Scandinavian Journal of Educational Research
Vol. 49, No. 3, July 2005, pp. 297–323

Computer-supported Problem-based Learning in the Research Methodology Domain

Sirpa Lehti* and Erno Lehtinen
University of Turku, Finland

The purpose of this research was to look for possible benefits and drawbacks of the use of computer-supported simulation in the teaching and learning of experimental research methodology and statistics. In the study three research methodology groups were compared. The results show that there were significant differences in favour of the computer-supported simulation group, called the ALEL group. During the course the conversations of two students' in the ALEL group were audiotaped. Although the ALEL students performed better than the other students and showed an improvement during the course, the conversations analysed showed that their learning outcomes should be still better in order to meet the learning goals of the methodology and statistics curriculum. An inadequate knowledge base was shown in the post-test and also in the discussions of the pair during the course.

Keywords: *Research methodology; Statistics; Problem-based learning; Computer-supported learning*

Introduction

Research methodology and statistics are important subjects in university studies, but motivational problems often interfere with learning. A long-standing aversion to quantitative and technical aspects of research methodology has also been detected (Cutler, 1987; cited in Winn, 1995). Even those graduates who have taken a few methodology courses and have completed pieces of research may still have a poor methodological understanding (Rui, Suntio, & Lehtinen, 1995). One reason why methodological learning is so difficult is that areas of knowledge remain a series of unconnected pieces (Lehtinen & Rui, 1995; Winn, 1995). This may be partly due to the complexity of the methodology domain (Lehtinen & Rui, 1995). Complex content areas have typically been divided into small content units that have been subsequently ordered in list form. These lists of goals and content units have proved

*Corresponding author. Educational Technology Unit, University of Turku, 20014 Turku, Finland. Email: sirpa.lehti@utu.fi

ISSN 0031-3831 (print)/ISSN 1430-1170 (online)/05/030297-27

DOI: 10.1080/00313830500109618

to be useful in teaching well-defined simple knowledge structures and fact lists, but inconvenient for learning complex knowledge structures and the skills needed in solving problems typical of professional practice (Lehtinen & Rui, 1996). Even good mastery of the content units typically does not lead to comprehension of the complex situation and the learning outcome remains on the level of memorized lists of isolated units. In some recently developed computer-based learning environments the complexity of the content area has consciously been considered.

Students consider statistics and quantitative research methods more difficult than their major subject studies (Murtonen & Lehtinen, 2003) and anxiety about statistics (Bell, 2001; Forte, 1995; Onwuegbuzie & Seaman, 1995) has been reported in many studies. Statistics is seldom taught using real data (Singer & Willett, 1990; Thompson, 1994) and this may be one reason for the difficulties experienced in applying the knowledge learned. Thus, many graduate students fear applying research knowledge and using research language (Onwuegbuzie, 1997).

To enhance students' interest in research methodology and statistics it appears to be important that the subject is taught in a meaningful context and that the data are real (see, for example, Magel, 1996; Singer & Willett, 1990; Thompson, 1994; Winn, 1995). If students are involved in a project where they cannot participate in all stages of the research process, substructures of the research domain may remain insular (Winn, 1995). On the other hand, if the whole research process is carried out by all students there is a massive amount of work involved and the data may still be inadequate for some statistical analyses.

The internal theory structures of a discipline have traditionally been set as the basis of the curriculum in higher education. Here it has relied on well-defined and segmented study exercises but has not managed to develop students' abilities to apply their knowledge in complex, ill-defined practical situations (Actenhagen, 1994). This type of higher education often fails to provide students with the knowledge and skills which would be applicable to the different problem-solving situations and activities of working life (Mandl, Gruber, & Renkl, 1995). This has led to a strong emphasis on learning experiences in practical situations that are embedded in cultural contexts or communities of practise. Although this approach has led to important progress in the planning of learning environments (Bransford, Brown, & Cocking, 1999), it has also often neglected the importance of learning formal knowledge (see, for example Boshuizen, Smidt, Custers, & van de Wiel, 1995) and the construction of abstract ideas (Ohlsson & Lehtinen, 1997). The problems of higher education cannot simply be solved by cutting studies of formal, theoretical knowledge out of the curriculum and replacing them with direct studies of informal knowledge related to the domain in question.

Problem- and Case-based Approaches

A variety of new approaches to teach students to learn more applicable knowledge structures and manage complex and ill-defined tasks has been developed. Different case- and problem-based methods have turned out to be effective approaches in many learning contexts. Most of these models stress authentic cases or problems, students' self-directed learning, and collaborative processes (Schmidt & Moust, 2000).

In its established form problem-based learning is typically organized according to a sequence of distinguishable steps (Barrows & Tamblyn, 1980). For example, in a typical tutorial group at Maastricht University problem based learning follows the so-called "seven jump" procedure: (1) clarify unknown terms and concepts in the problem description; (2) define the problem, that is, list the phenomena to be explained; (3) analyse the problem, brainstorm, and try to produce as many different explanations for the phenomena as you can using prior knowledge and common-sense; (4) criticize the explanations proposed and try to produce a coherent description of the processes that, in your opinion, underlie the phenomena; (5) formulate learning issue for self-directed learning; (6) fill the gaps in your knowledge through self-study; (7) share your findings with your group and try to integrate the knowledge acquired into a comprehensive explanation for the phenomena, checking whether you now know enough (Schmidt & Moust, 2000, p. 23). One principle which seems to be common to different approaches is to acquaint students with the structural complexity of the tasks from the very beginning of their study career. Instead of teaching sequences of isolated content units these environments present the students with complex problems while they are studying the sub-elements of problems.

Computer Environments for Methodology Studies

Several computer simulations have been created to help with the teaching and learning of experimental research. Most simulations were designed in the 1970s and early 1980s, for example EXPER SIM (Forbach, 1979), LABSIM (Eamon, 1980; Edwards, 1996; Kissler, 1974), LESS (Thurmord & Cromer, 1975), CLASCONSIM (Benedict, 1979) and Project Simulation (King & King, 1988; King, King, & Williamson, 1984). Most of them were developed to help with pragmatic problems in teaching psychological research: laboratory experiments are time consuming and expensive, they need space and equipment, and many laboratory experiments are ethically questionable. Most of the programs have a similar structure. Students get a list of several independent and dependent variables from which they choose a few. The purpose is to obtain results in connection with a research problem. Students choose the statistical analysis and the results are presented on a computer screen. The purpose is that designs are constructed quickly, so that there is time to construct many designs during a short period of time. It is difficult to profoundly analyse these programs, since articles provide only short descriptions and there is a lack of research on how students have learned research methodology while using these programs. For example, Edwards (1996) has already published version 9 of LABSIM, but little research has been done on its use and impact. Lately, net-based experiment generators have been constructed, mostly in the domain of psychology (Ransdell, 2002). They provide opportunities for designing experiments and reporting data, but are mostly designed to teach a variety of phenomena in psychology (Levy & Ransdell, 1999).

Steinkuehler, Derry, Hmelo-Silver, & Delmarcelle (2002) have built a fairly strict sequence of different individual and collaborative activities forming a problem-based learning procedure in a computer environment. In this approach the idea is that a strongly structured network environment guides students into a meaningful problem-based learning process even in situations where the tutor is inexperienced in guiding tutorials or has too little time for all the small groups.

In a computer-supported problem-based learning environment there are many opportunities for pre-structured activity and interaction sequences. It is challenging to develop computer-supported environments in which the pedagogical ideas and procedures of problem-based learning are implemented in the structure of the software. The STEP program, developed by Steinkuehler et al., is a good example of an attempt to facilitate problem-based learning by highly specialized tools. They have managed to create a network environment in which students are guided through a series of collaborative and individual activities forming a typical problem-based learning process. The problem, however, is that students do not always interpret and use the computer environments in the expected ways (see Järvelä, Lehtinen, & Salonen, 2000). This problem is likely to occur if we try to implement too complex pedagogical procedures in learning environments.

In conventional problem-based learning models problems are typically presented in the form of written case descriptions. These case presentations give a limited amount of information and are static in nature. This means that reading the descriptions, which are only a few sentences long, is the students' only direct contact with the cases underlying the problems to be dealt with in the learning process.

Although many studies of problem-based learning have shown that it is possible to formulate written case presentations that serve as inspiring starting points for individual and collaborative problem solving, we should consider the opportunities to create new forms for problem presentations using technological tools (Lajoie, Lavigne, Guerrera, & Munsie, 2001).

Computer simulations provide very rich opportunities to present authentic problems for learning. Using a well-designed, simulated, computer-generated environment a student can focus her/his attention on those questions that are necessary for the theoretical and practical management of the task at hand. Covering complex problems typical of real life work environments is very difficult for many fields of higher education, because dealing with real complex problems demands a substantial investment of resources, takes a considerable amount of time, and may be problematical due to ethical and safety issues involved. Computer-based applications simulating these complex and practical problems do, however, provide us with promising opportunities for developing higher education to meet the challenges that a developing society places on future academic experts (Lehtinen & Rui, 1996). In a simulated environment it is also possible to practice complex problem situations which are very rare in practical work but which experts should be able to cope with immediately they are faced with them (Lesgold, Lajoie, Bunzo, & Eggan, 1992). In many recent studies simulated environments have been used as a tool for problem presentation in problem-based learning. For example, Bergland, Klyczek, Lundeberg, Mogen, Nelson, & Johnson (2001) simulated DNA electrophoresis to create a starting point for student discussions as part of a course on genetics, while Lajoie et al. (2001) developed a system called BioWorld to facilitate problem-based learning among high school biology students. One limitation of typical simulation environments is that students tend to view the simulations as artificial games.

In our own studies we have combined computer simulations and collaborative network environments to enrich (distributed) problem-based learning. In this approach information on real cases is presented with the help of a simulation environment that makes it possible for a collaborative problem-based learning group to obtain further information about the particular case through multiphase interaction with the simulation environment. The ALEL (artificial laboratory for exploratory learning) program is designed to facilitate learning about scientific experiments and analysis of the results by statistical methods (Lehti & Lehtinen, 1999). NerveGame is for first year anatomy students learning the structure and function of muscles and nerves (Salmi, Lehti, & Lehtinen, 1999). In addition, we have developed a simulation for presenting cases of children with infectious diseases (Lehtinen, Nurmela, & Salo, 2001). All the environments share three common features: (a) they present information on real cases; (b) students have to carry out multiphase activities typical of professional practice to obtain information; (c) the simulations provide students with different representations of the information and the problem-solving path they have carried out. Preliminary results show that these dynamic problem presentations facilitate collaborative problem-solving processes and more precise definition of the learning issues by providing the student group with a joint point of reference that develops according to their progress in the problem-solving process.

Computer-assisted Environment for Problem-based Learning Studies in Research Methodology

We started to develop a computer application for learning research methodology at the beginning of 1990. Early studies on the use of ALEL have shown that students' learning outcomes improved significantly compared with a control group in which research methodology and statistical inference were taught more traditionally (Lehti & Lehtinen, 2000; Lehtinen & Rui, 1995). The researchers concluded that an improvement in higher order learning in particular was observed in the ALEL group. Students in the ALEL group were better able to apply methodological knowledge to complex and ill-defined situations. Differences were found in total achievement, procedural knowledge, epistemologically complex knowledge, and practical design skills, but not in factual knowledge.

The Research Problems

Based on the earlier good results at the group level (Lehti & Lehtinen, 1999; Lehtinen & Rui, 1996) two problems were addressed. The first was to determine the efficiency of the ALEL method compared with two other methods of teaching experimental research methodology, the second to determine in detail the procedural features of ALEL and the reasoning and problem solving used by students following this method.

The first problem was attacked by comparing the results for three experimental research methodology courses, each with an expert and experienced teacher of the particular course method. The second problem was explored by following a pair of students using ALEL, who describe their strategies and behaviour, thus detecting possible benefits and drawbacks of this kind of simulation to learning experimental research methodology and statistics.

Methods

Subjects

The participants were 32 university students of educational science. The students were assigned to three matched groups based on scores in a pre-test and prior credits in educational science. There were no significant differences between the groups in the pre-test. In all groups the students worked in pairs or as a threesome. In the ALEL group one pair was chosen for closer analysis. The students' work as a pair was audiotaped.

Materials

Pre- and post-tests. The questionnaires used in pre- and post-tests consisted of 15 open-ended questions dealing with different aspects of experimental research methodology. Nine items asked the students to give a short definition of basic concepts within empirical research methodology (these variables were summed and termed factual knowledge) and statistics. Three items asked them to present a general description of different procedures related to the planning and carrying out of an experimental design (the sum score of these variables was termed procedural knowledge). One item gave the students a half-page description of a complex practical problem and asked them to design an experimental study to improve understanding of this problem (practical design). One item showed the subjects an experimental design and asked them to decide which statistical analysis would be most appropriate to analyse the data. The last item dealt with making statistical inferences. Subjects were shown an experimental research design and an example the output of a statistical analysis. Subjects were then asked to make statistical inferences on the basis of the analyses. Questions that concerned statistics were all summed and termed statistical knowledge.

The ALEL learning environment. The ALEL learning environment has been described in detail by Lehtinen and Rui (1995). Only a brief description of the basic elements of ALEL is presented here.

The domain of the ALEL learning environment is empirical research methodology, especially experimental design and statistical inference. ALEL is meant to be used on intermediate and advanced university courses on research methodology. It covers the content dealt with in the experimental methodology and statistical inference chapters of widely used methodology books for undergraduate and graduate programmes of the educational and social sciences.

ALEL consists of a planning and problem-solving environment integrated with a methodological hypertext reader. Students start to work with the ALEL program by selecting one of its research topics. Then they are provided with a theoretical introduction to the selected research topic. The description of the topic is based on real experiments that have been reported in international journals. In the beginning students are only given general information about the theoretical ideas and aims of the model experiment, with no description of the experimental design used by the original researchers. Guided by this introduction, the student starts planning and conducting his or her own experiments in the simulated environment. After formulating empirical hypotheses for the experiment, the student selects an appropriate population to be used in testing the hypotheses. The system contains simulated populations, treatments, and measurement instruments for the selected research topic. The simulated populations are generated on the basis of the statistical data for the original model experiment.

The experiment-specific tools (treatments and measurement instruments) are constructed to be similar to the instruments used in the model experiment. The general tools available in the program consist of sampling methods, grouping procedures, and statistical techniques that can be used by the students.

As the students plan and implement an experiment the system generates, step by step, an external representation of the activity structure. This representation is displayed on the computer screen as a hierarchical tree diagram. Students create experimental designs by defining sequences of actions. Every action forms a node in the tree diagram, which describes the activity structure of the student while planning and carrying out an experimental design (see Figure 1).

ALEL is integrated with the SPSS™ statistics package and the student can use the statistical operations of this package through the ALEL interface. The statistical tools of ALEL help the students to plan and conduct statistical analyses that are closely connected to the experimental design (Figure 2). This is done by moving (via a drag-and-drop function) groups from the tree diagram to the statistical design

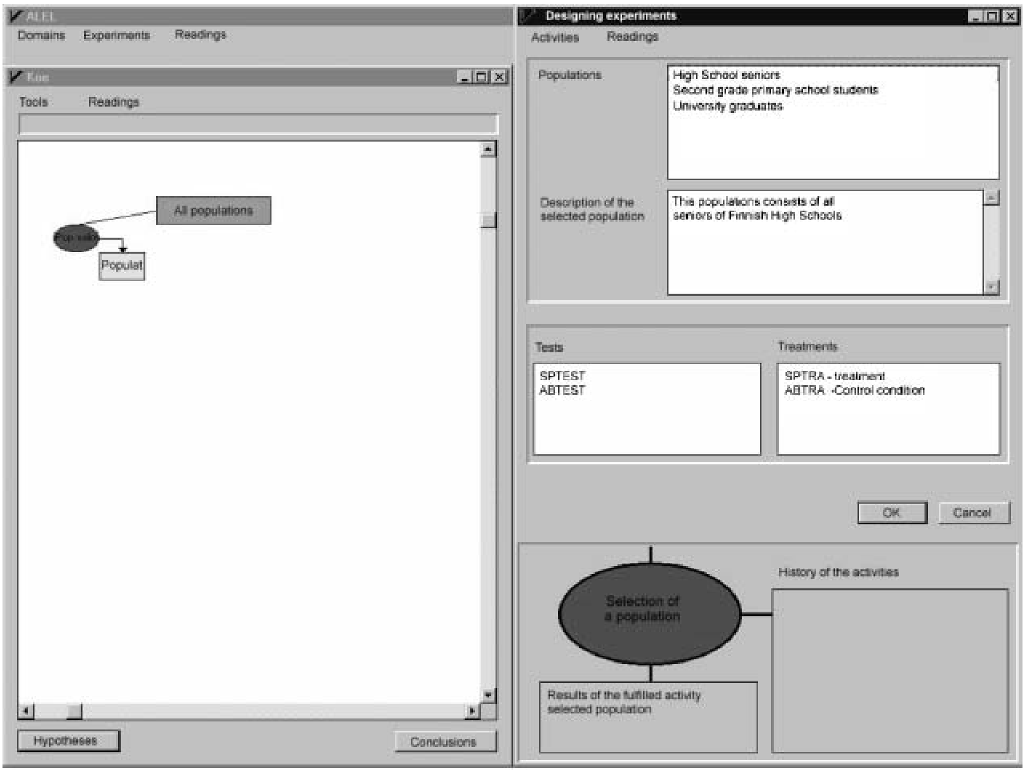

Figure 1. Tools for planning experimental designs in ALEL

form. The results of statistical analyses are presented in a special window of ALEL that provides the students with online help in interpreting the meaning of the different elements presented in the tables.

During planning and problem solving the student has access to a hypertext document which gives information about and online assistance with the different experimental designs and statistical methods. Access to the hypertext is related to the student's current activity. If he or she is, for example, planning the sample, the system provides a path to the hypertext with basic information about the sampling method and links to related topics.

Experimental research implemented in ALEL. Students start working with ALEL by selecting one of its research topics. In this study students used only one research topic throughout the course. The research problems that students had to deal with were modified from a study by Fong and Nisbett (1991). The research problems dealt with in the Fong and Nisbett study were as follows.

- What kind of effect would formal training have on the subjects' statistical reasoning in everyday life events, where people also have a great deal of knowledge and beliefs based on their everyday experiences?

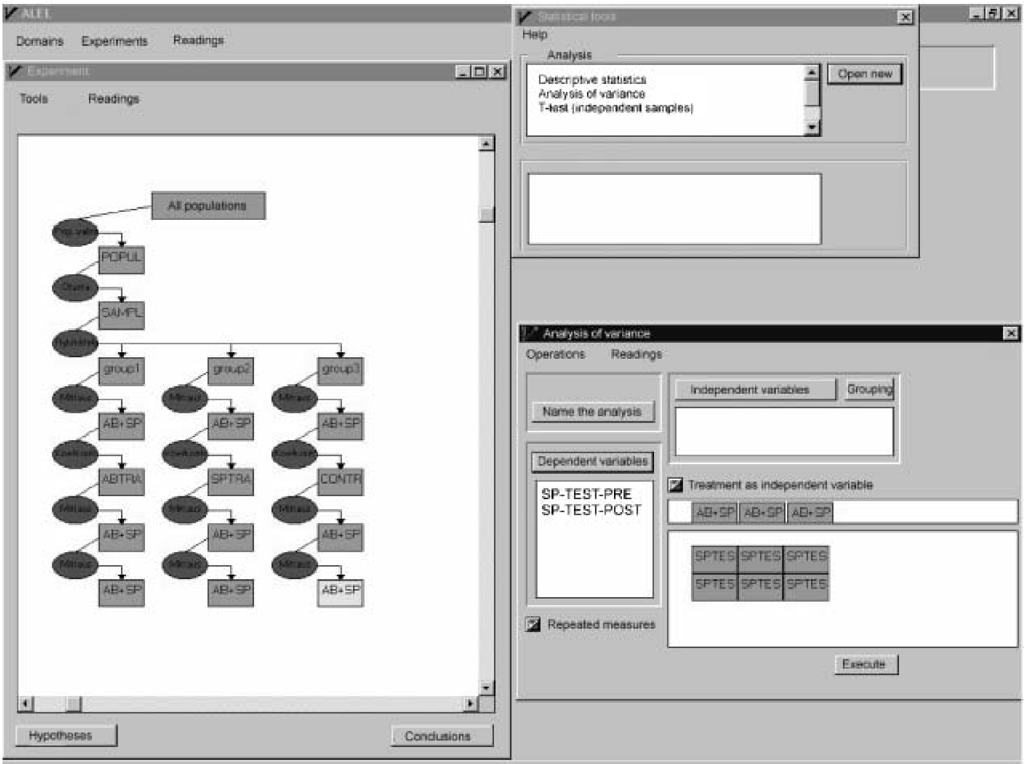

Figure 2. Tools for designing statistical analyses in ALEL

- Does training transfer to a domain other than the domain used in training?
- How permanent are the possible training effects?

When students constructed their designs they were able to choose from two training conditions and from two tests to measure statistical inference. Subjects could be trained in statistical reasoning either in the domain of sports (Sport Training) or in the domain of ability (Ability Training). For testing reasoning, the two different tests were available, based on sports examples (Sport Test) or ability examples (Ability Test).

Procedures

In order to obtain knowledge about the effectiveness of ALEL, three groups were formed. The first group used ALEL. The second, the article group, were provided with authentic, case-based tasks that were anchored to real world research. The third, the statistics group, were taught with the help of a statistics computer program.

These three groups had different teachers who all had a different perspective on teaching experimental research methodology and statistics. The goal of the course was discussed in detail with the teachers. Teachers saw the pre- and post-tests before beginning the course and were able to contribute to the test. When the goals of the course had been discussed, it was the teacher's duty to plan the course. Beside the purpose of teaching experimental research methodology and statistics, it was agreed that all the teachers would use student pairs as a part of their teaching and that the course would last 16 hours.

ALEL group. The first group worked with the ALEL computer program, which is designed to teach research methodology (Lehtinen & Rui, 1995). The course lasted for 16 hours. Two students' discussions during the course were audiotaped while they worked with ALEL as a pair. Pairs used the program four times, for 3 hours at a time, for a total of 12 hours. The system created a log file containing information about students' designs and statistical analyses. This information was used in analysing students' discussions and to answer the project's second main problem.

Article group. The idea was that the second group would study experimental research methodology by reading two articles and analysing and critiquing those articles using Tuckman's (1994) criteria. Tuckman's article consists of 25 questions to answer in analysing and critically evaluating a research study. The questions concern research problems, hypotheses, variables, operational definitions, control and manipulation, design of the study, results, discussion, and literature review.

Statistics group. For this group the purpose was to introduce the student to experimental design and statistical thinking from the statistical point of view. The aim was to become acquainted with experimental design and and the use of a user-friendly Finnish statistics program that allows students to analyse data without performing complex manual computations or command lines.

Students in all three groups were given a pre-test and a post-test. The post-test took place within a week after the last training session. The students had 1.5 hours to perform the test. The time between the pre-test and post-test was six weeks.

Results

The Three Groups Experiment

The results show that the ALEL group was significantly better than the other two groups in terms of total achievement. There was an interaction of group and development [$F(2,29)=4.54$, $p<.05$]. Only the group that used the ALEL computer program gained from the course (see Figure 3).

In the pre-test all groups performed equally well in terms of total achievement, factual knowledge, procedural knowledge, statistical knowledge, and practical design skill. In the post-test students in the ALEL group achieved significantly higher scores

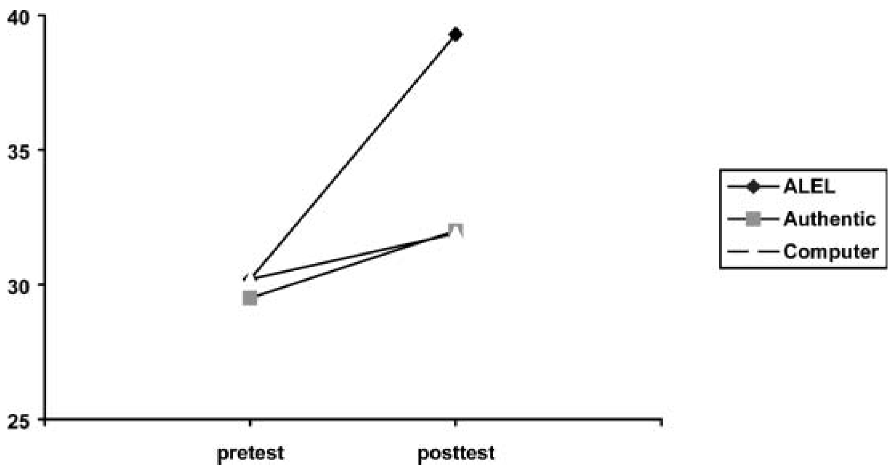

Figure 3. Comparison of three research methodology courses (maximum score 62)

Table 1. Achievement scores in pre- and post-tests

	ALEL group		Authentic group		Computer group		
	Mean	SD	Mean	SD	Mean	SD	*F* value
Factual knowledge (max. 22 points)							
Pre-test	16.0	4.1	14.3	3.8	14.7	4.1	0.517
Post-test	13.9	2.7	15.3	1.7	13.5	3.5	1.229
Procedural knowledge (max. 14 points)							
Pre-test	7.2	3.2	7.4	2.9	7.7	2.8	0.092
Post-test	10.7	1.0	8.0	2.7	7.6	2.7	6.142[b]
Statistical knowledge (max. 20 points)							
Pre-test	5.2	2.8	6.2	3.5	6.1	2.2	0.417
Post-test	12.1	3.6	7.4	3.4	8.6	2.9	5.713[b]
Practical design skills (max. 6 points)							
Pre-test	1.8	1.8	1.6	1.6	1.6	1.5	0.541
Post-test	2.5	1.9	1.3	1.3	2.2	1.6	1.572
Total achievement (max. 62 points)							
Pre-test	30.2	9.4	29.5	9.0	30.2	7.2	0.021
Post-test	39.3	7.5	32.0	5.4	31.9	7.9	3.877[a]

df=29. [a]$p<.05$. [b]$p<.01$.

on procedural and statistical knowledge, whereas for factual knowledge and practical design skills students performed equally well in all groups (see Table 1).

Description of Group Interaction Using ALEL

One of the ALEL groups consisted of Tiina and Maija. It was decided to select this pair for closer study of students' interactions and understanding during the learning process. Tiina and Maija seemed to represent the ALEL student group quite well on the variables measured by the pre-test. They were only slightly better in statistical

Table 2. Achievement scores in pre- and post-tests for the whole group (n=11) and for Tiina and Maija separately

	Whole group (n=11)		Tiina	Maija
	Mean	SD	Mean	Mean
Factual knowledge (max. 22 points)				
Pre-test	16.0	4.1	17	16
Post-test	13.9	2.7	12	14
Procedural knowledge (max. 14 points)				
Pre-test	7.2	3.2	7	7
Post-test	10.7	1.0	10	11
Statistical knowledge (max. 20 points)				
Pre-test	5.2	2.8	7	9
Post-test	12.1	3.6	13	15
Practical design skills (max. 6 points)				
Pre-test	1.8	1.8	2	1
Post-test	2.5	1.9	3	3
Total achievement (max. 62 points)				
Pre-test	30.2	9.4	33	33
Post-test	39.3	7.5	38	43

knowledge than the rest of the students. In the post-test Maija showed more improvement in statistical knowledge and also in total knowledge than Tiina (Table 2).

The gain in total knowledge was in procedural and statistical knowledge and in practical design skills. Students' performance in factual knowledge tasks was weaker in the post-test than in the pre-test.

Description of ALEL Lessons

During the first two lessons the teacher explained the basic principles and features of the program and gave concrete guidance on its use. The rest of the time, students worked in pairs. Each of these audio-taped lessons is summarized below. Descriptions of the two chosen students' actions during the lesson are given in brackets.

Lesson 3. When the students start working with ALEL they open the hypertext and read about the purpose and research questions of the study they are going to conduct. When they begin their own design, they are uncertain where to start. In the episode below the students discuss the difference between population and sample. Even though the students had completed a course in research methodology just prior to this, they are not certain what population and sample mean. They discuss it and seem to agree about the concepts, but later Maija wonders whether sampling should be done before selection of the population.

Tiina: Let's take selection of the population for example, let's start from there.

Maija: Isn't it the population from which we choose the subjects?
Tiina: Yeah.
Maija: It's supposed to represent the entire population.
Tiina: Yeah, or at least the population is the whole of that group, for example, that we're going to do the experiment with. Or … let's see what the population is.
[She suggests they read the hypertext, but the other student continues]
Maija: So if we were going to research young people, then young people it is.
Tiina: Yup.
(After a while)
Maija: Maybe we should have done the sampling before the population.
Tiina: No, sampling is done from the population.

It is not a simple matter for them to choose the population and sample. At first they choose an inappropriate population, but after talking with the teacher they change it and continue working with a new design (design 2). In the next episode the students are reading the hypertext about different sampling methods. Students read from the hypertext about random and systematic sampling, but they choose the sampling method on the basis of superficial signs in the text. They do not think about what the different sampling methods mean and how they affect their design.

Tiina: Random, systematic sampling: Which one of those is it? How are we supposed to know?
Maija: Maybe we should read.
[They open the hypertext]
Maija: … 'Its use as such is rare.'
[They read about random sampling]
Maija: Not that one at least!
Tiina: There it was … let's take systematic.
Maija: Systematic.
Tiina: Yes, here we take for every ….
[They read about systematic sampling]
Maija: Advantages, aren't there any disadvantages?
Tiina: …'easy to perform.' [mumbling when reading the text]
Maija: 'One disadvantage is that in the population and the variables under study there may be repeating characters which can make the results questionable. Repeating characters may appear ….' Not in ours.
Tiina: [Laughing] Maybe we should try that.
[They choose systematic sampling]

After choosing the sampling method they start thinking about dividing the sample into experimental and control groups and thinking how many groups there are in their design. The computer gives them critical feedback, since they make too many, too small groups. The students decide to start a new design (design 3). In the next episode the students are again dividing the sample into experimental and control groups. The students choose the grouping method and the number of groups in a trial and error manner. When the students do not get positive feedback, they immediately change their design, without discussing why their design is inappropriate. The feedback does not advise the students what to do. It just gives a clue that some decisions may not be the best ones or that the students need to think about them further.

> Tiina: And based on that, experimental and control groups. Yeah. Let's try. And the grouping method is random.
> Maija: Was it? I don't remember. And four.
> [They are trying to remember what kind of design they had the previous time]
> Tiina: Now it started, yeah.
> [They get feedback from the computer]
> Maija: Was it? Hmmm. What's wrong with it? Should we have taken another … (grouping method)?
> Tiina: Well, let's go there again. Can we do it that way?
> Maija: Why doesn't it take it? Now it took it. Let's change.
> Tiina: Let's try matching.
> Maija: Lets take matching now.
> Tiina: Let's try that because that didn't work either.
> Maija: Do we have it now …. Should we try with four groups, and if that's not OK, let's try with two.

After getting their first complete design (design 3) done, they start reading the hypertext about statistical analyses. They decide to carry out a χ^2 test and *t* test, because they read that those tests are appropriate for testing the difference between two groups. They start with the χ^2 test, but they do not know how to classify the variables. The teacher advises them to start with *t* test. They discuss with the teacher what the difference is between the *t* test with independent and dependent samples.

To summarize, during the first ALEL lesson the students struggled with trying to make a design piece by piece and learning to use the program. It seems that they expended all their effort on surface level actions and, because of that, the students' goal during the third lesson was not so much to learn research methodology as to get a task, a design, done.

They read the hypertext a few times, but they did not really concentrate on understanding what was written. They made their decisions on the basis of superficial signs in the text and it seems that they just wanted quick hints to further their design. They leant on computer feedback and changed their design until the feedback was positive. While constructing designs they did not try to answer the research questions, they just tried to get the task, the design, done. The students did not think about how the population, sampling, and grouping methods and sample size affected the design and the results. They did not seem to have developed an understanding of controls.

Lesson 4. The students start making a new design, since they had forgotten to save their design at the end of the previous lesson. In the next episode the students choose the measurements for their design. This time they try to more clearly construct a

DESIGN 1: **Hypothesis:** - **Population**: university students Students change the population after discussing with the teacher. **DESIGN 2:** **Hypothesis:** - **Population**: High school students **Sampling**: systematic, n=500 **Grouping**: matching based on school achievement, 5 experimental and 5 control groups (10 students in every group) → computer gives them negative feedback, so the students change their design
DESIGN 3: **Hypothesis:** Statistical training improves reasoning in practical problems **Population**: High school students **Sampling**: random, n=40 **Grouping**: random, 4 groups, 2 experimental and 2 control groups (10 students in every group) → since the computer gives them negative feedback, students change their design and choose 2 groups which are matched based on school achievement: one experimental and one control group **Measurements**: ability-testing of both groups in pre- and post tests **Treatment**: Experimental group is to receive ability treatment

Figure 4. Students' designs constructed during the third lesson

design that will answer the research problems. They are discussing how to answer the transfer problem. They wonder whether they should do a pre-test with one test and a post-test with another, so that they could answer the transfer question. They come to the conclusion that they cannot answer that question with their design.

Tiina: How do we test whether it will transfer [talking about the second research problem: does training transfer to a domain other than the domain used in training?]. If the pre-test is this ….
Maija: It would almost need another study
Tiina: We don't do it with this study
Maija: We cannot
Tiina: How do we get the transfer? What if we do a pre-test with this test and a post-test with another test. No!! What was the training … ?
Maija: It is the same as the treatment, that is the training.
Tiina: Yes, of course.
Maija: I think that after this whole thing we should make that transfer.
Tiina: Yes, we cannot do it in this same… let's do that. Let's do this first.

During this lesson the students spend a lot of time doing statistical analyses. They read the hypertext about them. With the help of the teacher they decide to do a *t* test. The following episode shows how the students have difficulties in choosing independent and dependent variables for the test. They do not know what the independent and dependent variables mean, even though they were able to define them in the pre-test. They discuss the concepts and display a clear misconception about what independent and dependent variables are. Tiina explains that the

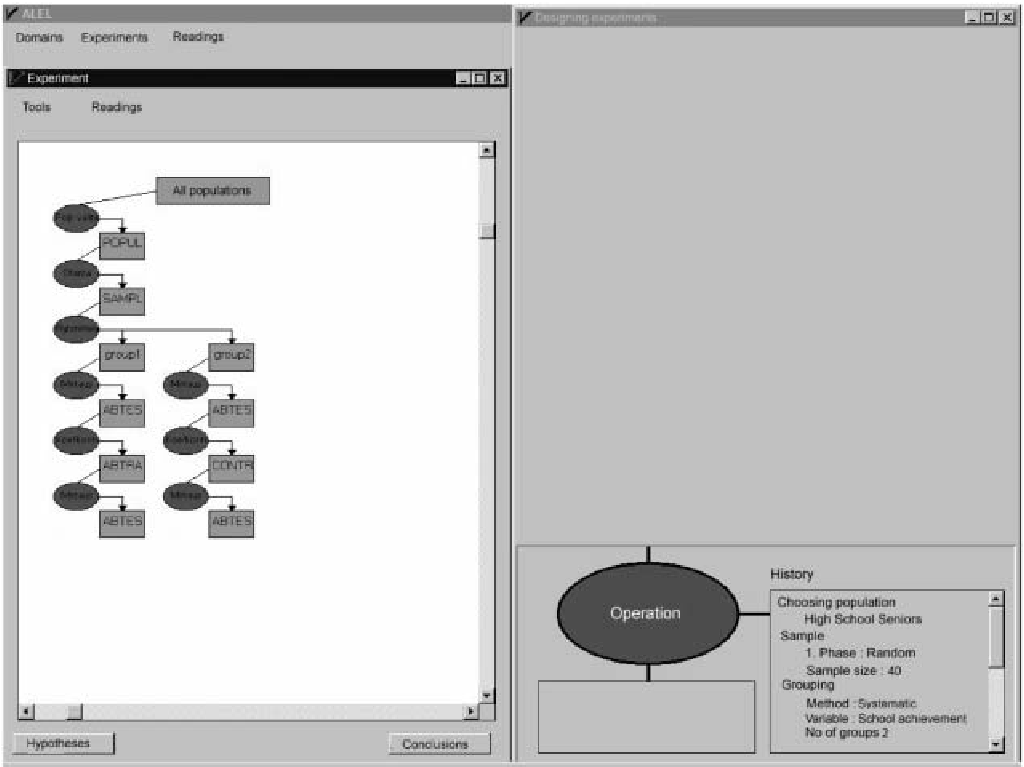

Figure 5. Students' first complete experimental design implemented in ALEL (design 3)

independent variable is the pre-test because it is independent of the treatment. The treatment does not affect the pre-test. Then she thinks that the post-test is the dependent variable, because the post-test is dependent on the treatment. Maija agrees with that.

The computer gives them feedback about their different choices, but the students do not understand what they did wrong. The teacher comes to help and explains that there are two possibilities in the program to choose independent variables: to take certain groups as an independent variable (for example, experimental and control groups) or classify a certain variable (for example, school achievement). He does not comment on the fact that the students have pre-test as an independent variable. The students and the teacher try to do the *t* test again, but the computer crashes. The students restart the computer and start doing the *t* test. They choose groups (one control and one experimental group) as an independent variable without discussing the meaning of these concepts. (The meaning of these concepts was not mentioned during the rest of the course, so it is not certain whether the students had changed their conception of the variables or had just mechanically chosen groups as independent variable because the teacher had mentioned it.)

The students do a *t* test and the teacher helps them to look at the data output. They are able to interpret the results. When the students draw conclusions they notice that they cannot obtain an answer to the transfer problem with their design. In the next episode the students try to reason how to improve their design, but their thinking is still quite vague. They are not certain how to make the changes and whether the changes are sensible. They have got results that confirm that the treatment has not been effective, that there are no differences between the experimental and control groups. Because of these results they wonder whether it is sensible to test those groups with another test to obtain answers to the transfer question.

Maija: And then about this second [problem] we can't say anything, because we haven't researched it (laughing)
Tiina: Well (laughing), well yes. How can we research it?
Maija: We could delay the post-test.
Tiina: Yeah, and do it with another test then.
Maija: So that it would have had time to affect then.
Tiina: Hey, but can't we do it still. Look, if we ….
Maija: But it has to do with the experimental design.
Tiina: But can't we still fix it?
Maija: Choose 'Making experimental design' and then 'Functions', 'Measurement'. [Advising Tiina to use the program] But we can't measure it with another test, can we?
Tiina: But how can we research that transference? If our result is that there is no effect, then what do we gain from this? (laughing)
[Their result was that the treatment had no effect]
Maija: Nothing, so let's go to the conclusions, because ….
Tiina: Let's see if it would accept anything (laughing). But then at least we get a hypothesis there as well. Now in a way we can't make a hypothesis for that second problem. Or we can't reject it.
Maija: No we can't.
Tiina: Let's do it and a Sport Test.
Maija: But can we measure it with that?
Tiina: But with what then, if it …. It feels kind of crazy to suddenly ask some questions about sports.
Maija : Yes, but it means that we can somehow generalize it into something else.
Tiina: Yes, let's try.

The students do *t* tests and learn to make inferences from the data output. They finish their research by writing down results and drawing conclusions.

To summarize, compared with the third lesson the students were not so much concerned about how to use the computer and the program, but concentrated more on making a design that would give an answer to the research questions. They had moved from a surface level (trying to develop a design quickly) to a problem level (trying to answer research problems). The students had good discussions about how to obtain an answer to the transfer question. They had some problems with doing the statistics, since they did not completely understand what the independent and dependent variables were, even though they were able to define them in the pre-test. The teacher helped the students with carrying out a *t* test and interpreting the output.

Lesson 5. At the beginning of the lesson the teacher comes to the students and asks them to explain their design (design 4). They explain their population, sampling method, and grouping into experimental and control groups. In the next episode the students explain their design further. It is easy for them to explain their design to the teacher and for the teacher to follow their explanation because of the tree diagram in the ALEL interface. The students explain that they can get an answer to the first question with their design, but that the transfer question is more difficult. They explain that they have performed a delayed post-test with another test (Sport Test), so that they could answer the transfer question. They also wonder whether they have performed appropriate *t* tests. The teacher asks the students to explain what kind of comparison they should make in order to obtain results for the transfer problem. During the discussion the students and teacher go through important information about what to take into account when constructing a design.

Teacher: What statistical tests did you already do?
Maija: A *t* test
Teacher: And between which groups?
Maija: We did it ….
Tiina: First between this experimental and control ….
Teacher: Was it?
Tiina: Independent variables.
Teacher: And did you get statistical significance?
Maija and Tiina: No.

DESIGN 4:
Hypothesis: 1. Statistical training improves practical reasoning 2. General statistical reasoning ability can be transferred to practical reasoning
Population: High school students
Sampling: random, n=40
Grouping: matching based on school achievement, 4 groups: 2 experimental and 2 control groups (10 students in every group)
Measurements: ability testing with both groups in pre- and post tests
Treatment: experimental group to receive ability treatment
Statistical analysis: t-test
Conclusions: Statistical training does not improve practical reasoning. General statistical reasoning ability does not transfer.

→ Later when the students were writing up the conclusions of the study, they realized that they could not get results on the transfer problem. In order to get results for that problem, students made a delayed measurement with a sport test for all experimental and control groups.

Figure 6. Students' design constructed during the fourth lesson

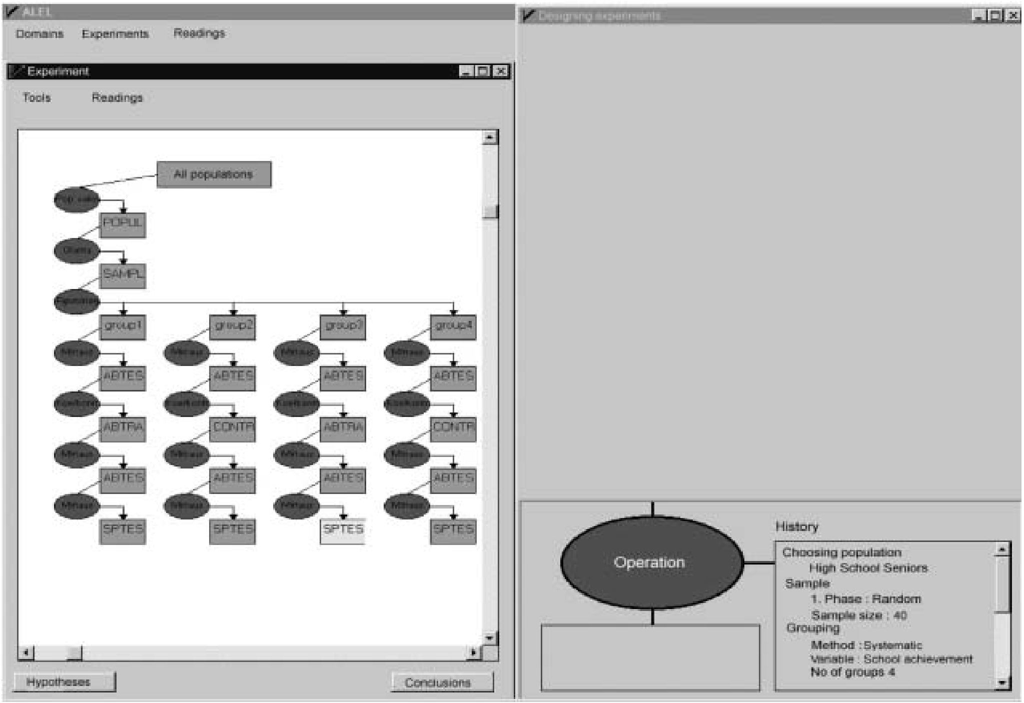

Figure 7. Students' second complete experimental design (design 4)

Teacher: Well, it might be that you have too small groups. So with such small groups you need to have a really big difference to get a statistically significant result, because statistical testing is quite sensitive to the size of the group. In other words, when you had that 40. Well, it's not that small a group, but when you divide it into four groups, then all of those begin to be quite small.
Tiina: Yeah
Maija: Then we did a test between these (statistical test).
Teacher: Yeah
Maija: I don't know whether you can interpret it then. I don't even know whether this test was even a good idea.
Teacher: Which one of these tests? This last one?
Tiina: No, as you can see, between this one and that one.
Teacher: It is a little .... What if you have two different tests (Sport Test and Ability Test) and then you compare the averages?
Tiina: On the other hand, how could we have researched that transfer, if not with two different tests?
Teacher: Yeah, but what should be compared if you look at whether the effect of the treatment affected the other test? You can study it from that. What do you need to compare?
(Long silence)
Teacher: How did you do this test in the same area ... how did you work that out? Did you compare those two?

Maija: Yeah
Teacher: Well. How can you compare whether or not there will be a difference from the treatment to a test in another domain?
Maija: That was the problem. We did not get any statistically significant difference ….
Teacher: Yes, but not in principle, because it's probably the smallness of the group and transfer that can be researched by comparing these two because there was treatment in one and not in the other. We can conclude that this is a group where there has been change and this is the group where, if there has been change, treatment will have an effect.
Tiina: But doesn't it matter that in our group it hasn't had an affect? Can we still compare those to each other?
Teacher: Of course you can. It can affect and then you could do the statistical analysis and then you can start thinking about how this design could be developed so that you could research all those elements that are here … what would be an economical way.

The teacher suggests that the students should carry out a certain *t* test. The students do it and draw conclusions about the study. Then they start a new design. During the following episode the students seem to have a more developed understanding of experimental controls. Even though the discussion is far from sufficient, the students are capable of comparing different designs and discussing different possibilities for getting results on the research problems.

Maija: But what should we put then ….
Tiina: As a treatment?
Maija: Yeah, well …. One (test) for one (group) and one for the other.
Tiina: But we can then, you see, if we study with one group whether the treatment will help and with another if it will transfer. So, with one group we do the Ability Test at the beginning and at the end and then with the other group we do the Ability Test at the beginning but a Sport Test at the end.
Maija: Yeah, or could we do them at the end like we were just trying to do?
Tiina: As well, you mean?
Maija: Like opposite tests?
Tiina: So we would do the same to both in a way?
Maija: So that we could be more certain.
Tiina: Yeah, but then we don't need the pre-test.
Maija: Or do we? What was it, what did you say first?
Tiina: So that for this first group we do an Ability Test for instance as a pre-test and an Ability Test as a post-test and with the other group we could do an Ability Test as a pre-test, so we could study transfer in the same design, 'cos in a way we have already got an answer from this first experimental group as to whether the treatment helps.
Maija: But then we would have to know the level of knowledge of the other group, where it has started to transfer, because they haven't had any treatment or ….

They complete their design and are enthusiastic, because they are now certain that they can obtain answers to all the research questions with their design. After that they carry out *t* tests without any problems and they know how to read the data output. The students are surprised by the results. Then they ask the teacher if there is a better analysis, because the *t* test does not seem to be suitable to obtain answers to the original research questions. The teacher explains analysis of variance to the whole group. The students start to carry out an analysis of variance, but it seems very difficult. They appear troubled and disturbed. They make comments like: "I cannot think any more" and "This is annoying". They do not know what to do. The teacher does the analysis for them, but they do not have time to look at the analysis more closely.

In sum, during the fifth lesson the students had an important discussion with the teacher. The teacher elucidated further about constructing designs, but also questioned the students and asked them to explain their design. After the discussions the students seemed to have a more developed understanding of experimental controls, even if it was far from sufficient. The students' design was also now better than earlier in the course. During this lesson the teacher explained analysis of variance to the whole group and demonstrated it to the students. Analysis of variance seems to be difficult to understand and do and the students expressed this verbally.

Lesson 6. The students start where they finished off the previous time. They read about analysis of variance from the hypertext. They do their own analysis, but they do not know how to read the data output. They make an error in their analysis and the teacher comes to correct it. They do an analysis together with the teacher. The teacher explains how to read the output. Now they seem to understand the difference between the *t* test and analysis of variance. The students write down the results and draw conclusions. They find out that their experimental group was better in the pre-test than the control group, so they plan another design controlling for the same level of knowledge of the experimental and control groups are at the beginning of the course. They complete the design and do an analysis of variance. Now the students seem to be able to read the data output.

DESIGN 5:
Hypothesis: 1. Statistical training improves practical reasoning 2. Statistical reasoning ability transfers in practical reasoning
Population: High school students
Sampling: random, n=90
Grouping: matching based on school achievement, 3 groups: 2 experimental and 1 control group (30 students in every group)
Measurements: ability testing and sport testing to both groups in pre-, post- and delayed tests.
Treatment: The students gave ability training to one experimental group, and to another experimental group, sport training. The control group did not get any training.
Statistical analysis: t-test and analysis of variance
Conclusions: ABILITY treatment yields statistically significant results and it also transfers. It is not possible to draw conclusions on the effect of SPORT treatment, since the control group had a low mean in pretest.

Figure 8. Students' design constructed during the fifth lesson

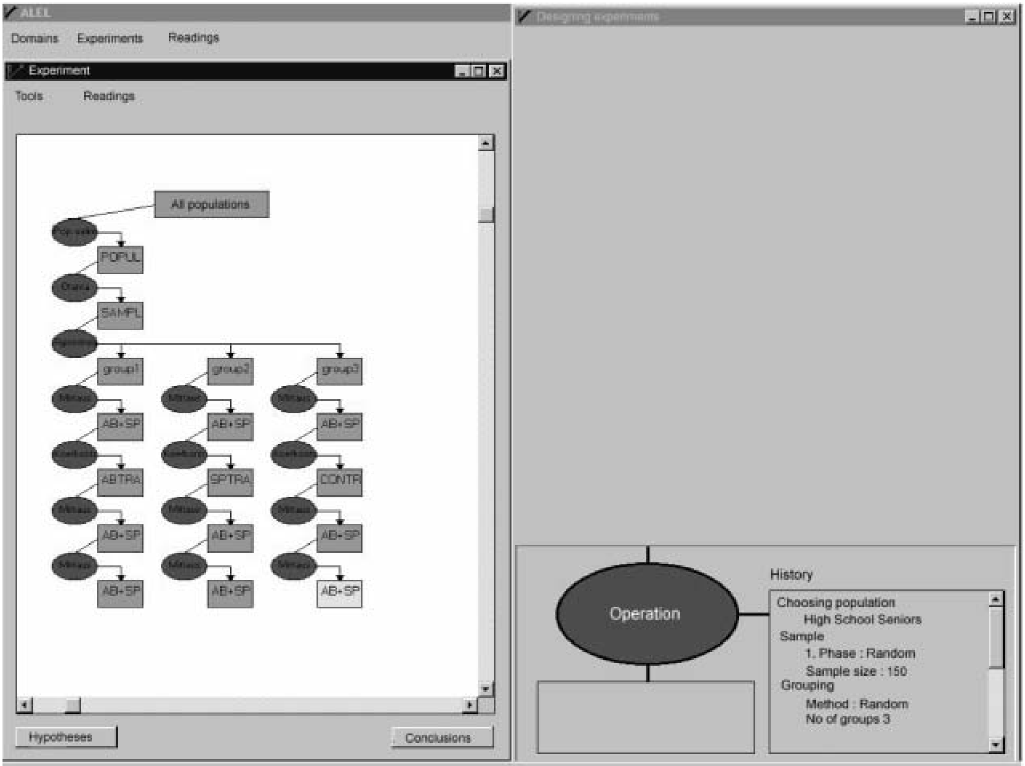

Figure 9. A later experimental design constructed by the case students (design 5)

During the sixth lesson students learnt the basics of analysis of variance and were able to compare the *t* test and analysis of variance. It did not seem to be so difficult and annoying any more. Constructing designs and controlling variables also seemed to be easier.

Conclusions

One aim of this study was to compare three ways of teaching experimental research methodology. According to the results the ALEL group outperformed the so-called

DESIGN 6.
Hypothesis: Treatment program has an effect on reasoning, which also transfers
Population: High school students
Sampling: random, n=240
Grouping: matching based on school achievement, 3 groups: 2 experimental and 1 control group (80 students in every group)
Measurements: ability testing and sport testing for both groups in pre-, post- and delayed tests.
Treatment: the students assign ability training to one experimental group and sport training to another experimental group. The control group did not get any training.
Statistical analysis: Analysis of variance
Conclusions: Treatment has a statistically significantly effect on reasoning: School achievement does not make any difference to success in training.

Figure 10. Students' design constructed during the sixth lesson

Article and Statistics groups in total achievement and also in procedural and statistical knowledge tasks, but not in factual knowledge and in tasks where knowledge must be applied, where it was only slightly but insignificantly ahead.

The surprising result that the students' performance in factual knowledge tasks deteriorated may have been due to another methodology course held just before the ALEL course on which students had to learn definitions of concepts and thus they were able to write them down in the pre-test. During the ALEL course students applied their knowledge instead of concentrating on learning definitions and concepts. The students' increased knowledge was shown in procedural and statistical knowledge and in practical design skills.

The ALEL teacher's role was crucial to the construction of the students' methodological knowledge. The teacher gave systematic guidance during the course. The tree diagram generated by the system made it easy for the teacher to immediately recognize the students' current problem solving situation and to focus the discussion on relevant questions. On some occasions more questioning would have been needed from the teacher in order to detect students' misconceptions. It is clear that misconceptions may develop and even prevail in students' conversations in such exploratory environments if the teacher does not question the students' knowledge. It appears that it is impossible to design feedback that can substitute for a teacher (see Lepper, Woolverton, Mumme, & Gurter, 1993).

In the beginning it appeared that the students' goal was not so much to learn research methodology and statistics, but to get a task, the design, done. It was difficult for the students to construct a design and they did not have the vocabulary to discuss it. At the end of the course they had more task-related and relevant discussions that showed a deeper understanding of research methodology.

The students' improvement was mostly seen in an increase in statistical knowledge. In the ALEL environment the learning of statistics is embedded in the whole research procedure. Statistics is then learned in a meaningful context. Students try to get results to research problems that they have themselves planned. This possibility appears to be very important for the learning of statistics. The teacher's advice was also crucial to the learning of statistics.

As for the second aim of the study, to obtain further insight into the interactions and cognitive behaviour in the ALEL group that helped or, alternatively, complicated the learning process, the discussions of one student pair clearly indicated development of a deeper understanding of methodological and statistical knowledge during the repeated design processes. The possibility of easily constructing several designs appeared to be very important for the knowledge construction process. As the students' knowledge deepened, they constructed more elaborate designs. However, the discussions also showed how complex the learning of research methodology is. Even though the students just prior to the course had completed a lecture course on research methodology, they had many problems in applying their knowledge. For example, the students found it difficult to choose independent and dependent variables for statistical analysis, even though the concepts had been explained in detail in the pre-test.

The study also revealed other merits of the ALEL method. Compared with other experimental research simulations ALEL differs in how quickly the experiments are carried out. In ALEL students carry out many steps in designing the experimental research. The aim is that the students start from the very beginning, thinking about the hypothesis, population and sample, different measurements, and treatments. When the design is ready, students carry out statistical analyses and interpret the results. When their design is ready, the aim is that the students elaborate on it, trying to improve it to better answer the research questions. When students using, for example, LABSIM construct complete designs in 5 minutes (Eamon, 1980) it is doubtful that they obtain a complete picture of the research as a whole.

The feedback in ALEL did not optimally help students to reflect on their thinking and actions. On the contrary, if the students did not get good feedback, they immediately changed their design, without discussing why their design was inappropriate. On the other hand, feedback helped the students to construct a possible and decent design, a process which would have been slower and more frustrating without the feedback. In Rieber and Parmley's (1995) study the performance of the subjects who only received an unstructured (pure) simulation fell short of that of subjects who received a tutorial.

The students used the hypertext regularly. It was easy for them to use, because it always opened a page related to the students' activity. Because of this quality in ALEL the students did not have any navigation problems. Other problems did occur in their use of the hypertext. At the beginning of the course especially students made their decisions about the design on the basis of superficial signs in the text. They were just trying to find quick hints in order to proceed with their design. It is a great challenge to write hypertext that encourages students to try to comprehend. It is also important that information is available at all times. Leutner (1993) found that information that is available at all times helped learners to acquire domain knowledge, whereas information that is provided before the simulation was ineffective.

The results show that the learning of research methodology and statistics is complicated. ALEL helps the learning process, but an expert tutor is also needed to detect misconceptions and to help the students with constructing the design and doing the statistical analysis. The results also show that the feedback and hypertext were not enough to "teach" the subjects. They show how important it is to follow students working with computer based-learning environments in addition to using pre- and post-tests.

This study, as well as that of Lehtinen and Rui (1995), have shown that, with the help of ALEL, methodology and statistics learning can be helped considerably and the problems of traditional methodology and statistics teaching can be at least partially solved. The purpose of this study was also to obtain more information about the use and important features of ALEL.

The results from the process descriptions show that although the ALEL students performed much better than the other students, their learning outcomes could be still better in terms of meeting the learning goals of the methodology and statistics curriculum. An inadequate knowledge base was shown in the post-test and also in the discussions of the pair during the course.

The best way to use ALEL is a question for further studies to research. On the basis of an expert–novice study on constructing experimental designs, Schraagen (1993) suggested that students should be presented with high level control structures in such a way that the teacher would design experiments for the students and, at the same time, teach explicitly strategic design knowledge. What would have happened if the principles of good design had been explained to the students first and only after that had they constructed their own designs in different areas?

References

Achtenhagen, F. (1994). How should research on vocational and professional education react to new challenges in life and in the working place. In W. J. Nijhof, & J. N. Streumer (Eds.), *Flexibility in training and vocational education*. Utrecht: Lemma.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*. New York: Springer.

Bell, J. A. (2001). Length of course and levels of statistics anxiety. *Education*, *121*(4), 713–716.

Benedict, J. O. (1979). CLASCONSIM: A computer program to simulate experiments in classical conditioning. *Behavior Research Methods & Instrumentation*, *11*(6), 603–604.

Bergland, M., Klyczek, K., Lundeberg, M., Mogen, K., Nelson, M., & Johnson, D. (2001). Case it!: Enhancing case-based learning in biology education through computer simulation and Internet conferencing. *Teaching with Technology Today*, *7*(9), Retrieved from http://www.uwsa.edu.ttt/bergland.htm.

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

Boshuizen, H. P. A., Schmidt, H. G., Custers, E. J. F. M., & Van de Wiel, M. W. (1995). Knowledge development and restructuring in the domain of medicine: the role of theory and practice. *Learning and Instruction*, *5*(4), 269–284.

Cutler, S. J. (1987). The A.C.E. freshman survey as a baseline instrument for survey projects in research methodology courses. *Teaching Sociology*, *15*, 121–127.

Eamon, D. (1980). LABSIM: A data-driven simulation program for instruction in research design and statistics. *Behavior Research Methods & Instrumentation*, *12*, 160–164.

Edwards, R. E. (1996). *LABSIM. Experimental design and data analysis simulator*. (version 9). Pacific Grove, CA: Wadsworth Publishing.

Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology*, *120*(1), 34–45.

Forbach, G. B. (1979). EXPER SIM: Review and update. *Behavior Research Methods & Instrumentation*, *11*(5), 519–522.

Forte, J. A. (1995). Teaching statistics without sadistics. *Journal of Social Work Education*, *31*, 204–219.

Järvelä, S., Lehtinen, E., & Salonen, P. (2000). Socioemotional orientation as a mediating variable in teaching learning interaction: Implications for instructional design. *Scandinavian Journal of Educational Research*, *44*(3), 293–306.

King, A. R., & King, B. F. (1988). The redesign of PROJECT SIMULATION for microcomputer-assisted instruction in psychology and research methodology. *Social Science Computer Review*, *6*(1), 75–89.

King, A. R., King, B. F., & Williamson, D. A. (1984). Computerized simulation of psychological research. *Journal of Computer-Based Instruction*, *11*(4), 121–124.

Kissler, G. R. (1974). Evaluation of computer-based laboratory simulation models to teach scientific research strategies. *Behavior Research Methods & Instrumentation*, *6*(2), 124–126.

Lajoie, S. P., Lavigne, N. C., Guerrera, C., & Munsie, S. D. (2001). Constructing knowledge in the context of BioWorld. *Instructional Science*, *29*, 155–186.

Lehti, S., & Lehtinen, E. (1999). *Computer-based authentic instruction in teaching empirical research methodology*. Paper presented at the Proceedings of Managing Learning Innovation Conference, Lincoln, UK.

Lehtinen, E., Nurmela, K., & Salo, A. (2001). *Case-based learning in CSCL environment*. Paper presented at the Ninth European Conference for Research on Learning and Instruction, Fribourg, Germany.

Lehtinen, E., & Rui, E. (1996). Computer-supported complex learning: An environment for learning experimental methods and statistical inference. *Machine Mediated Learning*, *5*, 149–175.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurter, J.-L. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie, & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 75–105). Hillsdale, NJ: Lawrence Erlbaum.

Lesgold, A., Lajoie, S. P., Bunzo, M., & Eggan, G. (1992). A coached practice environment for an electronics troubleshooting job. In J. Larkin, & R. Chabay (Eds.), *Computer-assisted instruction and intelligent tutoring systems: Shared goals and complementary approaches* (pp. 201–238). Hillsdale, NJ: Lawrence Erlbaum.

Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. *Learning and Instruction*, *3*, 113–132.

Levy, M. C., & Ransdell, S. (1999). *Laboratory in cognition & perception*. (version 3). Gainesville, FL: Psychology Software.

Magel, R. (1996). Increasing student participation in large introductory statistics classes. *American Statistician*, *50*(1), 51–56.

Mandl, H., Gruber, H., & Renkl, A. (1995). Knowledge application in complex situations. In S. Vosniadou, E. De Corte, & H. Mandl (Eds.), *Technology based learning environments. Psychological and educational foundations* (pp. 40–47). Berlin: Springer.

Murtonen, M., & Lehtinen, E. (2003). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education*, *28*(2), 171–185.

Ohlsson, S., & Lehtinen, E. (1997). Abstraction and the acquisition of complex ideas. *International Journal of Educational Research*, *27*(1), 37–48.

Onwuegbuzie, A. J. (1997). Writing a research proposal: the role of library anxiety, statistics anxiety, and composition anxiety. *Library & Information Science Research*, *19*(1), 5–33.

Onwuegbuzie, A. J., & Seaman, M. A. (1995). The effect of time constraints and statistics test anxiety on test performance in a statistics course. *Journal of the Experimental Education*, *63*(2), 115–124.

Ransdell, S. (2002). Teaching psychology as a laboratory science in the age of the Internet. *Behavior Research Methods, Instruments, & Computers*, *34*(2), 145–150.

Rieber, L. P., & Parmley, M. W. (1995). To teach or not to teach? Comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. *Journal of Educational Computing Research*, *14*, 359–374.

Rui, E., Suntio, S., & Lehtinen, E. (1995). *University students' knowledge of experimental research methodology*. (Tech. Rep.). Joensuu, Finland: University of Joensuu Research Center for IT in Education.

Salmi, S., Lehti, S., & Lehtinen, E. (1999). Using multimedia in integrating problem and discipline based approaches. Multimedia simulation in medical education. *Proceedings of Managing Learning Innovation Conference*. Lincoln, UK.

Schmidt, H. G., & Moust, J. H. C. (2000). Factors affecting small-group tutorial learning: A review of research. In D. Evenson, & C. E. Hmelo (Eds.), *Problem-based learning* (pp. 19–51). Mahwah, NJ: Lawrence Erlbaum.

Schraagen, J. M. (1993). How experts solve a novel problem in experimental design. *Cognitive Science*, *17*(2), 285–309.
Singer, J. D., & Willett, J. B. (1990). Improving the teaching of applied statistics: Putting the data back into data analysis. *American Statistician*, *44*(3), 223–230.
Steinkuehler, C. A., Derry, S. J., Hmelo-Silver, C. E., & Delmarcelle, M. (2002). Cracking the resource nut with distrubuted problem-based learning in secondary teacher education. *Distance Education*, *23*(1), 23–39.
Suter, W. N., & Frank, P. (1986). Using scholarly journals in undergraduate experimental methodology courses. *Teaching of Psychology*, *13*(4), 219–221.
Thompson, W. B. (1994). Making data analysis realistic: Incorporating research into statistics courses. *Teaching of Psychology*, *21*(1), 41–43.
Thurmond, J. B., & Cromer, A. O. (1975). Models and modeling with the Louisville experimental simulation system (LESS). *Behavior Research Methods & Instrumentation*, *7*(2), 229–232.
Tuckman, B. (1994). *Conducting educational research*. (4th ed.). Fort Worth, TX: Harcourt Brace.
Winn, S. (1995). Learning by doing: Teaching research methods through student participation in a comissioned research project. *Studies in Higher Education*, *20*(2), 203–214.