Journal of Computer Assisted Learning (2000) **16**, 229-242

Learning physics with a computer algebra system

E.R. Savelsbergh, T. de Jong & M.G.M. Ferguson-Hessler
Eindhoven University of Technology and University of Twente, The Netherlands

Abstract To become proficient problem-solvers, physics students need to form a coherent and flexible understanding of problem situations with which they are confronted. Still, many students have only a limited representation of the problems on which they are working. Therefore, an instructional approach was devised to promote students' understanding of these problems and to support them in forming associations between problem features and solution methods. The approach was based on using the computer algebra software *Mathematica* as a tool for problem solving and visualisation. An electrostatics course module was implemented based on this instructional approach, and this module was compared with a usual paper-and-pencil based one. Learning outcomes for both courses were not significantly different. The experimental course was found to impose a high cognitive load on the students. Based on the outcomes, proposals are made for ways in which the course could be improved.

Keywords: Cognitive load; Control group; Interactive learning environ-ment; Physics; Post-secondary; Problem solving; Tutorial; Visualisation

Introduction

In introductory physics education a rich understanding of situations is estimated higher than procedural ability (Rutherford & Ahlgren, 1991; Gravenberch, 1996). When students start to learn calculus-based physics the emphasis is shifted. Although situational understanding and the ability to identify a problem remain crucial to deep understanding and proficient problem solving (Chi *et al.* 1981; Larkin, 1983; see also Nathan *et al.* 1992), learning to carry out solution procedures simply consumes students' attention and takes up the available time. Therefore, it has been unavoidable that more challenging situations are postponed until procedural mastery has been achieved. Recent developments in user-friendly computer algebra software may offer new opportunities to do some more substantial situation analysis in calculus-based physics.

Accepted: 23 January 2000

Correspondence: Elwin R. Savelsbergh, Centre for Science and Mathematics Education, University of Utrecht., PO Box 80.008, 3508 TA Utrecht, The Netherlands. Email: e.r.savelsbergh@phys.uu.nl

A computer algebra system (CAS) in itself is no more than a high level programming language for symbolic and numerical computation. The first computer algebra systems, which became available in the late 1980s, were mainly of only theoretical interest. Over the last decade, some of these packages have evolved into more practical computation and visualisation tools that can take over many routine problem solving tasks. At the same time the required hardware has become more affordable. *Derive*, one of the more limited programs, is already implemented on a pocket calculator, and more extensive packages, such as *Mathematica* and *Maple*, run on any modern desktop computer. In several branches of mathematics, physics and engineering, computer algebra packages have an increasing popularity as a tool for constructing proofs and solutions. Also in introductory mathematics courses at the university level, there is an increasing use of computer algebra software. There are fewer examples where computer algebra is integrated throughout an introductory physics course. That is not to say that computers have been non-existent in physics courses, but their use has been restricted to numerical applications, which are not central to an introductory calculus-based course. This implied that the central part of the course — introducing the theory, and proving the formulas — had to be done by hand, like many student assignments. In this paper, it will be argued that a CAS could be used to promote the students' understanding of problems and to support the formation of associations between problem representations and solution information and a didactic approach for using such software will be suggested.

These ideas were applied to an introductory electrostatics course. This paper first discusses the current course and its problems; then the design of a learning environment and an experiment in which the effectiveness of the module was tested will be described. Finally, there is a discussion of design improvements and design guidelines inferred from the outcomes of the experiment.

The electrostatics curriculum

The domain of the study is an electrostatics course module taken from the standard curriculum for first-year physics majors. The module is taken as part of a longer course on electrodynamics. Topics covered in this module include charge distributions, symmetries, Coulomb's law, Gauss' law, dipoles, multipoles, conductors, computation of potentials with given boundary conditions, dielectrics, and polarisation.

The course has three major components: lectures, work groups, and homework. Lectures last two hours, and a typical audience is about one hundred students. In the lectures, the theory is presented and examples of typical problems are worked out. During work groups, small groups or individual students are assigned a set of problems to solve. One tutor for every 20 students assists if problems arise. Students are expected to solve additional problems at home and to study the course text (Griffiths, 1989). The projected total workload for the course is 80 hours for the average student.

The main aim of the course is to give students a thorough understanding of fundamental concepts and approaches. Here, the groundwork is laid both for more advanced theoretical courses and for application-oriented technical courses. The concepts taught are discussed in a simplified way, and the methods presented are only practical for some idealised problems. However, because later courses build on the material, students need to become fluent with the basic concepts, relations between them, applicability of the methods, and assumptions underlying them.

It is a well-known problem that students, even though they have learned the methods, fail to see how and when these methods can be applied in new situations. More specifically, in this course, students fail to see how they can use the geometrical properties of a situation to simplify the problem. This might be explained from the more general finding that novices in this field do not integrate solution information in their mental problem representations (Savelsbergh *et al.*, 1998a). A further problem is that weak students frequently seem to misunderstand problem descriptions and often fail to make proper drawings of situations (Van Heuvelen, 1991; Feiner-Valkier, 1997). This might be explained by the students' inability to form a coherent understanding of the problems (Savelsbergh *et al.*, 1998a), and by the weak students failing to switch between propositional and pictorial representations (de Jong & Ferguson-Hessler, 1991). A final, related problem is that students pay little attention to problem analysis, instead starting to manipulate formulas right away (de Jong & Ferguson-Hessler, 1984).

Based on these findings it can be assumed that students could benefit from a training procedure promoting the construction of problem representations. The primary goal of this study was to help students construct an integrated model of the situation, and to connect this situation representation to solution information.

Design of the learning environment

Current learning theories (such as constructivism, schema theory, and production rule theories) suggest that problem representations are best constructed by students themselves, and that an adequate problem representation has to be constructed in the context of real problem solving activity (Glaser & Bassok, 1989; Corbett & Anderson, 1992). Therefore, the approach used in this study was to support the formation of problem representations during practice problem solving aiming to make a proper situation analysis intrinsically rewarding, rather than having it imposed by a teacher. A review of several learning tools led to the conclusion that a CAS may offer the right functionality to achieve this goal (Savelsbergh *et al.* 1998b).

Three properties of CASs are of importance: first, CASs demand precise specification of problems, in a highly constrained formal specification language; second, CASs take over algebraic calculations; and finally, most CASs have visual-isation facilities. The required precise specification of the problem and the assistance in algebraic calculations can be used to direct students' attention to the properties of the problem situation. Once a first case has been worked out, situation properties can easily be manipulated, and the solution of a first case can be reused in a following case, provided that the situations match. In addition, the algebraic support may help students to focus on the main line of the solution rather than algebraic calculation details. Finally, visualisation facilities may be used to construct graphical representations of situations and solutions that would otherwise remain abstract.

Because the intended participants in the study already had some experience working with *Mathematica*, it was decided to use that program. *Mathematica* (version 3.0) has an advanced hypertext interface that can present pictures and conventionally formatted mathematical formulae alongside the user-typed input. A course module written in *Mathematica* can be set up as an 'interactive book', with all the necessary information presented on-screen. It could thus provide an integrated learning environment that would present the theory in brief, then worked examples, after which various types of assignments would follow. Theory was presented only briefly, because a book is more convenient for extensive reading. The first assignments on a topic were highly structured, requiring the learner to modify something in a worked example or to complete an incomplete solution. Later assignments were more open, requiring the learner to construct the entire solution. This set-up was intended to minimise extraneous cognitive load (cf. Sweller, 1988).

Worked examples helped to reduce the effort that goes into mastering the programming language (Sweller & Cooper, 1985; Zhu & Simon, 1987; Paas & Van Merriënboer, 1994). The structured assignments were intended to focus on variations of situation properties within a class of problems, and on the consequences the properties have for the solutions. Thus, structured assignments may help the students to identify relevant properties of a situation. The visualisation assignments may help to connect a concrete physical representation to the abstract formalism of both the situation and the solution. Finally, practice problems require the problem solver to elaborate the problem statement and, moreover, they provide a training opportunity.

The experimental course was intended to represent 8–10 hours of workload for the average student. The general subject of the course was "special techniques for calculating potentials." There were four sections: general introduction and instruction on *Mathematica*; introduction of E-field and potential; image charges; dipole and multipole expansion.

Figure 1 presents a brief example of a topic in the experimental course (Savelsbergh, 1998; for more extensive examples see Savelsbergh et al., 1998b). The example deals with computing the field of a point charge near a planar conductor. The examples starts with a summary of the relevant theory, followed by a worked example. In this worked example the text in lines labelled 'In' were already present when a student started working with the material. When the student executed the commands in these lines, by pressing SHIFT + ENTER, the computer-generated output labelled 'Out' appeared. The student could also modify the input lines and then execute the commands again to examine the effects of a different situation. In the figure, the worked example is followed by parts of two structured assignments. The first assignment required the student to check that the solution found in Out [11] satisfies the condition that, at a large distance from the point charge, the potential approaches zero. This required using the limit function, as shown in In [13]. The second assignment required plotting a graph of the solution found in Out [11]. Here, the student had to identify the region of interest, and to think of a suitable type of picture. The input shown in these examples (In[13] and In[14]) has to be thought up and typed in by the students themselves. To facilitate entering symbols and expressions, a floating 'palette' was provided in a separate window. The students could pick an element from the palette by clicking on it.

In the actual course, the assignments shown in Fig. 1 would be followed by an open assignment. This assignment would be of the same type as those the students worked on at their regular work groups, or at the final test (Fig. 2).

Classroom evaluation

The experimental course was expected to improve students' understanding of physical situations and to strengthen the relation they see between solution methods and situation features. To test these expectations, learning outcomes of students in the experimental course were compared with those of students who attended extra training on their normal practice problem. So, there were two conditions:

- a computer group where the experimental course module was used and

- a control group where the students used paper and pencil to solve problems similar to the ones presented as open assignments in the computer course.

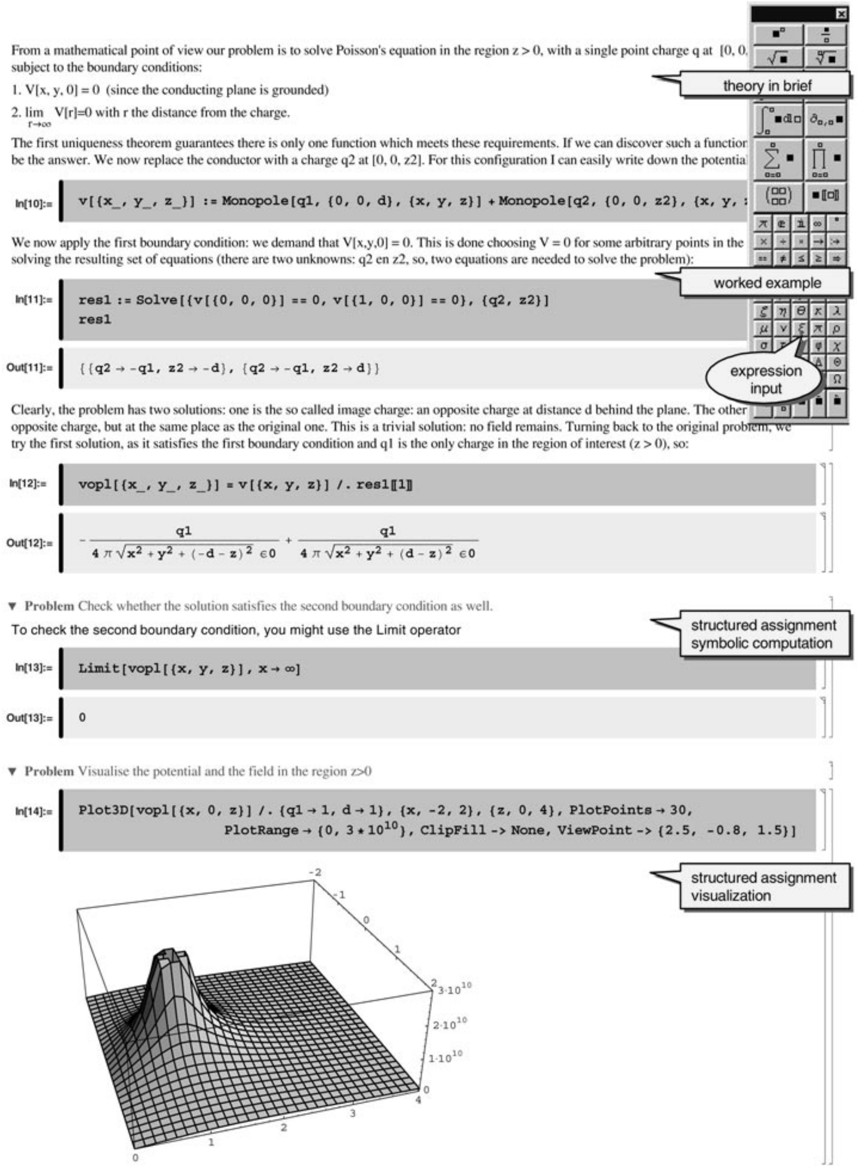

Fig. 1. Brief example from the section on image charges in the experimental course (translated).

The experimental course was expected to improve students' understanding of physical situations and to strengthen the relation they see between solution methods and situation features. To test these expectations, learning outcomes of students in the experimental course were compared with those of students who attended extra training on their normal practice problem. So, there were two conditions: a computer group where the experimental course module was used and a control group where the students used paper and pencil to solve problems similar to the ones presented as open assignments in the computer course.

In a spherical region in vacuum (radius R, permittivity ε_0), an inhomogeneous spatial charge density is given:

$$\rho(r) = \rho_0 \left(\frac{r}{R}\right)^n$$

Here r is the distance from the center ($r < R$), ρ_0 is a positive valued constant, and n is a whole number greater than or equal to zero. The potential is defined as zero at infinity. For convenience we introduce:

$$Q_0 = \tfrac{4}{3}\pi R^3 \rho_0 \text{ and } E_0 = \frac{Q_0}{4\pi\varepsilon_0 R^2} = \frac{\rho_0 R}{3\varepsilon_0}$$

(a) Determine the total charge in the spherical region; express the answer in Q_0 and n
(b) Determine the electric field for points outside the sphere; express the answer in E_0, n, and R/r
(c) What is the potential at the surface? Express the answer in E_0, R, and n.
(d) Determine the electric field for points inside the sphere, express the answer in E_0, n, and R/r
(e) What is the difference in potential between the center and the surface? Prove that

$$V(\text{center}) = \left(\frac{n+3}{n+2}\right) V(\text{surface})$$

Fig. 2. Example of a final test item

Method

The experiment was conducted at the Faculty of Physics at Utrecht University, The Netherlands. There are approximately 90 first-year physics students. All these students were contacted individually and invited to participate. In total, 42 students agreed to participate. These students were assigned to the two experimental groups by the experimenters. On the basis of available prior performance scores, two groups were formed of equal ability. Because of the limited number of computers available, the computer group had to be split up into three equally sized subgroups taking the course at different times. To maintain equal group sizes, the control group was also split into three subgroups. In the first session 33 of the 42 students attended (17 in the computer group and 16 in the control group). All remaining students attended the next sessions. Students in both groups were paid Dfl 50 (approximately US$ 25) at the end of the experiment.

Spread over three sessions, on different days, participants had 7 hours available for the learning task, which is somewhat shorter than the projected time required to complete all assignments, so that even the quicker students had to use all the time available. Students in both groups worked under the supervision of a tutor, who was available to answer any questions. Students in both groups were free to co-operate or to work individually.

Results

Learning outcomes

As a measure for learning outcome, the final examination that students have to take at the end of the regular electrostatics course were used. This test has three assignments, consisting of 14 subproblems in total. A typical example of a final test item is presented in Fig. 2. In total, 56 students took this final test: except for one student from the computer group, all students who had participated in the experiment also took the final test. To assess the reliability of the test score, Cronbach-α was used taking the three main problems in the test as the independent items. With all 56 test-takers included a reliability of $\alpha = 0.85$ was found. To assess the significance of any between-group difference, the students' prior performance level as a covariate was included, to correct for prior differences between the groups. Prior performance was determined on the basis of high school final examination scores for physics and mathematics, and previous university examination scores for mechanics and relativity theory. In this case, data were on 57 students and the scale was shown to be reliable, $\alpha = 0.87$. With the inclusion of this covariate in the analysis, no significant effect of the experimental treatment remains, $F_{1,29} = 0.26$, $p = 0.77$. A second, and equally important issue is whether the treatment reduces or increases the difference between weak and proficient students. Here, again, no interaction between exper-imental condition and general performance level was found, $F_{1,28} = 0.70$, $p = 0.41$.

The results of participants in the experiment were compared with those of the students who had not participated. Comparison of prior performance scores indicates that mainly proficient students volunteered, $F_{1,55} = 8.17$, $p = 0.006$. Although the participants in the experiment were not drawn randomly from the population, the study can give some idea of the main effect of attending an extra 7 hours of instruction by comparing the final test scores of participants and non-participants. The prior performance level has to be included as a covariate again to compensate for the general difference in achievement level. Surprisingly, no evidence of a gain for taking extra instruction remains, $F_{1,53} = 0.50$, $p = 0.48$.

Students' evaluation

Apart from direct performance outcomes, students' opinions about the two versions of the course, and the learning process that took place in the computer group, were of interest. A questionnaire was constructed to assess students' opinions but because of limited resources, the only way to keep track of the learning process in the computer group was by informal notes kept by the experimenter who taught the course. There were two versions of the questionnaire, one for the computer group and one for the control group. Both versions of the questionnaire addressed the following major issues: number of assignments completed; difficulty of the subject; attractiveness and instructiveness; navigation and confusion; help and collaboration. Table 1 provides an overview of these items.

In the questionnaire for the computer group, there were additional items to assess the use and appreciation of features of *Mathematica,* such as graphing and computing integrals, and instructional elements, such as worked examples and practice assignments. Except for the item on the number of assignments completed, all items had to be answered on five-point Likert scales. Statistical analysis and further interpretation were done at the level of single items.

Certain items can be compared across groups. The first issue addressed was the number of assignments completed. Students in the control group had to solve the same assignments that were presented as open assignments in the computer group. The question was about how many of these assignments the student had completed. The worked examples and completion exercises that students in the computer group had to study prior to starting with open assignments were not included in the comparison. The average number of assignments completed was significantly lower for the students in the computer group, compared to students in the control group, $F_{1,31} = 29.1$, $p < 0.001$ (M = 3.6, sd = 2.1, and M = 7.0, sd = 1.5, respectively). This indicates that students in the computer group spent much of their time on worked examples and completion exercises.

Table 1. Evaluation items presented to both groups

Difficulty of the subject	
1.	Electrostatics is a difficult subject
2.	The electrostatics covered in the experimental course is difficult
Attractiveness and instructiveness	
3.	I learned a lot about electrostatics during the experimental course
4.	I liked the experimental course
Navigation, confusion	
5.	While I was working with the system (solving the exercises), I often forgot what I did before
6.	The system (I) usually could solve the problem, once I had entered (formulated) it correctly
7.	While I was working with the system (solving the exercises), I often lost an overview of things that appeared on the screen (I had already written)
8.	While I was working with the system (solving the exercises), I often had to look up things in earlier work
Help and collaboration	
9.	I often needed the tutor to help me
10.	I often discussed solutions with other students
11.	I often needed other students to help me
12.	I often gave help to other students
13.	When I needed help, it generally was about physics understanding

Answers were given on a five-point Likert scale with labels *agree* and *disagree*. For items where the wording differs across groups, adaptations for the control group are given between brackets.

There was no evidence that students preferred one version over the other, $F_{1,31} = 0.20$, $p = 0.89$ (Table 1, Item 4). In the further items that could be compared across both groups there were a number of differences. The students in the computer group judged the physics content of the course to be simpler than the students in the control group, $F_{1,31} = 4.48$, $p = 0.042$ (Table 1, Item 2). Because the difference was not reflected in the students' judgements about the difficulty of their compulsory course (Item 1), it must be concluded that the difference is in the content of the course module used in the experiment. This is easily understood because students who worked with the computer course spent so much of their time on the introductory assignments. Given this difference, it is not surprising that the students in the control group score higher on the question whether or not they had learned much about physics content.

Navigation and confusion was also explored. Students in the computer group clearly had more trouble finding their way than students in the control group, as becomes clear from the scores on Items 5 to 8 (Table 1), where the computer group had higher scores on all four items. The differences were marginally significant for Item 5, $F_{1,31} = 3.39$, $p = 0.075$, and significant for Items 6 to 8, $F_{1,31} = 7.31$, $p = 0.011$; $F_{1,31} = 10.34$, $p = 0.003$; and $F_{1,31} = 9.40$, $p = 0.004$, respectively.

There were marginally significant differences in the help required and the type of help required. The computer students required slightly more help, $F_{1,31} = 3.28$, $p = 0.08$ (Item 9), and this help was less focused on physics understanding, $F_{1,31} = 3.07$, $p = 0.09$ (Item 13). Scores for the students in the *Mathematica* group indicated that they needed more help on *Mathematica* than they did on conceptual physics problems, $F_{1,16} = 5.54$, $p = 0.03$ (repeated measures ANOVA).

The next part of the evaluation concerns specific features of the *Mathematica* learning environment. Several features were assessed on the following aspects: clarity, attractiveness, difficulty and instructiveness. Two of the items addressed inherent features of *Mathematica* itself, namely visualisation, and symbolic computation. Visualisation refers to the possibility of representing a found result in a graph, like the one in Fig. 1. Symbolic computation refers to the possibility of having the system evaluate integrals, gradients, divergences, etc. For these features students were asked how frequently they were used. Three other items addressed the elements that were built into the learning environment: worked examples, completion tasks and open assignments. The items are presented in Table 2.

Table 2. Some of the evaluation items that were administered to the computer group only

1.	I often used the possibility to visualise the fields and potentials that I had computed	true—not true
2.	I consider the possibility to visualise the fields and potentials that I had computed	(a) clear — unclear (b) boring — fun (c) easy — difficult (d) not instructive — instructive
3.	I often used the possibility to have gradients, divergences, integrals, etc. evaluated by the computer	true — not true
4.	I consider the possibility to have gradients, divergences, integrals, etc. evaluated by the computer	(a) clear — unclear (b) boring — fun (c) easy — difficult (d) not instructive — instructive
5.	I consider the worked examples	(a) clear — unclear (b) boring — fun (c) easy — difficult (d) not instructive — instructive
6.	I consider the completion assignments	(a) clear — unclear (b) boring — fun (c) easy — difficult (d) not instructive — instructive
7.	I consider the open assignments	(a) clear — unclear (b) boring — fun (c) easy — difficult (d) not instructive — instructive

Answers were given on a five-point Likert scale with labels as indicated

Outcomes are summarised in Table 3. It immediately becomes clear that, overall, visualisation is judged more favourably than gradients. The significance of the differences was tested by using a repeated-measures

Table 3. Evaluation outcomes for two features of *Mathematica* itself

	Amount of use	Clarity	Attractiveness	Difficulty	Instructiveness
visualisation	4.06 (0.97)	4.53 (0.62)[a]	4.06 (0.97)	2.71 (1.05)	3.94 (0.75)
symbolic comp	3.71 (1.21)	3.87 (0.96)[a]	3.65 (0.79)	2.29 (1.05)	2.41 (1.12)
worked examples		4.53 (0.80)	3.88 (0.70)	2.06 (0.56)	3.53 (1.33)
completion assignments		3.88 (0.86)	3.59 (0.51)	3.12 (1.11)	3.53 (1.12)
open assignments		3.12 (1.11)	3.53 (0.72)	3.88 (0.78)	3.82 (0.72)

Also includes the three elements implemented in the learning environment ($n = 17$). Values represent mean score and, between brackets, standard deviation

[a] $n = 16$ because of one missing value.

ANOVA. A marginally significant difference for clarity, $F_{1,14} = 3.65$, $p = 0.076$, and a clear difference in instructiveness, $F_{1,15} = 26.3$, $p < 0.001$ were found. For the comparisons of worked examples, to completion assignments, and to open assignments, there were significant differences in clarity, $F_{2,32} = 14.3$, $p < 0.001$, and difficulty, $F_{2,32} = 27.1$, $p < 0.001$. All pair-wise comparisons gave significant results so, as expected, the three elements were found to become progressively more difficult.

As the final part of the evaluation, remarks collected from the evaluation forms were examined with particular focus here on those collected in the *Mathematica* group. Some students expressed their general feeling about the course:

> Working with *Mathematica i*s a good supplement to problem solving on paper (ID8).
>
> After all it was fairly instructive, . . . A problem is, however, that while sitting behind the computer I can't concentrate on physics problems too well (ID15).
>
> It was fun to participate, but I did not find it very instructive . . . The paper-and-pencil work groups are far!! more boring (ID18).

Two students commented that the learning environment evoked a passive attitude:

> Solving the exercises tend to come down to using the 'copy' and 'paste' options of *Mathematica*. This did not contribute to understanding what really happened (ID6).
>
> [. . .] because of these worked examples you knew exactly what you had to do, so little initiative was required, whereas initiative should be important (you must be able to do it yourself, not to copy) (ID31).

There were some positive remarks specifically about the visualisation facilities:

> The benefit of the method is that the pictures give a good insight into what's going on. This may be helpful when you later come across a similar exercise. Pictures are easier to remember than formulas are (ID7).
>
> Working with *Mathematica* is a good supplement to problem solving on paper. The pictures give an insight into what you are working on (ID8).
>
> After all it was fairly instructive, especially the visualisation . . . (ID15).

Finally, some students commented on problems they had with *Mathematica*:

> Problems with *Mathematica* syntax cause a loss of time, especially during the first session (ID15).

> It is easy to lose your way; exercises were not hard but took a lot of time, because of irritating *Mathematic*a (ID30).
>
> It was instructive with regard to *Mathematica* (ID32).
>
> The difficulty was more in *Mathematica* than in physics. The longer I used *Mathematica,* the faster I worked, and the more I could concentrate on physics problems (ID33).

Discussion

The aim of the experiment was to compare the effectiveness of the newly designed form of tutoring to the traditional one. A benefit had been predicted for students taking the computer course, but no significant difference were found between the approaches. Firstly, the appropriateness of the instrument for assessing learning outcomes deserves attention. The final test had a strong focus on procedural ability, whereas in the course module the representation of the problem and the choice of a solution approach were central issues. Still, even with this limited test, any major differences between both groups would have been detected. So, the effect of the experimental treatment must have been limited.

An important factor explaining the lack of a gain for the newly designed instruction might be the high cognitive load in the computer course. On theoretical grounds learners were given full control of the learning environment and of their learning processes. This places a heavy cognitive load on the students, so that any additional extraneous cognitive load should be avoided. Yet both closed evaluation items and students' remarks suggest that students in the computer group were distracted from the physics content. Navigating through the hypertext environment already required some effort. More importantly, the approach to problem solving was quite new for the students, and, as most students had limited programming experience, they were not very systematic debuggers. Many failed to diagnose even expressions with non-matching parentheses. Cryptic error messages and unstable software might have aggravated matters.

Students' prior attitudes and expectations might have worked against the computer learning group too. They had worked with a previous version of *Mathematica* (version 2.2.3) in a programming course, and most students disliked it. Moreover, several students expressed the opinion that physics problem solving is best done on paper, and that you cannot learn physics via a computer. This belief must be reinforced at all those examinations where the first requirement is that students can fluently write out solutions.

The evaluation of the different components of the *Mathematica* learning environment indicates that the students found the visualisation facilities rather instructive. This is in line with observations by the experimenter, who had several discussions about common misconceptions that were triggered by visualisation outcomes. As an example, consider the following: one of the assignments was to draw a field plot for the field of a physical dipole, and compare this to the field of a mathematical dipole. One of the steps in the exercise was to zoom in on the physical dipole. Several students failed to understand why the plot essentially remained the same, and where they had to look for the changes. Such an impasse provides a good starting point for tutoring discussion. Although students gave high ratings for clarity of the visualisations, the quality of the plots could still be improved. Moreover, even though students considered the visualisations only moderately difficult, it appeared that students spent too much time on figuring out details of graphics commands.

The symbolic computation features were valued less positively. Although students judged them easy to use, they did not find them very instructive. A major cause could be that students fail to examine the computer-generated expressions critically, so that application of the symbolic computation facilities remains a trick. As indicated both by remarks on the evaluation form and by the experimenter's observations, too many of the assignments could be solved by mindless copying of the worked examples. This may be because students spent most of their time on the introductory, structured assignments. In any case, the opportunity to solve problems by just copying entire solutions must have worsened the students' attentiveness to computer-generated formulas.

Conclusions and recommendations

This paper has reported on the development and evaluation of a learning environment for a first-year university course on electrostatics. The goal was to support students in gaining an intuitive understanding of situations, solution methods and the relations between them. The demands of such a learning environment were analysed and a learning environment was built based on available software. Among the distinctive features of the learning environment are a precise language for specifying problems, visualisation support and symbolic computation support.

The first version of the experimental course was tested on a sample of first-year physics students and no significant differences between learning outcomes with the new approach and with the usual approach were found. Results indicate that the approach takes a considerable amount of time to get used to. Moreover, students did not see how the abilities taught in the computer course were relevant to their final examination, where they would have to solve the problems by hand anyway. If the students are to be won over to the approach, some items in the final test should be clearly related to the new approach. One option is to model these items after those in Hestenes' and Wells' Mechanics Baseline Test: that is, to place more emphasis on conceptual understanding of the situation (Hestenes & Wells, 1992); one could also construct items asking for a solution approach, rather than for a worked out solution. Allowing the use of a CAS at the test would be a final possibility.

Although no significant learning gain over the old approach was found there were several instances where the computer was helpful in addressing misconceptions, clarifying concepts that underlie solution methods, and supporting the construction of situation models. As there was no observation data from the control group, this result cannot be compared between groups. In the learning episodes in the computer group, discussions with the tutor played a central role. Therefore, it can be concluded that the role of the tutor in the experimental approach is as important as it is in the usual problem solving workgroups.

The learning environment in a normal classroom setting was chosen, in order to obtain a realistic image of its use. This approach also poses limitations, however, on the assessment of individual learning processes. It became evident, for instance, that students paid too little attention to computer-generated solutions, but there is no detailed information on their behaviours. Therefore, it could be helpful to have students work in pairs, and to observe them more closely, either in a classroom context or as a single pair.

To explore this approach further, it is necessary to improve the quality of the software and to refine the educational design of the course. Some problems, such as the instability of the software have been solved in the current version of *Mathematica* already. More comprehensible error messages are needed and it is necessary to see if other educators will implement similar courses.

A first reduction of programming effort could be easily achieved by offering tools that provide those graphics routines that do not contribute to the students' understanding of electrostatics, such as drawing the location of a point charge in a field plot. A similar tool could be constructed to screen the input for frequently occurring beginner mistakes, such as non-matching parentheses. A further improvement could be the introduction of dedicated calculator interfaces for tasks such as solving Laplace's equation. This could relieve students of keying in the commands, prevent them from making syntax errors, and moreover, it could provide a way to structure the output. Such a calculator interface could be a straightforward extension of the palettes already provided in *Mathematica*, such as the expression-input palette shown in Fig. 1.

Improvements to the educational focus for the course module should be aimed at increasing the students' reasoning about the physics background while they are studying worked examples and solving completion assignments. A possible approach to stimulate studying the worked examples would be to pose interpretative questions about their physics content. The completion assignments could be revised in a way that requires more changes to be made by the students. Care should be taken, however, not to raise new programming problems. The assignments should be modified to prevent students from mindless copying. It is intended that students could copy the structure of the example solution, but then several minor adaptations should be required to solve the problem so that students have to work actively with the solution.

This research indicates that with these improvements, a revised course may provide a valuable supplement to practising electrostatics problem solving by hand. Likewise, in other physics domains, such as mechanics, similar courses may help the students to gain an intuitive understanding of the abstract situations and methods with which they are working.

Acknowledgements

The experiment described in this paper was conducted at the Faculty of Physics and Astronomy, University of Utrecht, The Netherlands. The following members of the staff contributed substantially to the successful implementation of the experiment: Cees Alderliesten, Rudy Borkus, Ger Engelbertink, and Jan Kuperus. We also thank Frans Blom (Eindhoven University of Technology, The Netherlands) for commenting on early versions of the testing materials.

References

Chi, M.T.H., Feltovich, P.J. & Glaser, R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science*, **5**, 121–152.

Corbett, A.T. & Anderson, J.R. (1992) LISP intelligent tutoring system: research in skill acqusition. In *Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches* (eds. J.H. Larkin & R.W. Chabay) pp. 73–109. Lawrence Erlbaum, Hillsdale, NJ.

de Jong, T. & Ferguson-Hessler, M.G.M. (1984) Strategiegebruik bij het oplossen van vakinhoudelijke problemen in een semantisch rijk domein. Electriciteit en Magnetisme [Strategy use while solving problems in a semantically rich domain: Electricity and Magnetism]. *Tijdschrift Voor Onderwijsresearch*, **9**, 3–16.

de Jong, T. & Ferguson-Hessler, M.G.M. (1991) Knowledge of problem situations in physics: a comparison of good and poor novice problem solvers. *Learning and Instruction*, **1**, 289–302.

Feiner-Valkier, S. (1997) Achtergronden bij de invoering van een nieuwe methode voor de propedeusevakken Mechanica en Elektriciteit & Magnetisme van de faculteit Technische Natuurkunde in het collegejaar 1996/1997 [*Backgrounds to the introduction of a new method for the first-year course on Mechanics an Electricity and Magnetism in 1996/1997.*] (Internal Report). Technische Universiteit Eindhoven, Eindhoven.

Glaser, R. & Bassok, M. (1989) Learning theory and the study of instruction. *Annual Review of Psychology*, **40**, 631–666.

Gravenberch, F. (1996) Voorlichtingsbrochure HAVO/VWO: Actuele stand van zaken invoering tweede fase [*Information brochure secondary education: Introduction of a new structure for the upper level.*] Instituut voor Leerplanontwikkeling, Enschede.

Griffiths, D.J. (1989) *Introduction to Electrodynamics*. Prentice Hall, London.

Hestenes, D. & Wells, M. (1992) A mechanics baseline test. *The Physics Teacher*, **30**, 159–166.

Larkin, J.H. (1983) Understanding, problem representations, and skill in physics. In *Thinking and Learning Skills: Research and Open Questions*, *Vol. 2* (eds. J.W. Segal, S.F. Chipman & R.Glaser) pp. 141–159. Lawrence Erlbaum, Hillsdale, NJ.

Nathan, M.J., Kintsch, W. & Young, E. (1992) A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, **9**, 329–389.

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

Rutherford, F.J. & Ahlgren, A. (1991) *Science for All Americans*. Oxford University Press, New York.

Savelsbergh, E.R. (1998) Electrowerkboek: technieken voor het berekenen van de potentiaal [*Electro workbook: techniques for computing the potential] (Version 0.1) [Mathematica 3.0 notebook file*]. Eindhoven University of Technology, Eindhoven. Available at: http://www.phys.uu.nl/~savelsbe

Savelsbergh, E.R., de Jong, T. & Ferguson-Hessler, M.G.M. (1998a) Competence-related differences in problem representations. In Learning with multiple representation. (eds. M.W. van Someren, P. Reimannm, H.P.A. Boshuizen & T. de Jong) pp. 263-282. Pergamon Press, Amsterdam.

Savelsbergh, E.R., Ferguson-Hessler, M.G.M. & de Jong, T. (1998b) *Physics Learning with a Computer Algebra System: Towards a Learning Environment That Promotes Enhanced Problem Representations (IST-MEMO-98–01)*. University of Twente, Enschede. (ERIC Document Reproduction Service No. ED421346).

Sweller, J. (1988) Cognitive load during problem solving: Effects on learning. *Cognitive Science*, **12**, 257–285.

Sweller, J. & Cooper, G.A. (1985) The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, **2**, 59–89.

Van Heuvelen, A. (1991) Learning to think like a physicist: a review of research-based instructional strategies. *American Journal of Physics*, **59**, 891–897.

Zhu, X. & Simon, H. (1987) Learning mathematics from examples and by doing. *Cognition and Instruction*, **4**, 137–166.