Computer Based Revision

Edmund Furse

Swansea Metropolitan University

A computer system on the web known as xplus12 has been developed which supports KS3 revision in Number and Algebra. This has been evaluated with a year 8 top-set class on two trials in Number with a little Algebra. Since all the data is stored in a database including timings and all student attempts at answers, it is possible to identify a number of behaviour patterns among the pupils. These include the identification of groups such as "rubbishers" and "rushers" who essentially abuse the system. Although an effect size of 0.9 was found in the first trial it was not statistically significant due to the low numbers of pupils completing the test, and no improvement was found in the second trial. Misconceptions were also identified by the system. A number of suggestions are discussed for improvements of the system including improved examples with animation and explaining answers. Also included are techniques to handle rubbishers and rushers.

Keywords: web, revision, computer based learning

Introduction

The use of computer systems to assist in the teaching of mathematics is growing worldwide, especially in the USA. Anderson's Cognitive Tutor (Algebra) from Carnegie Melon University is installed in over 2000 schools. Computer remedial teaching of algebra has been a growing market in American Universities for some time, and in the last few years this is now filtering down into schools. It is convenient to use computers for remedial teaching since the pupils can do it at their own pace and institutions may use cheaper staff.

Here we report on an evaluation of the xplus12 revision system that runs on the web (www.xplus12.com). The system has been developed over a number of years and uses AI techniques both in the automatic synthesis of worksheets, and in sophisticated answer checking. The development system is built in LISP and Java, and the delivery system runs on all computer platforms that support Java. The website has sections on worksheets, starters, games and revision and has over 40 worksheets. A unique property of the system is automatic transition from type-in questions into multiple-choice questions after a suitable number of wrong attempts. Most systems available only support multiple-choice questions but these are too reliant on recognition memory. But the danger of type-in questions is that a pupil can get stuck on a question and not be able to progress.

Mathematics is an ideal subject for a computer based learning system since the computer is able to correctly check if a pupil's answer is correct. This can even be done for algebraic answers by use of suitable AI techniques. Most systems fudge this and cannot handle all the possible algebraic answers. This study finds an impact of the revision software in xplus12 but it is not statistically significant. Too many pupils had to be excluded from the trials due to non-completion of the test or writing rubbish. The paper discusses in particular two important groups of students: "rubbishers" and "rushers". The discussion section looks to a number of improvements of the system, which should make learning more effective.

Literature Review

Several studies have been done on the effectiveness of the use of computers in the teaching and learning of mathematics. The results appear to be mixed, with some finding no effect and others a small positive effect. Tienken and Wilson (2007) have a review of the impact of CAI on mathematics achievement. They found an effect size of 0.12, which was statistically significant $p < 0.05$. Whereas, The Electronic Education Report (2007) reports on a large study of seven maths computer packages and found none had a statistically significant improvement on performance. Zhu and Polianskaia (2007) provide a comparison between traditional lecture and computer-mediated instruction and find no statistically significant difference over a ten-year period. Paper and pencil was more effective. Clearly computer systems differ widely in many aspects and this has impacts on their effectiveness. One of the highest effect sizes (of 1) is reported by Anderson (1995) for his Cognitive Tutor in Geometry, but surprisingly there is no difference in performance with his Algebra Tutor between controls and students using the system. Anderson attributes this to poor transfer from the computer system to paper and pencil solutions.

One of the problems of some of the positive results studies is that the students were self-selected. This is particularly true for revision systems, like SAM Learning. It is also an issue that is recognised in the literature, for example Biesenger & Crippen (2008). However, several systems have been used with whole classes that then encounter problems of non-engagement by some students. Egan, Jefferies and Johal (2006) introduce the classification of lurkers, workers and shirkers in an online teaching system. There have been a number of studies of guessing including Beal et al (2008). Several systems, including Aleven et al (2006) attempt to get the computer system to identify the students abusing the system and get it to react appropriately; for example students who click through hints at speed are told to slow down. Abusing the computer system is part of a more general problem of disaffected pupils and those who avoid engaging with lessons (Dowson & McInerney 2001).

Most studies of the effectiveness of computer systems compare with a control group that uses paper and pencil. But paper and pencil questions do not always easily translate into a computer form. This is especially a problem with assessment and is one of the reasons for the prevalence of multiple-choice type questions. There is also the problem of the assessment of partially correct answers, since this is usually dealt with in an explicit way in paper-based assessments; but it is difficult to do on computers. Lindsay (1999) compared the use of a computer algebra system with paper and pencil techniques. Paper and pencil was more effective. Threlfall et al (2007) analysed the differences between paper versions and computer versions of KS2 and KS3 questions. Pupils sometimes do better with the computer versions because there is more opportunity to explore different answers even though the system did not give them feedback to tell them if it was correct or not. Ashton et al (2006) argue the importance of computer systems being able to evaluate partially correct answers to mathematics questions, but there are few if any that can do this. Computer systems may also teach a subject in a completely different way to how a student may do it in class using pencil and paper. For example, Anderson's Algebra Tutor is used in over 2000 American schools but uses box diagrams to explain algebraic simplification which does not naturally translate to a paper based method.

The quality of feedback from a teacher or a computer system is an important factor in how well pupils learn. Many systems only give a correct/incorrect response, which may not be very effective to aid learning unless they can eventually discover the correct result. Some systems give the correct answer and this can be done in a three ways: (1) immediately when the pupil gets the question wrong; (2) after a fixed number of attempts; (3) via a help system which has run to the end of helpful advice and just tells the student the answer. A few systems go further and give an explanation of why the answer is correct. There is also an issue of the timing of feedback. Some systems, including MyMaths, do not give feedback until all the questions have been answered. But Tallent-Runnels et al (2005) give a review on how to teach online and identify the importance of prompt feedback.

Received wisdom in Tutoring Systems suggested that users should be given a large amount of user control in navigation of the system (e.g. Wenger (1987)). But more recent research suggests that weaker students need stronger control by the system in order to learn effectively (Kopcha and Sullivan 2008). Ketamo and Alajaaski (2008) found that in the use of a multiple-choice based system there was a lot of guessing and students did not have the skill to choose appropriate materials to study. Similarly, Mezirow (1995) found that students tended to choose topics to study that they were already familiar with rather than the ones they needed to work on.

Computer System and Method

The xplus12 revision system covers a subset of KS3 Number and Algebra by means of interactive worksheets. The pupil is given simple feedback to each question on typing a return after the answer in the field. They are normally given five attempts, after which a multiple-choice version of the question replaces the type-in question.

When the pupil has completed the first worksheet and attempted all the questions, the computer determines which are the pupil's weakest sections. The pupil then does a short remedial worksheet for each topic they are weak on. In Number 3 these worksheets just contain 4 interactive questions, but in Number/Algebra 2 there was an example as well.

Once the pupil has completed the remedial worksheets the pupil then goes on to the last worksheet that is very similar to the first one. It has the same sections and also 2 questions per section. But there is no feedback and it works as a test.

Qualitative Study Results

A questionnaire was given to the pupils and there was also a discussion with four pupils. The closed questions had a five point Likert scale response. The question "How easy do you find xplus12 to use?" had a median of 4 as did the question "How much did xplus12 help you to learn maths?". The results indicated that the pupils liked the xplus12 system, found it easy to use, and felt they learned mathematics from it. There were also open questions such as "What do you particularly like about xplus12?" and several pupils indicated they liked the type-in questions turning into multiple-choice questions. On the negative side they thought there was more need of colour, animation and more games.

Quantitative Study Results

First Trial

22 students did the revision exercises but only 14 students provided reliable data that can be used for analysis. Eight students had to be excluded since they did not complete the revision exercises and so it is not possible to compare pre and post test results. The pre test scores are based on the pupils' first attempt at the questions. They may go on to attempt each question several times, and since the system eventually ends up with a multiple-choice question they eventually get every question correct. The post-test scores are just based on the answer since they are only given one attempt with no feedback. The pre test mean was 12.9 and the post-test mean of 15.6 with a standard deviation of 3, which represents an effect size of 0.9. Unfortunately this difference is not statistically significant at the 0.05 level.

One of the problems encountered was that some pupils would attempt to get to the multiple-choice version of a question as soon as possible by guessing or writing rubbish. Five pupils were identified using this tactic and they may be labelled with the name "rubbishers".

Second Trial

In order to try to decrease the number of pupils "rubbishing", the mechanism for deciding when to present the multiple-choice question was changed from a threshold of 5 attempts. Instead, whenever the pupil answered a question too fast (less than a second) and had the wrong answer, it was assumed they were guessing. The penalty for this was that the threshold number of attempts was increased by one. Unfortunately this change was a disaster with two pupils making 60 attempts at questions! Four pupils had at least 30 attempts at a question. We can operationalise this by identifying pupils who have at least two questions with 15 or more attempts as "rushers". An unfortunate consequence of the changed attempt threshold policy was that the rushers spent a great deal of time on the pre test and most did not complete the post-test. A few pupils also repeated the test thus invalidating the pre-scores.

One of the advantages of a computer system of this type is that it is possible to store all the attempted answers a pupil makes to a question and to identify misconceptions. A particular case of this occurred in question one of the pre test in BODMAS:

$3 - 12 + 4 = ?$

Eleven pupils gave the correct answer of -5 but ten students gave the answer of 13, and three -13. Clearly there is a misconception here and pupils are taking the smaller number from the larger one. Interestingly this misconception did not occur in the pre-test with the BODMAS questions:

$2 + 3 \times 5 = ?$

$2 + 7 - 3 = ?$

and since the post test question is easier, they are not matched questions.

Discussion

It is clear that there is a need for the computer system to be improved. Too many pupils are disaffected and not enough learning is taking place. I take the theoretical stance that if pupils want to learn, then it should be possible to design a computer system to help them to learn.

It can be argued that pupils behave as rushers and rubbishers because they cannot do the questions however hard they try. There is evidence of rusher and rubbisher personality types from the data since these behaviours persist, even across trials. Furthermore, several pupils type rubbish immediately they start a question, indicating no effort to try to answer the question correctly. On the other hand, no pupil is 100% a rusher/rubbisher (RR), but RRs are likely to attempt questions they find easier. It appears that they have a low threshold of effort and may resort to RR if their first 3 attempts are wrong. It is also true that rushers are likely to exacerbate rubbishing since the computer system was postponing the multiple-choice question as a result of their behaviour.

A simple solution to engaging and helping RRs is to provide interactive help. A later version of the system provides optional interactive help after three attempts. This is also the approach taken by Anderson’s Cognitive Tutor, although their help system is much more complex. It is also clear that the multiple-choice version should be triggered always after five attempts, and not penalise RRs by making it later.

Rushers can be slowed down by giving them a pop up message that tells them to slow down and they have to click to continue. It is easy for the system to identify if the pupil is answering questions too fast and incorrectly. It is more difficult to check if the pupil is writing rubbish but long answers of consecutive numbers are usually indicative. A strong rubbish checker would be safer than a weak one since a genuine pupil would take a dim view of being challenged at writing rubbish.

It is also clear that not enough pupils are learning from their mistakes. It is not clear if the pupils are actually benefiting from the multiple choice questions apart from letting them get onto the next question rather than being permanently stuck. The system does not currently record the timing data of the multiple-choice attempts but this could be changed. Thus, again the pupils could be slowed down with a suitable dialog if they are clicking the options too fast.

More importantly, the pupils need to have an explanation of the answer. It is not sufficient just to be told the correct answer, and rushers are likely to just want to get onto the next question rather than digest the answer. One simple solution is to leave the correct answer and its explanation up for 5 seconds before they can proceed. A more sophisticated approach is to use animated text. Also, it is probably desirable for the system to check if the pupils think they understand their mistakes.

Animation has also been used to improve the examples. Direct observation in class suggests that many pupils rush straight into answering the questions in a worksheet without reading the examples. Animating the examples makes them more interesting and encourages them to read them. Admittedly a disaffected pupil may not look at the animation, but there is nothing else for them to do on the computer. The system now also asks them how well they understand the example before proceeding to the questions. In principle, their answer can inform future navigation.

It is straightforward to incorporate misconception handling in the system if the author knows of suitable misconceptions. These can be included as alternative multiple choice questions and checked for also during the type-in stage.

It is clear from the qualitative study that the colour and appearance of the system needs to be improved. This came out clearly from the qualitative study, but it is possible that deeper issues may be more important. They also wanted more games. Some might argue that provision of games having completed the revision might provide motivation to do the work. But there is a danger that this will encourage some pupils to rush their work to get to the games. It is better to provide some form of intrinsic motivation within the system.

Conclusion

The xplus12 revision system improves learning but the quantitative results are not statistically significant. The new improvements of animated examples, preventing rushing, misconception handling and explanations of answers have been implemented and should improve learning further. If more schools use the system then further research should prove the advantages of these enhancements.

References

Aleven V., B. McLaren , I. Roll, and K. Koedinger. 2006. 'Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor. *International Jnl. of Artificial Intelligence in Education*; Jul2006, Vol. 16 Issue 2, 101-128

Anderson J.R., A. T. Corbett, K. R. Koedinger and R. Pelletier. 1995. Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*; 1995, Vol. 4 Issue 2, p167, 41p

Ashton, H.S., C. E. Beevers, A. A. Korabinski, and M. A. Youngson. 2006. Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology*, Vol 37, No 1, 2006, 93–119.

Beal, C. R., L. Qu, and H. Lee. 2008. Mathematics motivation and achievement as predictors of high school students' guessing and help-seeking with instructional software. *Journal of Computer Assisted Learning*, Dec2008, Vol. 24 Issue 6, p507-514

Biesenger K. & K. Crippen. 2008. The Impact of an Online Remediation Site on Performance Related to High School Mathematics Proficiency. *Journal of Computers in Mathematics and Science Teaching* (2008)27(1), 5-17

Dowson, M., & D. M. McInerney. 2001. Psychological Parameters of Students' Social and Work Avoidance Goals: A Qualitative Investigation. *Journal of Educational Psychology*, 93, 35-42.

Egan C., A. Jefferies and J. Johal. 2006. Providing Fine-grained Feedback Within an On-line Learning System – Identifying the Workers from the Lurkers and the Shirkers. *The Electronic Journal of e-Learning* Volume 4 Issue 1, 15-24.

Electronic Education Report. 2007. Study Reports Software Did Not Raise Test Scores in Reading and Math, *Simba Information*, Stanford, April 20 2007, 1-3.

Kopcha T.J. and H. Sullivan. 2008. Learner preferences and prior knowledge.

In *Learner-controlled computer-based instruction*. Education Tech Research Dev (2008) 56:265–286.

Mezirow, J. 1995. *Fostering critical reflection in adulthood : A guide to trans-formative and emancipatory learning*. Jossey-Bass.

Nguyen, D., J. Yi-Chuan, and D. Allen. 2006. The Impact of Web-Based Assessment and Practice on Students' Mathematics Learning Attitudes. *Jnl. of Computers in Mathematics & Science Teaching*; 2006, Vol. 25 Issue 3, 251-279

Tallent-Runnels M.K., S. Cooper, W. Y. Lan , J. A. Thomas, and C. Busby. 2005. How to Teach Online: What the Research Says. *Distance Learning* Vol 2. Issue 1

Thelfall J., P. Pool, M. Homer and B. Swinnerton. 2007. Implicit aspects of paper and pencil mathematics assessment that come to light through the use of the computer. *Educational Studies in Mathematics* (2007) 66:335–348.

Tienken C.H. and M. J. Wilson. 2007. The Impact of Computer Assisted Instruction on Seventh-Grade Students Mathematics Achievement. *Planning and Changing*, Vol. 38, No. 3 & 4, 2007, 181-190.

Wenger, E. 1987. *Artificial Intelligence and Tutoring Systems*. Morgan Kaufmann.

Zhu, Q and G. Polianskaia. 2007. A Comparison of Traditional Lecture and Computer-mediated Instruction in Developmental Mathematics, *Research and Teaching in Developmental Education*. Fall2007, Vol. 24 Issue 1, 63-82