Preliminary version of: Brinkman, W.-P., Rae, A., & Dwivedi, Y.K.(2007). Web-based implementation of the Personalised System of Instruction: A case study of teaching mathematics in an online learning environment, *International Journal of Web-based Learning and Teaching Technologies*, vol. 2, no. 1, pp. 39-69.

Web-based implementation of the Personalised System of Instruction: A case study of teaching mathematics in an online learning environment.

Willem-Paul Brinkman
Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom. email: willem.brinkman@brunel.ac.uk; tel:+44 (0) 1895 266043; fax: + 44 (0) 1895 251 686

Andrew Rae
Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom. email: andrew.rae@brunel.ac.uk

Yogesh Kumar Dwivedi
School of Business and Economics, Swansea University, Singleton Park, Swansea, SA2 8PP, Wales, United Kingdom. e-mail: ykdwivedi@gmail.com

Preliminary version of: Brinkman, W.-P., Rae, A., & Dwivedi, Y.K.(2007). Web-based implementation of the Personalised System of Instruction: A case study of teaching mathematics in an online learning environment, *International Journal of Web-based Learning and Teaching Technologies*, vol. 2, no. 1, pp. 39-69.

Web-based implementation of the Personalised System of Instruction: A case study of teaching mathematics in an online learning environment.

ABSTRACT

This paper presents a case study of a university discrete mathematics course with over 170 students who had access to an online learning environment (OLE) that included a variety of online tools, such as videos, self-tests, discussion boards, and lecture notes. The course is based on the ideas of the Personalised System of Instruction (PSI) modified to take advantage of an OLE. Students' learning is examined over a period of two years, and compared with that in a more traditionally taught part of the course. To examine students' behaviour, learning strategies, attitudes and performance, both qualitative and quantitative techniques were used as a mixed methodology approach, including in-depth interviews (*N*=9), controlled laboratory observations (*N*=8), surveys (*N*=243), diary studies (*N*=10), classroom observations, recording online usage behaviour, and learning assessments. The paper aims to increase understanding of whether PSI, supported by an OLE, could enhance student appreciation and achievement as findings suggest.

Keywords: CAPSI, Case Study, E-Learning, Keller Plan, Online Learning Environment, Online Test, PSI, Self-Paced Learning, Video-Based Learning, Web-Based Learning, WebCT.

INTRODUCTION

As Online Learning Environments (OLEs), such as WebCT® and Blackboard®, are becoming more widely used, the role of teachers changes as they adapt to their new mode of teaching (Coppola, Hiltz & Rotter, 2002). It remains a challenge however for teachers to use these technologies effectively (Hiltz & Turoff, 2002), and benefit from the suggested advantages of OLEs over traditional classroom learning. These include being more learner-centred, providing flexibility as to the time and the location of learning, being cost-effective for learners, and potentially serving a global audience (Zhang, Zhao, Zhou & Nunamaker, 2004). This paper arises from the experience obtained in delivering a mathematics module for both Computer Science and Information Systems undergraduate students in a UK based university. The first term of the module, which focuses on discrete mathematics, makes extensive use of OLE tools, such as online self-tests, video clips and a discussion board, whereas the second term, which focuses on statistics, is taught by a more traditional lecturer based approach. Comparing the data obtained in the two terms during the academic years 2003-2004 and 2004-2005, gives an insight into how students perceive these OLE tools and into how they affect students' learning strategy and the learning outcomes.

The teaching method used in the first term is based on the principles of the Keller Plan (Keller & Sherman, 1974), also known as the Personalised System of Instruction (PSI). Although these principles were already published in the sixties (Keller, 1968), the observations presented here suggest that they can be highly relevant when teaching is supported by an OLE. The principles of PSI can be summarized as "(1) mastery learning, (2) self-pacing, (3) a stress on the written word, (4) student proctors, and (5) the use of lectures to motivate rather than to supply essential information." (Keller & Sherman, 1974, p. 24). PSI has been applied to courses in various areas such as psychology (Kinsner & Pear, 1988; Pear & Crone-Todd, 2002), physics (Austin & Gilbert, 1973; Green, 1971), mathematics (Abbott & Falstrom, 1975; Brook & Thomson, 1982; Rae, 1993; Watson, 1986), and computer science (Koen, 2005). PSI has received extensive attention in the literature. For example, ten years after its introduction Kulik, Kulik and Cohen (1979) could base their meta-analysis on already 72 different papers, and today PSI is still a topic that receives research attention. In all these years teachers have successfully used PSI, although often they have made some modifications so that it fits into their academic environment (Hereford, 1979). The trend towards high marks has been a recurring observation. The original PSI description talks of a self-paced learning approach where students have to prove mastery of learning material that is divided into small learning units. For each learning unit students receive written material, which includes the learning objective for that unit. Students study the material on their own or in groups, and when they think that they have mastered the unit they take a test. An instructor or a student assisted, called a proctor, immediately marks this test in the presence of the students. If they answer all questions correctly, they receive the written material for the next unit. If they fail, the marker provides them with formative feedback and asks them to study their material again before they re-take the test. Passing the test also gives students the right to attend lectures as a reward. This is possible because no essential material is taught in the lectures; only a few lectures being scheduled and their main purpose is to motivate the students. The use of student proctors clearly has both economic and educational advantages, though care has to be taken to avoid misconduct by proctors.

Proctors are not always used as is shown by Emck and Ferguson-Hessler (1981) who reported that at the Technische Universiteit Eindhoven (The Netherlands) the proctors were replaced by a computer as early as 1970. In a mechanical engineering course, the computer randomly selected a number of questions from a question book. After the students entered their answers onto an answer card, they gave it to the test-room assistant who fed it into the computer. Within less than a minute the test results were printed on a computer terminal. In addition to checking students' answers, computers have also been used to facilitate the testing process. For example, Pear and Crone-Todd (2002) have used the computer to provide proctors, who were automatically selected by the computer, with the completed tests of other students, and afterwards to return the proctors feedback to the students. They relied on human proctors instead of a computer because they used essay style questions. The idea of Computer-Aided PSI (CAPSI) has also been picked up by others (Kinsner & Pear, 1988; Koen, 2005; Roth, 1993; Pelayo-Alvarez, Albert-Ros, Gil-Latorre, & Gutierrez-Sigler, 2000; Pear & Novak, 1996) and this might even become easier to implement if teachers could use an off-the-shelf or standard OLE. For example WebCT®, short for Web Course Tools, is an online course management system in which automatically marked quizzes can be set. It also has other online tools such as a discussion board, a calendar, student homepages, email, chat rooms, online submission of coursework, and a place to store files, such as lecture slides, which students can access online. Providing an effective teaching approach that suits such widely used OLE could therefore be of obvious benefit to teachers.

Although OLEs have been the topic of a number of publications, an OLE in combination with PSI has received limited attention (e.g. Pear, 2003). Research on standard OLEs is, however, advancing. For example, Hoskins and Van Hooff (2005) found a relation between academic achievement and the use of a discussion board. Johnson (2005) also made this observation. Additionally she found that an increased use of the OLE coincided with an increased feeling of peer alienation, i.e. the experience students might have of feeling isolated from other students. She also observed that students who felt more Course and Learning alienation (experiencing the course as irrelevant) made less active use of the OLE and also obtained lower grades. Again this relates to Hoskins and Van Hooff (2005) observation that their achievement oriented students made more active use of the OLE. They therefore worry that an OLE might only engage students that are highly motivated and academically capable. This concern is supported by the findings of Wernet, Olliges, and Delicath (2000) that graduate students valued OLE tools such as a course calendar, hyperlinks, email, and online quizzes more highly than undergraduate students. Other reports (Debela, 2004; Jones & Jones, 2005) on students' attitudes towards e-learning environments are more positive as students believe that these environments can improve their learning, and also that these environments are more convenient and accessible.

Another OLE tool is video. The use of video in a PSI-based course is not new. For example Rae (1993) has used videotapes as a key support to his written course material. On the videos, which students could watch in the university library, he solved exercises and gave short summaries of each learning unit. He found that this approach resulted in high examination marks while he could reduce the tutorial support to an economic level. Instead of using video to present the learning material, Koen (2005) has used it to increase the feeling of being present at the university for students who participate in a distance PSI-based course. He had placed video cameras in the student computer room, the room of the proctor and in the room of the professor. The distance learners on this course could see these live images via their OLE. Unfortunately in his report Koen did not provide results on the effect the streamed video had on the students other than that he was forced to remove the cameras in the student computer room after complaints by a handful of students who did not want to be observed. Of course video has been used in non PSI-based courses as well - for example in the medical field students responded positively on the use of streamed video (Green, et al., 2003; Schultze-Mosgau, Zielinski, & Lochner, 2004), and in a survey on courses in economics offered via the Internet in the USA in the fall of 2000 semester, Coates and Humphreys (2003) found that 18% of the 189 courses used streamed video.

To summarise, although research has been done on PSI, CAPSI and OLE tools there is currently limited understanding about their use and effect on student learning when they are combined in a CAPSI-based course that is supported by a standard OLE. This motivates the present study of a PSI-based module that makes substantial use of a standard OLE. The study looked at students' attitude, their learning strategy and their academic achievement in relation to OLE tools. Before presenting the results of the study, the following section will provide some background information on how the module was set-up; this is followed by a section describing the research approach used. The paper concludes by briefly discussing the main findings and the resulting modifications that have been made to the module.

ADAPTION OF PSI APPROACH TO OLE

The first year module, Foundations of Computing (School of Information Systems, Computing and Mathematics, Brunel University, UK), is taught over two terms. The CAPSI-based first term focuses on discrete mathematics and looks at logic and set theory, whereas the conventionally taught second term focuses on statistics and looks at probabilities, correlations and regression analysis. In the first term the principles of the PSI (Keller 1968; Keller & Sherman, 1974) were implemented in the following way.

- Students used specially written material, divided into four modules, each module including five learning units, each with clearly stated learning objectives. Units one through four were theoretical while units five, for motivational purposes, focused on the application of those theories in computer science.
- Instructors (graduate teaching assistants) in the seminars used specially developed written diagnostic tests for each module to examine, together with the students, their understanding of the material. In contrast to the original PSI principles, these tests were not part of the formal assessment, but students were advised to demonstrate sufficient mastery before they received the written material for the next module.
- The lectures were mainly motivational, covering only the application units, and aimed to give students study advice.

Computer-assisted learning tools and video have been developed to support these traditional elements of PSI. In the OLE students had access to learning tools such as online-self tests, streamed video clips, a discussion board, as well as to the written material, lecture slides and old exams. One hundred and twenty five video clips, mostly less than five minutes in length, were used as the main medium to give essential information (introductions, summaries and solutions to exercises). The videos were simply made and designed to give the students the impression of having the lecturer at their side explaining rather than imparting new information; experience over 25 years has shown such video clips to be remarkably effective with PSI (Rae 1993). For 16 of the learning units (the theoretical units) videos were available in both years that were examined in this case study, but the videos for the other four learning units (the applications units) were available only in the second year, thus providing an opportunity to study their effectiveness by comparing students attitude and learning between the two years. Because students could only access the streamed videos on campus, they could also obtain a DVD for use at home. Although the computer support services charged students £5 duplication cost for the DVD, students were encourage and allowed to copy the DVD freely from each other, or borrow it from the library.

Each learning unit also had a related OLE-based self-test. A test consisted of around five questions, and completion of the test would give students access to the self-test of the following learning unit. These questions were multiple-choice type questions marked automatically by the computer. Another OLE tool that students could use was the online discussion board. Instructors actively encouraged students to post their questions on the online discussion board and told students that they would not answer questions directly emailed to them. Although instructors invited students to answer questions posted on the discussion board, they promised that they would answer questions on a daily basis between Monday and Friday. This approach had advantages for both the students and the instructors. Students could read the questions and answers previously posted, and the instructor only had to answer a question once, instead of the alterative, responding to numerous individual emails regarding the same question.

Each week students had three contact hours, a one-hour lecture delivered by the lecturer in a large lecture theatre, a one-hour seminar in which they could take the written diagnostic tests or get help from the teaching assistant (TA), and a one-hour lab in which they could watch the videos, take the online self-test, and work on their coursework. Eight TAs ran the seminars and lab sessions, each TA being responsible for a group of between 18 and 25 students. To reduce the often-mentioned problem of procrastination in PSI (Hereford, 1979), the students had to take three formal assessments in the first term. They consisted of two pieces of coursework (project 1 and 2), which students could do in pairs, and a one-hour mid term test on module 1 and 2. Furthermore, students had to demonstrate mastery of the first term material as part of a three-hour final exam at the end of the year. The element of pair work was deliberately introduced to counteract concerns that PSI might mitigate against the social interaction between students in which they benefit from exploring conceptual problems with peers (Sheehan, 1978).

The second, more traditionally taught term centred on a weekly two-hour lecture that covered all material. In the one-hour seminar, students worked again in groups of 18 to 25 students, under the supervision of a TA on so called problem sheets, answers to which students could also find on the OLE at the end of each week. The OLE also provided the usual tools including a discussion board and weekly lecture slides. TAs also ran the one-hour lab session, in which students worked on exercises with the statistical software application SPSS® and Excel®, or could take computerised self-tests developed in Mathletics™ (Kyle, 1999). This environment was also used as part of the formal assessment of the second term, students taking two tests that the computer marked automatically. Students also had to submit a statistical report as a piece of coursework, and finally to take the three-hour final exam at the end of the year.

RESEARCH APPROACH AND INSTRUMENTS

Biggs' (2003) 3P (Presage, Process and Product) model of teaching and learning was used as a starting point to structure the research into categories of factors that could influence learning. The model follows the chain in learning. It starts with the factors before the learning takes place, which are split into ***student factors***, such as prior knowledge and interests, and ***teaching context***, such as assessment procedures, teaching sessions, and computer-assisted learning tools. These factors influence students' ***learning activities*****,** or in other words their approach to learning. The 3P model sees this engagement as an essential factor that eventually determine the ***learning outcomes,*** the new skills and knowledge that students master**.** To study these factors, data was collected in a variety of activities, including online student surveys, interviews, diary studies, observations in the class, a controlled usability test, tracking online behaviour, and assessment results. The online student surveys (appendix Table 8 and Table 9) were OLE delivered and students completed them anonymously at the end of each term. A total of 243 responses were collected, 85 in the first term survey of 2004 and 60 in same survey in 2005; 54 responses being collected in the second term survey of 2004, and 44 again in 2005. Please note that from now on the academic year 2003-2004 will be referred to simply as 2004 and 2004-2005 as 2005.

An important part of the research was to see whether CAPSI would change students' learning activities and strategies. A good teaching context should encourage deep learning, where students focus on underlying meaning and principles; the learning environment should move students away from surface learning (Biggs, 2003; Ramsden, 2003), where students act with the intention of passing the module with the minimal amount of effort or engagement. In a previous study on this module, Hambleton, Foster and Richardson (1998) had found that PSI could have a positive affect on students' learning. They had asked students to complete the Approaches to Studying Inventory (ASI) (Ramsden & Entwistle, 1981), an earlier instrument to measure approach to learning. After analysing this data they found that students scored significantly higher on the 'comprehension' learning scale for the PSI-based module than for a more traditionally lecture-based module on statistics. They therefore concluded that PSI could have a positive impact on students learning strategy. To follow up on this research the R-SPQ-2F inventory (Biggs, Kember, & Leung, 2001) was included in the end of term surveys, with the exception of the first term survey of 2004. The inventory is a 20-item questionnaire that scores students both on a deep approach and surface approach scale. These two scales are derived from two subscales, which the inventory provides for each approach: the students' motivation and their strategy.

While the survey data provides a general insight from a sizable sample, between 25% and 48% of the population, the semi-structured interviews conducted with nine students in the summer break of 2004 provided an in-depth understanding of students' learning strategies and attitudes. The students had all passed the module in 2004 and agreed to complete the R-SPQ-2F inventory and to be interviewed by a PhD student on the premise that he would not disclose their names. These students were interviewed for half an hour on the phone, and focused on the student approach to learning, the teaching approach, student characteristics, and the OLE tools (appendix Table 10). For their participation in the interview students received a £5 incentive, paid as all other incentives in this study from a university research grant that supported this research.

Another source of data was the diary study. Where surveys and interviews provided information from only one point in time, in the diary study 10 students agreed to provide weekly information throughout the course of 2005. Collecting data in the diary study however proved to be more difficult. Although initially 10 students started with weekly reports, only three students continued to do this into the second term. Still 105 weekly reports were collected. For their participation students received a £10 incentive for each term and they were promised that their names would not be disclosed to the instructors. The diary study was conducted by a PhD-student who also made weekly observations in the lab. The number of students that attended the lab and their activities were recorded in the logbook. Another PhD student was asked in the summer of 2004 to conduct a usability study of the OLE of the module. Eight masters students who had not attended the module were invited into a usability laboratory and asked to study the first learning unit of module one and to take the related online self-test. During the test, she watched the students from an observation room through a one-way mirror, and recorded any comments made by the students. Afterwards, she interviewed the students and asked them to highlight specific problems they had encountered. The ease of use and the students' satisfaction in using specific OLE tools were also examined in the usability test with a component-specific usability questionnaire (Brinkman, Haakma, & Bouwhuis, 2005) (appendix Table 11). This questionnaire included six ease-of-use questions from the Perceived Usefulness and Ease-of-Use (PUEU) questionnaire (Davis, 1989) and two questions from the Post-Study System Usability Questionnaire (PSSUQ) (Lewis, 1995). The test took around two hours and students were given a £12 incentive for their participation. Student behaviour was also followed by recording web traffic. In 2005 four weeks into the module, web access of several pages was tracked, including the home page, the self-tests, and the main page for the video clips. All these data collection activities were set up to get both a broad and in-depth understanding of how students perceived the module and how they actually engaged with it. Finally, the results of the coursework and the exam gave an insight into the students' academic performance. Because data was collected both in the CAPSI taught first term and in the traditionally taught second term, it was possible to examine the effect these different teaching contexts had on learning activities and outcomes.

FINDINGS

Student Factors

In 2004 176 students (73% male, 27% female) were registered for the module and in 2005 177 (79% male, 21% female). The students had various educational backgrounds. For example, in the four end-of-term surveys conducted in 2004 and 2005, 50% of the students stated they had an A-level, short for Advanced Level, a non-compulsory qualification taken by students in England, Wales, and Northern Ireland which is the usual university entrance qualification. Students usually take A-levels in the final two years of secondary education, after they have obtained a General Certificate of Secondary Education (GCSE), which is taken at an age of around 16. On the other hand, 17% of the students had a certificate or diploma of the Business and Technician Education Council (BTEC), 10% of the students had a General National Vocational Qualification (GNVQ), 14% of the students had done an Access course, and 9% of the students had overseas qualifications. But a basic requirement for the degree course was that students at least had a GCSE maths level or equivalent for example a graduation diploma from a good US high school. For some students however, it had been a while since they obtained their qualification. For example, one student in the in-depth interview mentioned: "…but things such as tables, and probability I did in GCSE. However, I have forgotten it." Some students were therefore invited to attend workshops to refresh their knowledge on basic mathematical concepts relevant for this module, such as linear equations, linear function, power, and logarithms. In 2005 this invitation was based on the results of an online OLE test that students took in the first lab. In the test, the students scored on average 24.96 (SD = 10.91) points out of a 40 points maximum. Students that obtained a score below 20 were advised by email to attend the workshops, which was also open to other students to attend.

Some of the students did not live on campus, and therefore spent time commuting. Attending class or doing group work was more difficult for some of these students. For example one student mentioned in the interview:

> I was off campus and I had to travel by bus, which took around one and a half hours each day. This was one reason why I was not attending the lectures, coming to university means wasting three hours in travelling and in that much time I can do some work. It did stop me from going to lectures….

Almost all students in the interview mentioned that access to a computer was vital for passing the module. Computer access outside the lab sessions seemed adequate. In the four surveys, only two responses indicated that they never had access to a PC or laptop outside the lab session, 21 responses indicated to have access sometimes, but the majority, 219 responses replied that they had regular access. Although this seems a reassuringly high number, the percentage of students that have never had access might be higher since the data from the online questionnaires could be biased. On the other hand, students that lived on campus had access to computer rooms, where they could work outside the normal class hours. Most students who were interviewed that lived off campus also stated that they had a computer at home, except for one student, who stated: "I did not have a computer at home, so I had to come to university everyday to study. So it was a waste of my time travelling." This is clearly a disadvantage. Where students in conventionally taught courses cannot study anywhere and anytime because they have to attend lectures, computer-assisted learning makes students dependant on computer access, which unfortunately goes against the idea of anywhere, anytime learning. The time students spent on learning, was also limited as students were engaged in other activities as well. For example one student mentioned: "I cannot spend enough time on each module, as I did not have free time left after doing my job. Also after doing my job, I would get tired and would not feel like studying."

Teaching Context

The difference in students' appreciation of the teaching context —things such as assessments, the OLE tools, and lectures— between the CAPSI-based first term and the conventionally taught second term shows that students on average were more positive about the first than about the second term teaching context. Table 1 shows the average scores obtained on the end of term survey items that related to the teaching context. The scores of these six items were used in a MANOVA, which showed a significant main effect for the terms ($F(df_{between\text{-}groups} = 6, df_{within\text{-}groups} = 204) = 26.91, p < .001$). ANOVAs on the individual items revealed this effect also in the scores of the overall module quality ($F(1, 209) = 18.00, p < .001$), the usefulness of lectures ($F(1, 209) = 70.92, p < 0.001$), the usefulness of seminars ($F(1, 209) = 11.00, p < .01$), and the usability of the OLE ($F(1, 209) = 55.28, p < 0.001$). Students rated all these items higher for the CAPSI-based term than for the conventionally taught term. Items only related to CAPSI, also received high average ratings, such as the usefulness of printed material ($M = 2.98$, $SD = 0.80$), the online self-tests ($M = 3.32$, $SD = 0.70$), and the video clips ($M = 3.07$, $SD = 0.86$). In the interviews, students were also very positive about these online tools. They liked the online-self tests to test and extend their knowledge and they suggested including more random and more difficult questions in the tests, because as one student put it "I do not like to do the same question ten times." The students found the discussion board also useful, especially those students that lived off campus. Some students post messages, whereas others (sometimes called 'lurkers') just checked it frequently to keep up to date, looked at the type of questions that were posted, or read previously posted answers to questions they also had. In the interview students were also positive about the video clips. They liked these because they helped them revise before exams, or when they had missed a lecture. One student frequently watched the video clips because he/she did not attend the seminars and tried to understand and solve the problems by watching the videos. Students also used the videos if they ran into problems as the following diary entry shows: "The difficulties [*sic*] encountered this week was rules of inference, however I resolved by watching some exercise videos…". On the other hand, one student who lived off campus mentioned that he/she had not purchased the videos on DVD because the five pounds charge for the disc was too expensive. Many students in the interview also mentioned the desire to have online access to the video off campus. Apparently students did not regard watching videos as a normal lab session activity, which became clear from their reluctance to bringing headphones (these were necessary because the computers in the lab were not equipped with speakers). However, the policy of bringing in your own headphone did not seem to work, as illustrated by the following remark in the observation logbook for the second lab of the first term: "Student suppose [*sic*] to bring headphones so they can listen Video but none of them brought [*sic*], instead they were trying to read the videos." In fact, during all his observations, the observer never noticed a student bringing in a headphone.

Table 1: Mean rating (*SD*) of teaching context of the first term (*N* between 134 and 145) and second term (*N* between 91 and 98) in the 2004 and 2005 surveys.

Item	Scale	Term 1	Term 2
Overall module quality	1 (poor) – 4 (very good)	2.98(0.78)	2.55(0.81)
Usefulness lectures	1 (useless) – 4 (very useful)	3.35(0.75)	2.44(0.83)
Usefulness seminars	1 (useless) – 4 (very useful)	2.85(0.89)	2.46(1.03)
Usefulness lab sessions	1 (useless) – 4 (very useful)	2.53(0.84)	2.45(0.90)
Usefulness discussion board	1 (useless) – 4 (very useful)	2.72(0.72)	2.82(0.90)
Usability OLE	1 (very low) – 5 (very high)	4.13(0.90)	3.12(0.87)

During the course of 2004 some students informally mentioned some concerns about the usability of OLE tools. A student in the interview mentioned: "…at the beginning of the year students should be given a demo of WebCT, how to use it and the things available on it. Many students did not use it because they do not know how to use it." This was therefore investigated further. A usability test was conducted in the summer break of 2004. Overall the students in this test were positive about the set-up of the OLE tools. However, their major usability concern related to the difficulty of finding and navigating to particular items in the site. This is not uncommon for OLE delivered courses (Engelbrecht & Harding, 2001). Therefore, after the test, the navigation panel and the overall structure of the site was redesigned in an attempt to make it more consistent. In the usability test participants were also asked to rate the ease-of-use and their satisfaction of a number of OLE tools. The results are given in Table 2. Since students rated items on seven point Likert scales, the rating of above 3.5 seems to indicate no serious usability problems with the six items.

Table 2: Mean (*SD*) ease-of-use and satisfaction rating of the OLE tools by eight students in the usability test.

Item	Ease of Use	Satisfaction
Study guide	5.48(1.09)	4.63(1.41)
Video	5.69(0.94)	5.31(1.36)
Lecture slides	5.62(0.69)	5.43(0.79)
Online self-test	5.48(0.72)	5.25(1.07)
Discussion board	5.73(0.79)	5.06(1.45)
OLE progress overview	5.40(0.96)	5.38(0.69)

Learning Activities

The first step in examining the student learning activities was analysing the scores on the R-SPQ-2F questionnaire that was included in the end of term survey for the first and the second terms in 2005. An ANOVA was conducted to see the effect of the learning context on the two learning approaches scales. The ANOVA with repeated measure took as between-subject variable the terms and as within-subject variable the scores on the deep approach and surface approach scales. The analysis reveals a significant effect ($F(1, 98) = 46.41$, $p < .001$) for scores between the two scales. Examining the means showed that the students scored, on a scale from 10 to 50, 28.8 ($SD = 0.73$) points on the deep approach scale and 23.2 ($SD = 0.58$) on the surface approach scale. However, the analysis failed to find a significant effect ($F(1, 98) = 1.42$, $p > .05$) for the terms, or for a two-way interaction ($F(1, 98) = 0.66$, $p > .05$) between approaches and terms. ANOVAs on the sub-scales, motivation and strategy, resulted in similar outcomes. This means that survey responders were on average more inclined to apply a deep instead of a surface approach through both terms. Thus the difference of the overall learning environment between the first and the second term did not seem to change the students learning approach.

The next step in the analysis was to look for possible relationships between the learning approach and elements of the teaching context. Table 3 and Table 4 show the Pearson correlations between the subscales of the learning approaches and the elements. Both the surface motivation and strategy seem to have a negative Pearson correlation with students' interest and perceived lack of difficulty of the first term subject matter. For the second term the only significant Pearson correlations between teaching context items and surface learning was a positive Pearson correlation between perceived usefulness of the seminars and motivation; and a negative Pearson correlation between perceived usefulness of the lectures and learning strategy. The relative small number of significant Pearson correlations in Table 3 may suggest that the surface learning approach was less driven by the teaching context, or perhaps that these relationships are not linear as a Pearson correlation assumes.

Table 3: Pearson correlation between surface learning approach and items of the teaching context of first term (*N* ranges from 58 down to 52) and second term (*N* ranges from 43 down to 39) in 2005.

Teaching context	Motivation		Strategy	
	Term 1	Term 2	Term 1	Term 2
Overall module quality	0.16	-0.21	0.08	-0.23
Lack of difficulty of subject matter	-0.26*	-0.26	-0.46**	-0.18
Interest in subject matter	-0.28*	0.00	-0.37**	-0.17
Previous familiarity subject matter	-0.23	0.22	-0.11	0.20
Usefulness lectures	-0.03	-0.17	0.14	-0.30*
Usefulness seminars	-0.21	0.34*	-0.25	0.28
Usefulness lab sessions	0.10	0.28	0.09	0.06
Attendance lectures	-0.04	-0.10	-0.09	-0.28
Attendance seminars	-0.10	-0.07	-0.08	-0.29
Attendance lab sessions	-0.19	0.09	-0.21	-0.21
Usability OLE	0.07	0.07	-0.06	-0.01
Usefulness discussion board	-0.01	0.10	0.10	0.08

**p*. < 0.05. ***p*.< 0.01.

The high number of significant Pearson correlations between the deep learning scale and teaching context items suggests the opposite for deep learning (Table 4). Both for the first and second terms, there are positive Pearson correlations between the deep approach on one side and on the other side: perceived quality of the module, perceived lack of difficulty and students' interest in the subject matter, previous familiarity with the subject matter, perceived usefulness of the lectures and seminars, and attendance at lectures. However, there are also some distinctions between first and second term. For example, the deep learning approach for the second term seems associated with attendance at lectures, seminars and lab sessions, whereas for the first term the deep approach is only associated with the attendance of the lectures. Therefore it seems that in the conventionally taught second term, classes were mainly attended by students that were intrinsically motivated and applied a deep learning strategy, while for the CAPSI-based term this factor was of less importance to explain class attendance. Table 4 also shows that the usability of the OLE and usefulness of the discussion board positively correlated with intrinsic motivations. This also seems to explain the perceived usefulness of the written study material, the online self-tests, the introduction, and Question & Answer (Q&A) videos that were provided in first term (Table 5). Apparently deep learners appreciate these online tools. However because no difference was found in the learning approaches between the terms, it is unlikely that these tools have a large impact on students' adoption of a learning approach.

Table 4: Pearson correlations between the deep learning approach and items of the teaching context of first (*N* between 52 and 58) and second term (*N* between 39 and 43) in 2005.

Teaching context	Motivation		Strategy	
	Term 1	Term 2	Term 1	Term 2
Overall module quality	0.39**	0.38*	0.30*	0.40**
Lack of difficulty of subject matter	0.33*	0.33*	0.21	0.43**
Interest in subject matter	0.53**	0.23	0.33*	0.33*
Previous familiarity subject matter	0.46**	0.29	0.31*	0.38*
Usefulness lectures	0.27*	0.42**	0.19	0.35*
Usefulness seminars	0.35**	0.37*	0.22	0.25
Usefulness lab sessions	0.04	0.21	0.08	0.26
Attendance lectures	0.31*	0.29	0.34*	0.41**
Attendance seminars	0.01	0.36*	0.23	0.41**
Attendance lab sessions	0.17	0.41**	0.23	0.42**
Usability OLE	0.29*	0.16	0.15	0.18
Usefulness discussion board	0.36**	0.37*	0.19	0.25

*$p < .05$. **$p < .01$.

Table 5: Correlation between learning approach and items of the teaching context of first term in 2005 (*N* between 48 and 58).

Teaching context	Surface approach		Deep approach	
	Motivation	Strategy	Motivation	Strategy
Usefulness written study material[a]	-0.26	-0.15	0.26*	0.28*
Usefulness online self-tests[a]	-0.26	-0.09	0.35**	0.18
Usefulness Q&A video[a]	-0.05	-0.18	0.28*	0.21
Usefulness introduction video[a]	0.15	-0.18	0.28*	0.23
Usefulness summary video[a]	0.10	0.14	0.26	0.13
Number of introduction video watched[b]	0.04	0.04	0.10	0.22
Number of exercise video watched[b]	0.01	-0.15	0.11	0.01
Number of summary video watched[b]	0.02	-0.05	0.12	0.09

[a]Pearson correlation, [b]Spearman correlation, *$p < .05$. **$p < .01$.

The ten students in the diary studies each week spent on average 9.3 (SD = 4.85) hours on the module. This was spread over lectures (M = 1.7 hours, SD = 0.78), seminars (M = 1.0 hours, SD = 0.60), labs (M = 1.3, SD = 1.26) and a considerable amount of self-study (M = 5.2, SD = 4.23). Applying an ANOVA on the survey data from the two years gave more insight into the impact of the learning context on class attendance. The analysis was conducted on the student's average percentage of the lectures', seminars', and lab sessions' attendance combined. An ANOVA with as independent variable years and terms revealed a significant main effect for terms ($F(1, 239) = 37.38$, $p < .001$) and for years ($F(1, 239) = 19.21$, $p < .001$), and also a significant two-way interaction effect ($F(1, 239) = 9.64$, $p < .01$) between terms and years. Examining the means of the first term shows that attendance remained stable over the two years. In the first term of 2004 it was 79.9% (SD = 16.2), and 83.2% (SD = 14.2) in 2005. Attendance at the second term classes was lower, but increased. In 2004, attendance was 56.4% (SD = 25.2) and in 2005, it was 75.5% (SD = 22.5). This agrees with students' attendance observed in the lab in 2005. In the first term this was on average 70 students, significantly ($F(1, 16) = 12.28$, $p < .01$) more than the 45 students on average in the second term. The difference between terms could suggest that CAPSI may change study behaviour. Of course there could also be other factors, such as previous knowledge. For example in the interview a student mentioned:

> Actually the first semester math was new for me so I attended regularly all lectures in the first semester. However most of the statistical stuff in the second semester was not new for me, therefore I only attended a few lectures in the second semester.

Another factor might have been the aim and style of the lectures. In the CAPSI-based term, lectures were for motivation, while in the other term, lectures were for covering the subject matter. Because of this, lectures in the first term were shorter, around one hour, while the lectures in second term were longer, around two hours. This point was illustrated by another comment made in the interview:

> The first semester lecture was extremely useful and interesting. However the second semester lecture was too lengthy, for me it was too much to attend a one and a half hour lecture without any break. After an hour I get tired, I lost my concentration and interest. Also in the statistics lecture there were less interaction between the students and the lecturer, students only ask questions at the end of the lecture not during the lecture. However it would be more useful to have questions and answer during the lecture.

Still the difference in attendance could also relate to students' tendency to initially follow all classes, and later on to stay away more often when they have become more familiar with university life and the course. A similar reduction was also found when examining the number of messages posted on the discussion board. In 2004 staff and students together posted 330 messages, whereas in 2005 this number had risen to 763. The promotion of the discussion board in 2005 seems therefore to have been successful. However, comparing the various periods in the year, the percentages of messages were rather similar. For example, in 2004, 49% of the messages were posted in the first term, whereas 47% in 2005. Next, the percentage for the second term was 37% for 2004 and 33% for 2005 (the remaining messages were posted in the revision and exam period). As with class attendance, students were more active on the discussion board in the first term than in the second term. The discussion board, however, could not replace face-to-face meetings entirely. For example, one student wrote down in his/her diary: "…even though there is webct's discussion board it is not the same as being explained a difficult problem over s [*sic*] few notes on the internet" referring here to the alternative of discussing it with the instructor in the seminar.

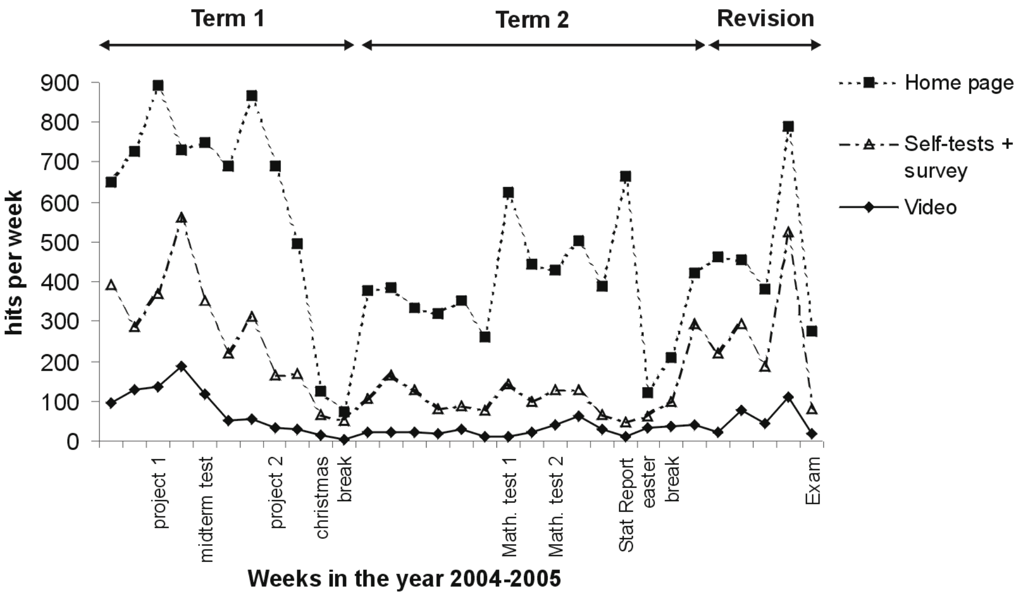

Fig. 1: OLE usage of module between week 4 and 34 in the year 2004-2005, with on the x-axis the assessments events (project 1, midterm test, project 2, mathletics test 1, mathletics test 2, statistical report, exam).

Students' behaviour was also affected by assessment deadlines, or as one student in the interview said: "I work harder when I have exams and not as hard when I do not have exams." This behaviour was also observed in web traffic. Fig. 1 shows the recorded accessing of the home page, the self-tests, and the main page of the online videos. The peaks of the home page hits clearly relate with the assessments events. The use of videos, and particular the use of online self-tests had peaks especially before the midterm test and for the final exam. During the Christmas and Easter break, online activity dropped and picked up again when classes resumed. The holidays were also quiet periods on the online discussion boards. Students in the interview were positive about the spread of coursework throughout the year, instead of only a single final exam. Even a student who applied mainly a surface approach said in the interview:

> I like very much the idea of having both coursework and exams. If I do not have the coursework then I would have left everything to study at the end. Because of the coursework, which was spread in several tasks, I studied this module constantly and therefore I learnt it better than other modules with only an exam at the end of the year.

The frequency of students' attempts at the online self-tests was relatively high for the first modules of term one, but it steadily declined, till halfway through the third module only around 25% of the students were still attempting the tests (Fig. 2). The initial access policy of having to pass previous online-self tests might have created this decline. Not all students liked this mastery policy. One student stated in his/her diary: "I find it a little overwhelming that I must get 100% [*sic*] in a lab test to regard that particular test as a pass!!!" Although it was never a 100% mastery policy, this was changed early on in the year. Instead of mastery, students should at least have attempted the previous self-tests. In the revision period, even this condition was dropped, to give students unconditional access to prepare for the exam. At the same time, students got access to a general self-test with 20 random questions taken from all tests. From Fig. 2 this seems to have attracted more attention from the students. Around a third of the students in 2005 took this test. Possibly students attempted the self-tests for module one and two as preparation for the midterm test. After that there was no immediate assessment deadline to motivate most students into taking the other tests, and their attention might also been drawn to coursework deadlines of other modules as the following diary entry clearly shows: "For weeks 7, 8 and 9 I haven't really been attending lectures, seminars or labs. Reasons for this are that I have been too busy with work and other modules." With the end of year exam approaching, students might have ignored the remaining unfinished self-tests and went immediately for the general test. Others might have completely ignored or forgotten the self-tests and used other means to revise for exams or focussed all their attention on the second term material.

Fig. 2: Percentage of students that attempted the online-self tests in 2005 (M stands for Module, and U for unit).

Learning Outcomes

Several reports can be found in the literature about the positive impact PSI has on students' marks. Kulik, et al. (1979) concluded in their meta-analysis of 75 comparative studies that students score on average 8% higher in a PSI course than in a conventionally taught course. The exam results of the module agree with these findings. Fig. 3 shows the exam results of the CAPSI taught part and the more conventionally taught part of the module. In 2004 the average difference between the two is 14 points on a scale of 100, and in 2005 it is 20 points. An ANOVA with repeated measure confirms this observation. The ANOVA with as dependent variables the exam marks related to the two terms of the module, and as independent variable the year of the exam, revealed a significant main effect for the two terms of the module ($F(1, 332) = 424.09$, $p < .001$). One reason for this might have been the differences in the teaching approach, or simply that one topic was easier than the other. The analysis also showed a significant difference between the years ($F(1, 332) = 13.09$, $p < .001$). On average students scored lower in the exam in 2005 than in the previous year. This could mean that the performance of the student cohort was different or the overall exam was more difficult. Still, the more interesting results of the analysis was a significant ($F(1,332) = 13.69$, $p < .001$) two-way interaction effect between the year and the topic of the questions. Where the exam scores for the CAPSI taught part remain more or less stable over 2004 ($M = 52.1$, $SD = 19.0$) and 2005 ($M = 48.9$, $SD = 19.2$), the exam scores of the other part drops from a 38.0 points ($SD = 16.9$) average to a 28.7 points ($SD = 15.0$) average. This could be attributed to several factors, such as a variation between the two years in the students' academic abilities and the ability of the teaching context to adapt to this; a temporary change of the lecturer in the first five weeks in the second term of 2005; or a combination of these factors.

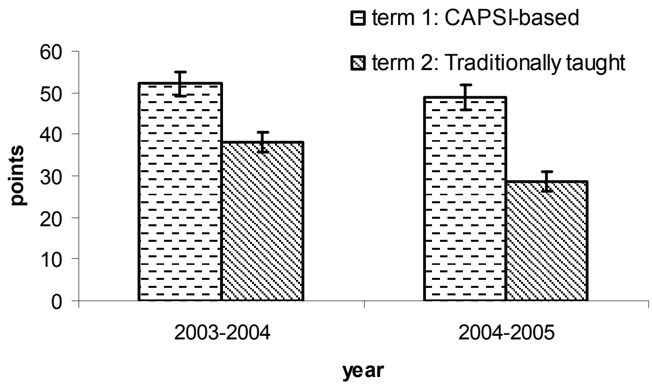

Fig. 3: Mean exam score, with a 95% confidence bar, on first term topic and second term topic obtained by students in academic year 2003-2004 and 2004-2005.

The impact of video clips and online self-tests was apparent in the midterm test of the first term. In this test students had to answer four out of six questions under exam conditions. Two of the questions related to the application units, which were taught only in the lectures and were supported by OLE tests and video material only in 2005. While in 2004, 71% of the students attempted one or both of these application related questions, in 2005 this had risen to 80%. Although it just fails to reach a significant level in a 2 × 2 Chi Square analysis, χ^2 (1, *N* = 345) = 3.81, *p* = .051, the trend shows that these OLE tools could have an impact on students confidence in attempting questions. Still, giving them without testing the material in the written diagnostic test does not seem to improve the exam results. An ANOVA with repeated measures on the average score obtained in the 2004 and 2005 exams on the four questions covering the first four theoretical learning units of the first term, and the score on a question covering the application units, units five, only revealed a significant main effect ($F(1, 332) = 74.38$, $p < .001$) for the topic of the questions. Students scored on average 10.5 (*SD* = 3.9) out of 20 points on the units one till four related questions, and 8.7 (*SD* = 4.9) out of 20 points on the units five question. But more importantly, no significant two-way interaction effect ($F(1, 332) = 0.87$, $p > .05$) was found between questions and exam year. This means that the introduction of the video clips and OLE self-tests in 2005 did not seem to improve students' performance in the end of the year examination.

The focus on the students learning strategy seems justified. In the end of term surveys students were asked to indicate their grades obtained for the midterm test and statistics report coursework. As mentioned in a previous section, the surveys also included the R-SPQ-2F (Biggs, et al., 2001) learning approach inventory. Table 6 shows Spearman correlations between the coursework grades and the learning approach. In the first term, based on the CAPSI principles, grades negatively correlated with a surface approach. Students that were motivated mainly by a fear of failure and applied a narrow target, rote learning strategy had a tendency to obtain lower grades in the midterm test. However no significant Spearman correlation was found between the midterm grade and the deep learning approach. Table 6 shows the precise opposite pattern when it comes to the statistical report coursework grade and learning approach. In the more conventionally taught term, the grade positively correlated with the deep approach. Therefore it seems that the CAPSI learning environment was less supportive when students applied a surface approach, in other words, when they picked items and rote learnt them, from a fear of failing, students were less likely to obtain high grades with CAPSI. However, success in a CAPSI course did not seem to rely on students' deep learning strategy. This might be explained by the behaviouristic tradition in which PSI was developed, placing an emphasis on the environment rewarding good behaviour (Keller & Sherman, 1974) and less on student comprehension and cognition. In the second term, students seem to get higher grades because of their deep learning strategy. In short, whereas in the CAPSI environment, the deep learning strategy is less important, in the traditional taught environment this seems to be a prerequisite for academic achievement.

Table 6: Spearman's correlation between learning strategy and coursework grades.

Scale	Midterm test (term 1)[a]		Statistical Report (term 2)[b]	
Surface approach	-0.28	*	-0.07	
Motive	-0.32	*	-0.15	
Strategy	-0.28	*	-0.02	
Deep approach	0.10		0.32	**
Motive	0.26		0.28	**
Strategy	-0.01		0.33	**

[a] obtained in end of the first term survey of 2005, with *N* between 51 and 53; [b] obtained in the end of the second term survey of 2004 and of 2005, with *N* between 88 and 91. **p* < .05. ***p* < .01.

Table 7 shows the Pearson correlations between OLE use and the exam marks for the first and second term. For both terms the access to the home page of the OLE site, the use of the self-tests and the discussion board correlated significantly with academic achievement. However, the size of the Pearson correlations is relatively small, and interestingly the use of the first term online self-test correlated significantly with the exam marks of the second term. Although the size of this Pearson correlation is smaller than the Pearson correlation between the exam marks of first term and use of the self-tests, it suggests that OLE use is perhaps mainly an indicator of students' overall motivation and only for a small portion a direct predictor of academic achievement.

Table 7: Pearson correlation between exam marks related to first and second term and OLE use (*N* = 334).

OLE use	Term 1	Term 2
Hits on the home page of the OLE site	0.23**	0.12*
Number of messages read on discussion board	0.32**	0.24**
Number of messages posted on the discussion board	0.20**	0.21**
Percentage of first term online self-tests attempted	0.34**	0.22**

p < .05. **p* < .01.

DISCUSSION

This study started from the question as to how CAPSI might affect student attitude, learning and performance. From the findings it seems that students were more positive about the CAPSI-based term than the traditionally taught term. They also found facilities used in the CAPSI term such as the printed material, video, online self-test and the discussion board, useful. Although the analysis of the survey data did not reveal that CAPSI changed the students' learning approach, class attendance and the use of the discussion board were higher in the CAPSI-based first term than the second term. Whereas lectures, seminars and lab attendance in the lecture-based term correlated with a deep learning approach, in the first (CAPSI) term this correlated only with attendance at the motivational lectures. In short, the CAPSI learning environment seems to engage students but not to change their motivations or learning approach noticeably. This is promising because in the traditionally taught term, engagement seems to be a function of the student's deep learning approach, which coincided with high academic achievement. In the CAPSI term this link was less strong; whether or not students applied a deep learning approach did not determine their academic achievement, though students that applied a surface approach in the CAPSI-based term obtained lower marks. In other words, for students to be successful in the CAPSI-based term, deep learning approach seems less important, however, applying a surface learning approach seems less effective. The correlations found in the first term between the deep learning approach and perceived usefulness of OLE tools (Table 5) seem to confirm the earlier report by Hoskins and Van Hoof (2005). The results support their concerns that the OLE might only be taken advantage of by highly motivated students. However, the Pearson correlations between OLE use and exam marks (Table 7), although significant, were relatively small. Therefore there seems little support to justify fears that an OLE would be an obstacle to student learning. Instead it seems that the teaching principles, rather than the use of OLE tools, are the determining factor. Marks for the CAPSI taught term were higher than those for the traditionally taught term. Therefore, the main lesson learned from the study appear to be that the principles of PSI can effectively be used in an online learning environment, creating higher marks and students' appreciation.

The findings related to the videos and the OLE self-tests suggest that students perceived them as useful, and students might feel more comfortable in taking an assessment on a topic discussed in the videos. However, these OLE tools alone do not have a large impact on student performance. It seems more likely that they are more successful in combination with diagnostic tests; bringing the student from only passively watching the videos to actively studying them to pass the diagnostic test. This agrees with the overall observation that students' learning activities were largely driven by assessment deadlines. The data on the online self-tests suggests that OLE tools were used more if students could clearly link them with the next upcoming assessment exercise. Kraemer (2003) also notices this effect. Completion rates of OLE material in her course were higher when librarian students were required to go through them for a grade. It seems that the call for aligning teaching, assessment, and learning objectives (Biggs, 2003) should include online teaching tools as well. Students should see the benefit in using these OLE tools for passing their assessments; otherwise, only the highly motivated students might use them.

Limitations of the Study

As in any empirical study, this study also has limitations. The main limitation is that the results are restricted to those students who responded to the surveys, interview and diary studies. These students were probably more motivated and performed better than average. For example the average module mark of the participants in the interview was 66.2 (*SD* = 11.6) compared to a class average of 52.5 (*SD* = 17.1). The results of the surveys could also have a positive bias towards OLE tools as the surveys were OLE delivered. Students that did not like or have problems using the OLE tools, might therefore be underrepresented in the survey response. Another limitation is the ability to control variables. For example, for obvious pedagogical and ethical reasons individual students could not be withheld or provided OLE access to study the effect these OLE tools had on student learning and performance. Furthermore, the material taught in the CAPSI-based term and traditionally taught term could be a confounding variable as students could be more familiar or interested in a particular topic which might have influenced their learning activities. However, the study applied a mixed methodology approach, grounding the analysis on data from various data sources. This has the advantage that limitations of one data source could be overcome by using another data source in some cases. Next, this is a case study about teaching mathematics to first year Computer Science and Information System students on a UK based university. Although this is very specific, it could be of wide interest as mathematics is taught in many courses in different fields including science, engineering, business studies and economics.

Further Development

The findings of the study have already led to considerable changes in the module. First of all, the second term has also become CAPSI-based, with the written material broken up into ten units, each with its own written diagnostic test, online self-test, and supporting videos. The videos again consisted of an introduction and a summary clip for each unit and a number of Q&A videos. Developing the videos took around one and half week preparation, a week shooting them in a recording studio, and one and half weeks editing and converting to MPEG format. The 57 clips, including a general introduction video, are again available online on campus and on a separate DVD. Eighty headphones have been bought and made available to students to borrow in the lab, since it was observed that students did not bring their own headphones. The question database for the first term has also been extended. Instead of the standard four or five questions, questions are now randomly drawn from a question database that has 20-30 questions for each unit. As others have reported (Engelbrecht & Harding, 2001), some problems were also experienced with displaying mathematical symbols in the OLE quizzes - unfortunately, some students are still reporting that they are unable to see some symbols on their computer at home. Developing the question database also took considerable time, but it is regarded as a long-term investment. Partly because of pressure to run the module with less staff, but also because of the findings, the question database is now used for supervised tests in the lab as part of the formal assessment. Students have four attempts throughout the year to take or retake these supervised OLE tests, and these replace the midterm test, and the project 2 coursework in the first term, and the two Mathletics tests in second term. For students this has the advantage that the online self-tests are directly aligned with the formal assessment —indeed they are drawn from the same question database. For staff it removes the burden of marking, while students seem to perform equally well on OLE-based and paper-based tests (Poirier & O'Neil, 2000).

Again data from online surveys is being collected and the use of the OLE is being tracked to improve the module and to run it on an economically viable level. The principles of PSI continue to be used in the module not only because of its good track record since its introduction in the sixties, but also because it provides an effective framework for taking advantage of the tools offered by OLE.

Conclusions

The main conclusion of the study is that PSI can be used to enhance both student appreciation and achievement in a course that is supported by OLE. Students' performance was significantly higher for the CASPI taught material than for traditionally taught material. Furthermore, the study found that while a deep learning approach was significantly correlated with good grades on the traditionally taught material, this was not the case for the CAPSI material; the conclusion is that even those students who are not taking a deep approach may in some way be helped to learn mathematics by a CAPSI course.

REFERENCES

Abbott, R.D., & Falstrom, P.M. (1975). Design of a Keller-plan course in elementary statistics. *Psychological reports*, 36(1), 171-174.

Austin, S.M, & Gilbert, K.E. (1973). Student performance in a Keller-Plan course in introductory electricity and magnetism. *American journal of physics*, 41(1), 12-18.

Biggs, J. (2003). *Teaching for quality learning at university: what the student does* (2nd ed). Berkshire, SRHE & Open University Press.

Biggs, J., Kember, D., & Leung, D.Y.P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British journal of educational psychology*, 71(1), 133-149.

Brinkman, W.-P., Haakma, R., & Bouwhuis, D.G. (2005). Empirical usability testing in a component-based environment: improving test efficiency with component-specific usability measures. In R. Bastide, P. Palanque, and J. Roth (Eds.) *Proceedings of EHCI-DSVIS 2004, Lecture Notes Computer Science 3425* (pp. 20-37). Berlin, Springer-Verlag.

Brook, R.J., & Thomson, P.J. (1982). The evolution of a Keller plan service statistics course. *Programmed learning & educational technology*, 19(2), 135-138.

Coates, D., & Humphreys, B.R. (2003). An inventory of learning at a distance in economics. *Social science computer review*, 21(2), 196-207.

Coppola, N.W., Hiltz, S.R., & Rotter, N.G. (2002). Becoming a virtual professor: pedagogical roles and asynchronous learning networks. *Journal of management information systems*, 18(4), 169-189.

Davis, F.D., 1989, Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

Debela, N. (2004). A closer look at distance learning from students' perspective: a qualitative analysis of web based online courses. *Journal of systemics, cybernetics and Informatics*, 2(6).

Emck, J.H., & Ferguson-Hessler, M.G.M. (1981). A computer-managed Keller plan. *Physics education*, 16(1), 46-49.

Engelbrecht, J. & Harding, A. (2001). WWW mathematics at the university of Pretoria: the trail run. *South African journal of science*, 97(9-10), 368-370.

Green, B.A, (1971). Physics teaching by the Keller plan at MIT. *American journal of physics*, 39(7), 764-775.

Green, S.M, Voegeli, D., Harrison, M., Phillips, J., Knowles, J., Weaver, M., et al. (2003). Evaluating the use of streaming video to support student learning in a first-year life sciences course for student nurses. *Nurse education today*, 23(4), 255-261.

Hambleton, I.R., Foster, W.H., & Richardson, J.T.E. (1998). Improving student learning using the personalised system of instruction. *Higher Education*, 35(2)¸187-203.

Hereford, S.M. (1979). The Keller plan within a conventional academic environment: an empirical 'meta-analytic' study. *Engineering education*, 70(3), 250-260.

Hiltz, S.R., & Turoff, M. (2002). What makes learning networks effective? *Communication of the ACM*, 45(4), 56-59.

Hoskins, S.L., & van Hooff, J.C. (2005), Motivation and ability: which students use online learning and what influence does it have on their achievement? *British journal of educational technology*, 36(2), 177-192.

Johnson, G.M. (2005). Student alienation, academic achievement, and WebCT use. *Educational technology & society*, 8(2), 179-189.

Jones, G.H, & Jones, B.H. (2005). A comparison of teacher and students attitudes concerning use and effectiveness of web-based course management software. *Educational technology & society*, 8(2), 125-135.

Keller, F.S. (1968). "Good-bye, teacher …" *Journal of applied behavior analysis*, 1(1) 79-89.
Keller, F.S., & Sherman, J.G. (1974). *The Keller plan handbook: essays on personalized system of instruction*. Menlo Park, W.A. Benjamin.

Kinsner, W., & Pear, J.J. (1988). Computer-aided personalized system of instruction for the virtual classroom. *Canadian journal of educational communication*, 17(1), 21-36.

Koen, B.V. (2005). Creating a sense of "presence" in a web-based PSI course: the search for Mark Hopkins' log in a digital world. *IEEE transactions on education*, 48(4), 599-604.

Kraemer, E.W. (2003). Developing the online learning environment: the pros and cons of using WebCT for library instruction. *Information technology and libraries*, 22(2), 87-92.

Kulik, J.A., Kulik, C.-L.C., & Cohen, P.A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American psychologist*, 34(4), 307-318.

Kyle, J. (1999). Mathletics – a review. Maths & Stats, 10(4), 39-41.

Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International journal of human-computer interaction*, 7(1), 57-78.

Pear, J.J., & Crone-Todd, D.E. (2002). A social constructivist approach to computer-mediated instruction. *Computer & Education*, 38(1-3), 221-231.

Pear, J.J., & Novak, M. (1996). Computer-aided personalized system of instruction: a program evaluation. *Teaching of psychology*, 23(2), 119-123.

Pear, J.J. (2003). Enhanced feedback using computer-aided personalized system of instruction. In W. Buskist, V. Hevern, B. K. Saville, & T. Zinn (Eds.). *Essays from e-xcellence in teaching* (vol. 3, ch. 11). Washington, DC: APA Division 2, Society for the Teaching of Psychology.

Pelayo-Alvarez, M., Albert-Ros, X., Gil-Latorre, F., & Gutierrez-Sigler, D. (2000). Feasibility analysis of a personalized training plan for learning research methodology. *Medical education*, 34(2), 139-145.

Poirier, T.I., & O'Neil, C.K. (2000). Use of web technology and active learning strategies in a quality assessment methods course. *American journal of pharmaceutical education*, 64(3), 289-298.

Rae, A. (1993). Self-paced learning with video for undergraduates: a multimedia Keller plan. British journal of educational technology, 24(1), 43-51.

Ramsden, P., & Entwistle, N.J. (1981). Effects of academic departments on student's approaches to studying. *British Journal of Educational psychology,* 51(Nov), 368-383.

Ramsden, P. (2003). *Learning to teach in higher education* (2nd ed). London, RoutledgeFalmer.

Roth, C.H. (1993). Computer aids for teaching logic design. *Frontiers in education conference* (p. 188-191). IEEE.

Schultze-Mosgau, S., Zielinski T., & Lochner, J. (2004). Web-based, virtual course units as a didactic concept for medical teaching. *Medical teacher*, 26(4), 336-342.

Sheehan, T.J. (1978). Statistics for medical students: personalizing the Keller plan. *The American statistician*, 32(3), 96-99.

Watson, J.M. (1986). The Keller plan, final examinations, and log-term retention. *Journal for research in mathematics education*, 17(1), 60-68.

Wernet, S.P., Olliges, R.H., & Delicath, T.A. (2000). Postcourse evaluations of WebCT (Web Course Tools) classes by social work students. *Research on social work practice*, 10(4), 487-504.

Zhang, D., Zhao, J. L., Zhou, L. & Nunamaker, J.F. (2004). Can E-learning replace classroom learning? *Communication of the ACM*, 47(5), 75-79.

APPENDIX

Table 8: Questions of the online surveys.

No	Included in survey	Question	Answer type
1	T1-04, T1-05, T2-04, T2-05.	All things considered, how would you rate the quality of the [first/second] term of this module?	A
2	T1-04, T1-05, T2-04, T2-05.	How would you rate the level of difficulty of [the second term of] the module in your case?	B
3	T1-04, T1-05, T2-04, T2-05.	How would you rate your interest for [the second term of] the module?	C
4	T1-04, T1-05, T2-04, T2-05.	How useless/useful did you find the lab sessions [in the second term]?	D
5	T1-04, T1-05, T2-04, T2-05.	How useless/useful did you find the seminar sessions [in the second term]?	D
6	T1-04, T1-05, T2-04, T2-05.	How useless/useful did you find the lectures [in the second term]?	D
7	T1-04.	Roughly what proportion of the lecture/lab/seminar sessions have you attended?	E
8	T1-05, T2-04, T2-05.	Roughly what proportion of the lecture sessions did you attend [in the second term]?	E
9	T1-05, T2-04, T2-05.	Roughly what proportion of the lab sessions did you attend [in the second term]?	E
10	T1-05, T2-04, T2-05.	Roughly what proportion of the seminars sessions did you attend?	E
11	T1-05,	How useless/useful did you find the end of the lecture questions?	D
12	T1-04, T1-05, T2-04, T2-05.	How much of the subject matter covered in the [first/second] term did you already know?	E
13	T1-04, T1-05, T2-04, T2-05.	What is your educational background?	F
14	T1-04, T1-05, T2-04, T2-05.	Do you have access to a PC / laptop outside the lab sessions to work on?	G
15	T1-05.	Do you have access to a PC / laptop with a DVD player outside the lab sessions to work on?	G
16	T1-04, T1-05.	Which grade did you receive for the mid semester test?	H
17	T2-04, T2-05.	Which grade did you receive for Mathletics test 1?	H
18	T2-04, T2-05.	Which grade did you receive for Mathletics test 2?	H
19	T2-04, T2-05.	Which grade did you receive for Statistical Report coursework?	H
20	T1-04, T1-05, T2-04, T2-05.	How would you rate the usability of the WebCT environment for this module [in the second term]?	C
21	T1-04, T1-05, T2-04, T2-05.	How useless/useful did you find the WebCT discussion board of this module [in the second term]?	D

22	T1-04, T1-05,	How useless/useful did you find the self-tests on WebCT?	D
23	T2-04, T2-05.	How useless/useful did you find the Mathletics self-tests (not the assessments)?	D
24	T1-04.	How useless/useful did you find the videos?	D
25	T1-05.	How useless/useful did you find the videos that discuss questions?	D
26	T1-05.	How useless/useful did you find the videos that gave an introduction?	D
27	T1-05.	How useless/useful did you find the videos that gave a summary?	D
28	T1-05.	Roughly how many video clips that give an introduction to a module/unit have you watched?	I
29	T1-05.	Roughly how many video clips that discuss a question have you watched?	J
30	T1-05.	Roughly how many video clips that give a summary of a unit have you watched?	I
31	T1-04, T1-05.	How useless/useful did you find the example questions of the mid semester test?	D
32	T1-04, T1-05.	How useless/useful did you find the written material (Modules 1-5)?	D
33	T1-04, T1-05.	How useless/useful did you find the book “Discrete mathematics with application” by S.S. Epp?	D
34	T1-04, T1-05.	How useless/useful did you find the book “Computer Science: an overview” by J. Glenn Brookshear?	D
35	T1-04, T1-05, T2-04, T2-05.	How useless/useful did you find the feedback [regarding the assessment of your coursework/ you received on your Statistical Report coursework]?	D
36	T2-04, T2-05.	How useless/useful did you find the Lecture Handouts used in the second semester?	D
37	T2-4, T2-05.	How useless/useful did you find the Lab Session Notes used in the second term?	D
38	T2-04, T2-05.	How useless/useful did you find the Seminar: Problem Sheets in the second term?	D
39	T2-04, T2-05.	How useless/useful did you find the example exam questions on WebCT?	D
40	T2-04, T2-05.	How useless/useful did you find the Study Guide?	D
41	T2-04, T2-05.	How useless/useful did you find the template document you could use to create your Statistical Report?	D
42	T1-04, T1-05.	Any comments that you like to make about the module?	K
43	T1-05, T2-04, T2-05.	I find that at times studying gives me a feeling of deep personal satisfaction	L
44	T1-05, T2-04, T2-05.	I find that I have to do enough work on a topic so that I can form my own conclusions before I am satisfied.	L
45	T1-05, T2-04, T2-05.	My aim is to pass the course while doing as little work as possible.	L
46	T1-05, T2-04, T2-05.	I only study seriously what’s given out in class or in the course outlines.	L
47	T1-05, T2-04, T2-05.	I feel that virtually any topic can be highly interesting once I get into it.	L
48	T1-05, T2-04, T2-05.	I find most new topics interesting and often spend extra time trying to obtain more information about them.	L
49	T1-05, T2-04, T2-05.	I do not find my course very interesting so I keep my work to the minimum.	L
50	T1-05, T2-04, T2-05.	I learn some things by rote, going over and over them until I know them by heart even if I do not understand them.	L
51	T1-05, T2-04, T2-05.	I find that studying academic topics can at times be as exciting as a good novel or movie.	L
52	T1-05, T2-04, T2-05.	I test myself on important topics until I understand them completely.	L
53	T1-05, T2-04,	I find I can get by in most assessments by memorising key	L

	T2-05.	sections rather than trying to understand them.	
54	T1-05, T2-04, T2-05.	I generally restrict my study to what is specifically set as I think it is unnecessary to do anything extra.	L
55	T1-05, T2-04, T2-05.	I work hard at my studies because I find the material interesting.	L
56	T1-05, T2-04, T2-05.	I spend a lot of my free time finding out more about interesting topics which have been discussed in different classes.	L
57	T1-05, T2-04, T2-05.	I find it is not helpful to study topics in depth. It confuses and wastes time, when all you need is a passing acquaintance with topics.	L
58	T1-05, T2-04, T2-05.	I believe that lecturers shouldn't expect students to spend significant amounts of time studying material everyone knows won't be examined.	L
59	T1-05, T2-04, T2-05.	I come to most classes with questions in mind that I want answering.	L
60	T1-05, T2-04, T2-05.	I make a point of looking at most of the suggested readings that go with the lectures.	L
61	T1-05, T2-04, T2-05.	I see no point in learning material which is not likely to be in the examination.	L
62	T1-05, T2-04, T2-05.	I find the best way to pass examinations is to try to remember answers to likely questions.	L

Note, Questions 43-62 were taken from the R-SPQ-2F inventory (Biggs, et al., 2001). Reproduced with permission from the British Journal of Educational Psychology, © The British Psychological Society. The responses on these questions where use to calculate the score on the learning approach scales in the following way: Deep Motive (DM) = Q43 + Q47+ Q51 + Q55 + Q59; Deep Strategy (DS) = Q44 + Q48 + Q52 + Q56 + Q60; Surface Motive (SM) = Q45 + Q49 + Q53 + Q57 + Q61; Surface Strategy (SS) = Q46 + Q50 + Q54 + Q58 + Q62; Deep Approach = DM + DS; Surface Approach = SM + SS. Question phrases within [] were adapted to the context of the term. T1-04 = Term 1 in 2004; T1-05 = term 2 in 2005; T2-04 = term 2 in 2004; T2-05 = term 2 in 2005.

Table 9: Answer types used in the online surveys.

Answer type	Answer options
A	1) poor; 2) fair; 3) good; 4) very good; 5) not applicable.
B	1) very difficult; 2) difficult; 3) average; 4) easy; 5) very easy; 6) not applicable.
C	1) very low; 2) low; 3) average; 4) high; 5) very high; 6) not applicable
D	1) useless; 2) some parts useless some parts useful; 3) useful; 4) very useful; 5) not applicable.
E	1) 0-20%; 2) 21-40%; 3) 41-60%; 4) 61-80%; 5) 81-100%; 6) not applicable.
F	a) A levels; b) BTEC; c) GNVQ; d) Access; e) Other.
G	1) never; 2) sometimes; 3) regular; 4) not applicable.
H	1) F; 2) E; 3) D; 4) C; 5) B; 6) A; 7) not applicable.
I	1) 0-4; 2) 5-9; 3) 10-14; 4) 15 or more.
J	1) 0-4; 2) 5-9; 3) 10-14; 4) 15-19; 5) 20-24; 6) 25-29; 7) 30 or more.
K	*Open answer*
L	1) never or only rarely true of me; 2) sometimes true of me; 3) true of me about half the time; 4) frequently true of me; 5) always or almost always true of me.

Table 10: Interview questions.

Category	No	Question
Students approach to learning	1	Many students have different ways how they approach their studies. There is not a single best approach, some people like going to lectures, seminars, and lab regularly, but other like to work on their own at home. Some people spread the learning across the year; others like to focus their learning activities at the moment before an assessment. Can you tell me how you approached your learning of this module, and also why did you approach it like this? Let's start how you began the year and progress through the year up to the days of the exams.
	2	What were the reasons for you to attend a lecture/seminar/lab session?
	3	What were your motivations to start studying this module? Or what were the thinks that stop you from studying?
	4	How much time did you spend studying on this module and why did you spend this amount of time to the module?
Teaching approach	5	In the first semester, [*name lecturer*] only looked at one part of the material in his lectures -units 5- the other parts, units 1-4, was covered by seminars and lab sessions. Whilst in the second semester [*name lecturer*] covered all topics in her lectures, while the seminars and lab session focuses more on details and practical issues. Please tell me how you perceived this approach of teaching?
	6	This module was assessed via both 50% coursework and 50% exams. Coursework included 6 tasks (Tarski, Mid term exam, Project two, Mathletics test 1 & 2, and Statistical Report). How do you perceive this approach of assessment?
	7	Why do you like or not like this approach; and what suggestions would you like to make?
	8	Were you clear about the assessment process before you undertook them?
Student Characteristics	9	Student of level one come from various educational backgrounds, for instance some might study math subject in GCSE and A level and some may not. What was your educational background?
	10	If you look at your background and the material in the module, which thing did you already know and which things were new for you?
	11	Students also differ in terms of where they stayed during their study period, on campus or off campus. Students that live off campus have different travelling rime to come to the university. What was your situation?
	12	How might this have affected your study?
	13	Some student only had access to a PC and Internet at the university; others have also access to these facilities at home. What is your situation?
	14	How might this have affected your study?
	15	Some students work or engaged in other activities during their study period, such as working, hobbies, sport, or other studies activities. What was your situation?
	16	How might this have affected your study?
	17	Some students have many friends and they study and do coursework in groups, whilst others do their study alone. What as your situation?
	18	How might this have affected your study?
OLE tools	19	Different modules have different way to support student learning activities. This module offered WebCT and printed material of lectures, problem sheets and lab notes at the beginning of the semester. Different student used these resources and facilities differently. How would you explain your way of using these and why did you use WebCT? What are the facilities you particular used and which didn't you use, and why?
	20	Did you use any of these facilities?
	21	Why did you use them?

	22	Do you have any suggestions to improve these facilities?
Closing question	23	Is there anything else what you like to mention, which was not covered by the previous questions?

Table 11: Questions used in usability test.

Scale	No	Question
Ease of use	1	Learning to operate *[name OLE tool]* would be easy for me.
	2	I would find it easy to get *[name OLE tool]* to do what I want it to do.
	3	My interaction with *[name OLE tool]* would be clear and understandable.
	4	I would find *[name OLE tool]* to be flexible to interact with.
	5	It would be easy for me to become skillful at using *[name OLE tool]*
	6	I would find *[name OLE tool]* easy to use.
Satisfaction	7	The interface of *[name OLE tool]* was pleasant.
	8	I like using the interface of *[name OLE tool]*.

Note, Questions 1-6 were adapted from the Perceived Usefulness and Ease-of-use (PUEU) questionnaire (Davis, 1989) (Reproduced with permission from MIS Quarterly). The answer option for questions 1-6 was a 7-point Likert scale ranging from Unlikely to Likely with 1) extremely; 2) quite; 3) slightly; 4) neither; 5) slightly; 6) quite; 7) extremely. Question 7 and 8 were adapted from the Post-Study System Usability Questionnaire (PSSUQ) (Lewis, 1995) (Reproduced with permission from Lawrence Erlbaum Associates, Inc. and the author). The answer option for question 7 and 8 was a 7-point Likert scale ranging from 1) strongly disagree to 7) strongly agree. The phrase "*[name OLE tool]"* in each question was replaced by name of the OLE tool. The score on the Ease-of-use scale was calculated by taking the average response on questions 1 till 6, and the score on the Satisfaction scale was calculate by taking the average response on question 7 and 8.