook.

The large-scale study that we conducted showed that implementing the program in schools led to a wide diversity of usages, dependent on local organization, structures and facilities. Schools differed considerably in their implementation efforts, sometimes shortening the program due to time constraints (many teachers in the experimental condition dropped the creation of subject-matter overviews), sometimes skipping computer exercises due to problems with the facilities (e.g., computers that were out of order, projectors that did not function or were not available). Of course, this threatens the experimental rigor, but it should also be recognized that the materials are likely to be used in these ways in everyday practices. In any case, we can safely conclude that the implementation of the experimental condition was not optimal in many schools.

Even under these challenging conditions the learning results of the experimental group on the posttest were encouraging. After correcting for pretest scores, the outcomes in the experimental condition equal the outcomes in the control condition in which students received the type of instruction they were used to and in which no major practical problems occurred. Exploratory analyses of two classes, one control and one experimental, that were more or less comparable for these external conditions, points even more strongly in this direction as posttest scores turned out higher for the experimental group.

Students from the control group turned out to score significantly better on procedural items for which, primarily, mastery of techniques is important. These students also

executed the test at a higher pace and succeeded to reach the end of the test more frequently (in the control group 82.9% of the students made one of the last questions, in the experimental group 67.3% did). This suggests that these students had automated their knowledge more, which was what we expected. In contrast, students in the experimental groups had higher scores (corrected for pretest scores) on conceptual (insight) items. This is in line with the idea that the experimental material focused more on building insight than the traditional material. Although the latter difference between conditions did not reach significance it was in the predicted direction. These results fit into a more general trend that is emerging from research (see, e.g., Rasmussen & Kwon, 2007).

An interesting finding in this study concerns the gender differences. Overall, and regardless of prior knowledge, girls performed better in the traditional classroom setting whereas boys profited more from the inquiry setting. One possible explanation for this effect is self-efficacy, that is, the students' handling of the inquiry environment may have been affected by their competency beliefs about learning mathematics in a more open-ended, guided inquiry learning manner. On the influence of gender on mathematics, the literature is equivocal. Some studies found no differences between girls and boys in self-efficacy beliefs toward mathematics (e.g., Chen & Zimmerman, 2007) whereas others report relevant gender differences. For example, Meece, Glienke, and Burg (2006) found that boys report higher interest and competency beliefs in mathematics than girls and Frenzel, Pekrun, and Goetz (2007) found that girls feel more insecure in mathematics than boys even when their knowledge is on the same level. To our knowledge, only few studies on the impact of gender on (guided) inquiry learning have been conducted. In an older study, Gennaro and Lawrenz (1992) found that girls performed better on inquiry tasks than boys. A similar finding is reported by Timmermans, Van Lieshout, and Verhoeven (2007) who compared guided instructions with prescribed, direct instruction. On the former, girls performed better and felt more at ease than boys. It is clear that more work needs to be done to unravel the relation between gender and inquiry learning in general and mathematics inquiry learning in particular.

An obvious question is how the implementation can be improved. One important constraint was the structure and quality of the textbook. We had to work in a set curriculum and therefore took the textbook of the schools as our starting point for developing the SimQuest applications. This turned out not to be optimal; among others because equations seemed to come "out of the blue." Whenever possible, learning materials should be simultaneously developed to realize a better integration of textbooks and the interactive materials (see chapter by Roschelle & Knudsen, this volume). Another important factor that we could not alter in the present study was the time available for this series of lessons. The realistic class situation required accommodating the existing time schedule for learning the topic of functions. However, having learners investigate mathematics themselves invariably costs more time. In the current situation, it may have demanded too much time. A different time schedule maybe necessary when students engage in inquiry learning, certainly when they do this for the first time. The third factor that clearly limited the implementation concerns the access to computer facilities. In many schools, it

was difficult to use computer applications comfortably in the lessons. Even in schools with a good computer infrastructure, problems repeatedly emerged due to organizational obstacles. In other words, a hefty check on computer facilities and organizational embedding is needed to ensure that these conditions do not form an obstacle.

Overall, the results of this study confirm a set of recent studies that indicate that traditional didactic teaching approaches achieve lower-order learning outcomes, whereas learner-centered and inquiry approaches, often enabled by technology, allow students to construct deeper and more conceptual understandings and enhanced problem-solving abilities. This justifies efforts to further investigate the conditions under which these types of learning experiences can be optimized.

References

Atkinson, R. K. (2005). Multimedia learning of mathematics. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 393–408). Cambridge, UK: Cambridge University Press.

Balacheff, N. & Sutherland, R. (1994). Epistemological domain of validity of microworlds - The case of Logo and Cabri-Geometre. *Lessons from Learning, 46*, 137–150.

Bottino, R. M., Artigue, M., & Noss, R. (2009). Building European collaboration in technology enhanced learning in mathematics. In N. Balacheff, S. Ludvigsen, T. de Jong, S. Barnes, & A. Lazonder (Eds.), *Technology enhanced learning - Principles and products* (pp. 73–89). Berlin: Springer Verlag.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (eds). (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C: National Academy Press.

Bruner, J. (1973). *Going beyond the information given*. New York: Norton.

Bruner, J., Goodnow, J., & Austin, A. (1956). *A study of thinking*. New York: Wiley.

Chen, P. & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *Journal of Experimental Education, 75*, 221–244.

Chin, C. & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*, 109–138.

Cobb, P. & McClain, K. (2006). Guiding inquiry-based math learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 171–186). Cambridge: Cambridge University Press.

Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program, description, and assessment data. *Educational Psychologist, 27*, 291–315.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project; Lessons in curriculum, instruction, assessment, and professional development*. Hillsdale, NJ: Lawrence Erlbaum Associates.

de Jong, T. (2006a). Computer simulations - Technological advances in inquiry learning. *Science, 312*, 532–533.

de Jong, T. (2006b). Scaffolds for computer simulation based scientific discovery learning. In J. Elen & R. E. Clark (Eds.), *Dealing with complexity in learning environments* (pp. 107–128). London: Elsevier Science Publishers.

Dewey, J. (1916). *Democracy and education. An introduction to the philosophy of education*. New York: MacMillan.

Eysink, T. H. S., de Jong, T., Berthold, K., Kollöffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: an analysis across instructional approaches, *46*, 1107–1149.

Falcade, R., Laborde, C., & Mariotti, M. (2007). Approaching functions: Cabri tools as instruments of semiotic mediation. *Educational Studies in Mathematics, 66*, 317–333.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics - A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*, 497–514.

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Gennaro, E. & Lawrenz, F. (1992). The effectiveness of take-home science kits at the elementary level. *Journal of Research in Science Teaching, 29*, 985–994.

Grabinger, R. S. (1996). Rich environments for active learning. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 665–692). New York: Simon & Schuster Macmillan.

Gravemeijer, K., van den Heuvel-Panhuizen, M., van Donselaar, G., Ruesink, N., Streefland, L., Vermeulen, W., et al. (1993). *Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek (No. SVO-6010)*. Utrecht: Freudenthal-instituut.

Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77*, 95–111.

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology: Research & Development, 39*, 5–14.

Kuhn, M., Hoppe, U., Lingnau, A., & Wichmann, A. (2006). Computational modelling and simulation fostering new approaches in learning probability. *Innovations in Education and Teaching International, 43*, 183–194.

Laborde, C. (2002). Integration of technology in the design of geometry tasks with Cabri-Geometry. *International Journal of Computers for Mathematical Learning, 6*, 283–317.

Linn, M. C., Lee, H.-S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration in science. *Science, 313*, 1049–1050.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice, 41*, 226–232.

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology, 44*, 351–373.

Noss, R. & Hoyles, C. (2006). Exploring mathematics through construction and collaboration. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 389–409). Cambridge: Cambridge University Press.

Novak, J. D. (1998). *Learning, creating and using knowledge: Concept map^{TM} as facilitative tools in schools and corporations*. Mahwah (NJ): Lawrence Erlbaum Associates, Inc.

Pea, R. D. (1987). Cognitive technologies for mathematics education. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 89–122). Hillsdale, NJ: Erlbaum.

Rasmussen, C. & Kwon, O. N. (2007). An inquiry-oriented approach to undergraduate mathematics. *The Journal of Mathematical Behavior, 26*, 189–194.

Roschelle, J. & Kaput, J. (1996). SimCalc MathWorlds for the mathematics of change: Composable components for calculus learning. *Communications of the ACM, 39*, 97–99.

Roschelle, J., & Knudsen, J. (this volume). From new technological infrastructures to curricular activity systems: advanced designs for teaching and learning. New York: Springer

Schoenfeld, A. H. (2006). Mathematics teaching and learning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 479–510). Mahwah, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic Press.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction, 25*, 161–217.

Timmermans, R. E., Van Lieshout, E., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction, 17*, 42–54.

Underwood, J. S., Hoadley, C., Lee, H. S., Hollebrands, K., DiGano, C., & Renninger, K. A. (2005). IDEA: identifying design principles in educational applets. *Etr&D-Educational Technology Research and Development, 53*, 99–112.

Vahey, P., Enyedy, N., & Gifford, B. (2000). Learning probability through the use of a collaborative, inquiry-based simulation environment. *Journal of Interactive Learning Research, 11*, 51–84.

van Joolingen, W. R. & de Jong, T. (2003). SimQuest: Authoring educational simulations. In T. Murray, S. Blessing & S. Ainsworth (Eds.), *Authoring tools for advanced technology educational software: Toward cost-effective production of adaptive, interactive, and intelligent educational software* (pp. 1–31). Dordrecht: Kluwer Academic Publishers.

van Streun, A. (1989). *Heuristisch wiskunde-onderwijs: Verslag van een onderwijsexperiment.* Groningen.

von Glaserfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems in the representation in the teaching and learning of mathematics* (pp. 3–17). Hillsdale, NJ: Lawrence Erlbaum.