

# Performance and Attitude of Children in Computer Based Versus Paper Based Testing

Gavin Sim  
Department of Computing  
University of Central Lancashire  
United Kingdom  
grsim@uclan.ac.uk

Matthew Horton  
Department of Computing  
University of Central Lancashire  
United Kingdom  
mplhorton@uclan.ac.uk

**Abstract:** This paper reports the findings of an investigation of children's performance and attitude towards a paper based and computer based test. Twenty children, aged between 7 and 8 of mixed gender, participated in this study using a commercial software application. The children's attitude towards the software was captured through the use of a smarty-o-meter to indicate their level of preference and the performance was based on their tests scores. The conclusions highlight the children's preference for using computer in their assessment and shows that there was no difference in performance between the two test modes.

## Introduction

There is growing increase in the use of computers for assessment purposes within Higher Educational institutions globally (Sim, Holifield, & Brown, 2004). The UK governments' aim is to introduce new online tests of information and communication technology skills at key stage 3 (age 13 to 14) by 2008. Currently pilot studies have been conducted within UK schools for the delivery of summative assessment via the web to determine the feasibility of this strategy (Ashton, Schofield, & Woodger, 2003; Nugent, 2003). This raises the questions of how acceptable the technique is to children, and is paper based comparable to computer based testing.

There has been considerable research conducted into the comparability of computer based versus paper based testing (Pommerich, 2004; Zandvliet, 1997) which has mainly focused on adult learners. Clariana & Wallace (2002) indicate that there is mounting empirical evidence that identical paper-based and computer-based tests will not obtain the same results due to a phenomena known as 'test mode effect'. Evidence to support this theory can be found in studies by (Russell & Haney, 1997) which showed students performing better on the computer test, whilst in another study students outperformed the computer based version when using paper (Federico, 1989). It has been argued that in presenting a test on computer, a qualitatively different testing experience is created (McDonald, 2002).

There are numerous variables that impact on student's performance when questions are presented on a computer, such as the quality of the monitor (Schenkman, Fukuda, & Persson, 1999), the way text is displayed on screen (Dyson & Kipping, 1997), reading from a monitor is slower than paper (Mayes, Sims, & Koonce, 2001) and the problems of obtaining a feel for the exam when only a single question is presented (Liu, Papathanasiou, & Hao, 2001). However, when including multimedia elements in the question these present additional issues, for example proving comparable alternatives for paper based versions (Bennett et al., 1999) and cognitive load (Cooper, 1990; Kirschner, 2002)

Evidence is presented signifying the differences between the two test modes and the variables that can impact the results. However, there is conflicting evidence suggesting that there is no difference between the two formats (Greenwood et al., 2000; Zandvliet, 1997). This paper aims to investigate the comparability of children's performance in a paper based and computer based test for key stage 1 (ages 5 to 7) science and examine their preference for a particular technique.

## **Method**

### **Sample**

The sample consisted of 20 children of both genders (9 boys and 11 girls), aged between 7 and 8, from a Year 3 class at a Primary School in Lancashire, England. The sample contained children with varying ability levels, some being able to read the questions and instructions and some requiring assistance. All the children had English as their first language and have been studying the National Curriculum since Foundation Stage. The children were familiar with the subject domain presented in the software as it is taught at key stage 1 of the National Curriculum.

### **Equipment**

Four laptop computers were used in this study to administer the tests. All were Ergo Preceptor 5 laptops with 1600 MHz Pentium processors, 512 Mb RAM, and 40Gb hard drives. The machines all had the same 15 inch TFT screen set at a resolution of 1024 x 768 pixels and a Genius Netscroll mouse attached, although each laptop did have a built in glide pad that could also be used. Ensuring the equipment is the same for all test participants is essential as monitor resolution is a factor that can influence test results (Pommerich, 2004). All the laptops were running Windows XP Professional as the operating system and had Camtasia Recorder 3.0.2 screen capture software installed.

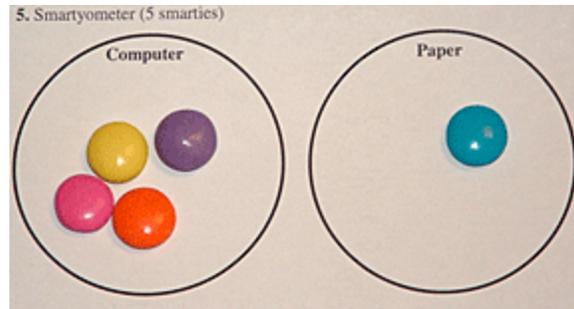
### **Procedure**

Four researchers took part in the study, each one supervising one of the laptops. Each child was asked to complete a short questionnaire before starting the test which was designed to gather information such as the gender of the child and their computer usage at home. The children were then given two tests, one using the Europress Key Stage 1 Science software and the other using a paper based test with questions in the same style as in the Europress software. A counterbalancing technique was used to minimise any learning effect caused by doing one test followed by another, with two of the researchers administering the paper tests first whilst the other two administered the computer test first.

The questions on the computer based test relate to ten different areas of the science curriculum and are randomly generated from a question bank. The paper based test was constructed by selecting a sample of 10 questions from each of the areas. The questions were then converted to a paper based version for example, a drag and a drop style question on the computer was converted to a fill in the blanks on paper.

When administering the paper based tests the researchers asked the children if they required help reading the questions and if so this help was provided. The computer based test required the researchers to start Camtasia Recorder before each child began the test in order to capture the interaction as part of usability testing. The researchers were given pre-designed sheets to record information about each question such as the type of question, whether it contained words or pictures, the number of options to choose from and also whether the child required help understanding the questions and whether they got them right or wrong. There was also a space for the researchers to record any other observations or comments.

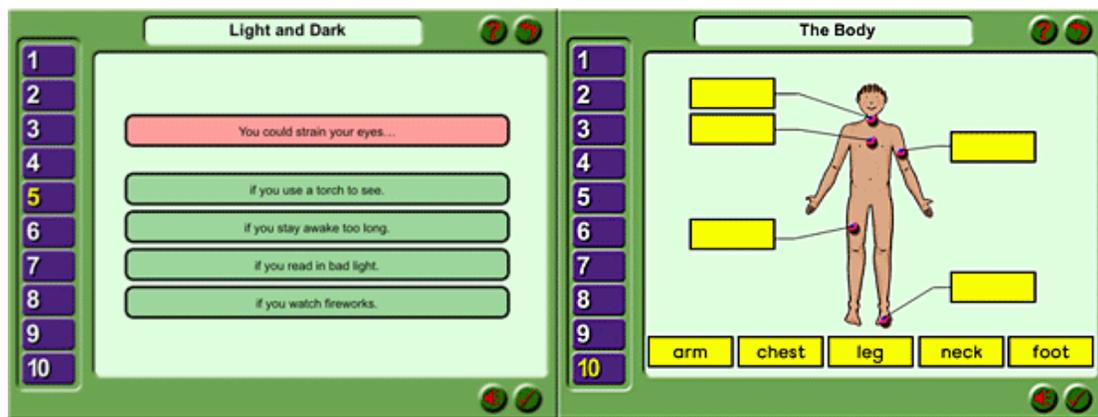
Once the children had completed the test they were asked to complete a smarty-o-meter (figure 1). They were given five Smarties (chocolate sweets) and asked to distribute them between the computer and paper based test. This was designed to measure their level of preference of one technique over another as apposed to simply asking them their preference.



**Figure 1:** Example of the smarty-o-meter used

They were also asked whether they would prefer to use paper based or computer based tests in school to see if this differed to the results from the smarty-o-meter.

The Europress Key Stage 1 Science software (Figure 2) is designed to test children in preparation for their Key Stage 1 Standard Attainment Tests (SATs). All instructions are given to the children verbally with a repeat option provided so their reading ability should not be an issue. Firstly the children are asked to enter their first name, then their surname, and press enter after each. They are then asked whether they would like to practice or do a test. The main difference between the two modes is that the child receives immediate feedback after each question using the practice mode whereas in test mode they receive feedback at the end and the results are saved. For this evaluation practice mode was used. Once the child has selected practice mode they are taken to the question screen where they must choose which question they would like to answer. The question then appears on screen and is read out to the child. Once the child has answered the question they are required to click on the tick in the bottom right hand corner of the screen. The child must go through this process until they have selected and answered all 10 questions.



**Figure 2:** The interface for the testing software showing two styles of questions.

## Results

The raw scores for the two types of tests were analysed to determine whether there was a correlation and significant difference between the two. There was a strong correlation between the children's performance on the computer based and paper based test ( $\rho = 0.647$ ,  $P=0.002$ ). The mean score for the paper based test was  $M=7.6$   $SD=1.35$  which is not significantly different ( $t=1.674$ ,  $df=19$ ,  $p=0.11$ ) from that of the computer based test  $M=7.0$   $SD=2.10$ .

The children's scores were also compared to determine whether gender differences affected test performance. On the paper based test the mean scores for boys was  $M=7.78$   $SD=1.39$ , which is not significantly different ( $t=0.52$   $df=18$   $p=0.61$ ) to the girls  $M=7.45$   $SD=1.36$ . Similar findings were shown with the computer based tests as there was no significant difference between the two groups, the means were identical for both genders  $M=7.0$ .

Further analysis examined the difference between the children's scores on both the paper based test and the computer based test. It was found that 50% of the children performed better on the paper based test, 25% on the computer based test with the remaining 25% showing no difference.

The results using the Smarty-o-meter indicated a strong preference for computer based testing with the mean score being  $M=3.5$ ,  $SD=1.1$  compared to  $M=1.5$ ,  $SD=1.1$  for paper and overall 70% of the Smarties were allocated to computer based test. The children were also asked which method they would prefer to use at school and 65% chose computer over paper. The performance of the children did not appear to influence their preferred test method as 70% of the children who performed better on paper stipulated they would rather use the computer.

The distribution of Smarties between the two methods were analysed to see if there was any gender difference. It was found that both genders had a similar preference towards computer based testing with the boys allocating 76% and the girls 66% of their Smarties to this test method.

## Conclusions

The results show that overall the majority of children performed better on paper than computer although there was no significant difference. There are numerous factors that may have attributed to this finding for example it could be a consequence of using a small number of questions, the children guessing the answers or the children having no prior experience of using software. However, this did not appear to affect their preference for the computer based test.

Overall it would appear that children as young as 7 have a preference for using computers in their assessment. This would go some way in supporting the introduction of computers assisted assessment into both primary and secondary schools within the UK.

Gender differences were also analysed and it was found that this was not an issue in both test performance and preference of delivery mode. However, a larger sample would be necessary to corroborate these initial findings.

This research study has shown that children as young as 7 would appear to find computer assisted assessment an acceptable technique to be integrated into the assessment strategy within schools.

## Further Research

The study will be carried out with a group of children from another primary school to see whether the results are replicated. Further analysis will be conducted to determine the reliability of the questions and any usability problems associated with the software.

## References

- Ashton, H. S., Schofield, D. K., & Woodger, S. C. (2003). *Pilot Summative Web Assessment in Secondary Education*. Paper presented at the 7th International Computer Assisted Assessment Conference, Loughborough.
- Bennett, R. E., Goodman, M., Hessinger, J., Kahn, H., Liggett, J., Marshall, G., et al. (1999). Using multimedia in large-scale computer-based testing programs. *Computers in Human Behaviour*, 15(3), 283-294.
- Clariana, R., & Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with test mode effect. *British Journal of Educational Technology*, 33(5), 593-602.
- Cooper, G. (1990). Cognitive load theory as an aid for instructional design. *Australian Journal of Educational Technology*, 6(2), 108-113.

- Dyson, M. C., & Kipping, G. J. (1997). The legibility of screen formats: are three columns better than one? *Computers & Graphics, 21*(6), 703-712.
- Federico, P. A. (1989). *Computer-Based and Paper-Based Measurement of Recognition* (Research Report No. NPRDC-TR-89-7): Navy Personnel Research and Development Centre Report.
- Greenwood, L., Cole, U. M., McBride, F. V., Morrison, H., Cowan, P., & Lee, M. (2000). *Can the same results be obtained using computer-mediated tests as for paper-based tests for national curriculum assessment?* Paper presented at the International Conference on Mathematics / Science Education and Technology, San Diego.
- Kirschner, P. A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction, 12*(1), 1-10.
- Liu, M., Papathanasiou, E., & Hao, Y. (2001). Exploring the use of multimedia examination formats in undergraduate teaching: results from the fielding testing. *Computers in Human Behaviour, 17*(3), 225-248.
- Mayes, D. K., Sims, V. K., & Koonce, J. M. (2001). Comprehension and workload differences for VDT and paper-based reading. *International Journal of Industrial Ergonomics, 28*(6), 367-378.
- McDonald, A. S. (2002). The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessment. *Computers & Education, 39*(4), 299-312.
- Nugent, G. (2003). *On-line multimedia assessment for K-4 students*. Paper presented at the world Conference on Educational Multimedia, Hypermedia and Telecommunications, Hawaii.
- Pommerich. (2004). Developing computerized versions of paper-and-pencil tests: mode effects for passage-based tests. *The journal of technology, learning and assessment, 2*(6), 3-44.
- Russell, M., & Haney, B. (1997). Testing Writing on Computers: An Experiment Comparing Students Performance on Test Conducted via Computer and via Paper-and-Pencil. *Education Policy Analysis Archives, 5*(3), 1-19.
- Schenkman, B., Fukuda, T., & Persson, B. (1999). Glare from monitors measured with subjective scales and eye movements. *Displays, 20*, 11-21.
- Sim, G., Holifield, P., & Brown, M. (2004). Implementation of computer assisted assessment: lessons from the literature. *ALT-J, 12*(3), 215-229.
- Zandvliet, D. (1997). A comparison of Computer-Administered and Written Tests. *Journal of Research on Technology in Education, 29*(4), 423-438.