Assessing the digital mathematics curriculum

Tandi Clausen-May

Department for Research in Assessment and Measurement, National Foundation for Educational Research, Slough, UK

In their literature review of e-assessment Ridgeway, McCusker and Pead note an "emerging gap between classroom practices and the assessment system" (2004, 17-18). This gap threatens to undermine the effective development of ICT in the teaching and learning of mathematics. An examination of currently available on-screen assessments indicates that stand-alone instructional programs, designed to teach a specific set of skills or topics, are relatively well supported by tests composed of constrained item-types which can be computer administered and marked. On the other hand, tool software such as dynamic geometry or computer algebra packages may be neglected in the classroom because their use does not form a focus construct within the current assessment system. In this paper some of the constraints on test development that have led to this situation are explored, and ways in which tool software usage might be incorporated into an effective mathematics assessment are considered.

BSRLM Keywords: assessment; digital; mathematics;

Assessment and change in the mathematics curriculum

In their literature review of e-assessment Ridgway, McCusker and Pead (2004) observe that

> there is a danger of an emerging gap between classroom practices and the assessment system… [In] mathematics and science, the use of graphics calculators, spreadsheets, computer algebra systems (CAS) and modelling software is commonplace (and universal in professional practice). Assessment systems that do not allow access to these tools are requiring students to work in unfamiliar and maladaptive ways. Non-ICT-based assessment can be a drag on curriculum reform, rather than a useful driver. (2004, 17-18)

Other writers have drawn attention to this 'gap between classroom practices and the assessment system' in England (Pimm and Johnston-Wilder 2005, Wright 2005). A similar argument has been put forward internationally, particularly in relation to the influence of assessment on teachers' readiness to accept and use computer algebra systems (Lokar and Lokar 2001; Meagher 2001). So there is a common view that unless assessment changes to support and encourage the use of computers in the classroom learners may not acquire the mathematical skills they will need in the future. There is also a general concern, however, that this may be difficult to achieve, as compromise and a change of focus are needed.

Approaches to the digital assessment of mathematics

Weist has developed a classification of mathematics teaching software which distinguishes between *instructional* software, which is 'designed to teach students skills and concepts', and *tool* software, such as dynamic geometry packages or computer algebra systems, which may be 'used as an aid towards another goal' (2001, 46-47). As Papert put it, in the first case 'the *computer is being used to program* the child', while in the second, '*the child programs the computer*' (Papert 1980, 5). These two types of software make different demands on teachers and learners. Many instructional programs may be used as they stand, sometimes without any direct input from the teacher. Tool software, on the other hand, generally requires all users to build up their experience of the programs before they can use them effectively, and this takes time and commitment.

Just as mathematics *teaching* software can be classified into two general categories, so also two broad types of *assessment* task may be identified – although there is a lot of overlap, with some activities showing some of the characteristics of both. Scalise and Gifford developed a "taxonomy" of computer-based item types, ranging between those with

> fully constrained responses…. which can be far too limiting to tap much of the potential of new information technologies, and fully constructed responses…., which can be a challenge for computers to meaningfully analyze even with today's sophisticated tools. (2006, Abstract)

Thus on the one hand there are items that have one, or at most a limited and definable range, of correct responses, so they can be instantly and automatically marked by the computer. Tasks composed entirely of constrained items of this type are somewhat akin to the simplest type of instructional software program in that they can be used by the learner with little input from the teacher and can provide immediate feedback to both. The most extreme example of this approach would be a test consisting entirely of simple four- or five-option multiple choice items – but the range of item types that a computer can be programmed to mark automatically is much wider than this (Clausen-May 2005). The question shown in Figure 1, for example, from the World Class Tests of mathematics, makes good use of the capacity of the computer to provide interactive graphics to allow learners to explore a situation and search for a solution to a problem, but is none-the-less computer markable. In effect, this is a multiple choice item – but one in which every intersection on the grid is a possible option, giving a total of 144 options.

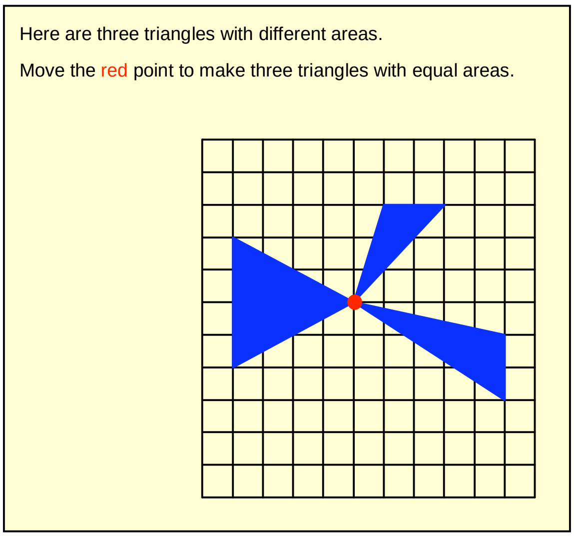

Figure 1 Threlfall and Pool (2004, 8)

Tasks that centre around the use of tool software, on the other hand, are likely to be more open, requiring constructed responses that leave many decisions to the individual learner. An assessment task of this type may involve significant elements of exploration, investigation and problem solving. The outcomes are therefore likely to be different for different learners, not just in terms of their success with different parts of the assessment but in the particular approaches and routes through the problem that they take. While it might be possible to program the computer to recognise and credit any completely correct solution however it was reached (as long as there are only a limited number of 'correct solutions'), it may be more difficult to award partial credit for a range of different responses which show some progress towards a solution but are not fully developed.

Constrained, computer-marked items on the one hand, and more open problems that may be explored using tool software on the other, offer different opportunities but also have different limitations. Teachers may welcome the support that a test composed of the former type of item can provide, offering immediate high quality data relating to the learner's knowledge and understanding of a range of mathematical skills and concepts. However, the restrictions imposed by the need for a closed set of possible responses may make this type of task less suitable for the assessment of such problem-solving skills as representing, analysing, interpreting and evaluating, and communicating and reflecting (Qualifications and Curriculum Authority 2007). These mathematical skills are highly valued, at least in the rhetoric of the school mathematics curriculum, but their assessment may demand a greater degree of flexibility than that offered by computer-marked test items.

Assessing mathematics with tool software

There is currently an explosion of computer-marked mathematics tests composed of items based in an *instructional* software mode. These range from simple multiple choice questions at one extreme to some of the very ambitious dynamic interactive items found in the World Class Tests at the other. The former can usually be created in a set format which makes them relatively cheap to develop, while the latter require individual programming for each item so they are likely to be significantly more expensive. Both, however, allow a closed set of possible responses, so they are computer-markable.

In contrast to these computer-marked tests, there is a dearth of well-constructed materials to support teachers in their assessment of learners' investigative and problem solving activities using commonly available *tool* software. This lack could relate both to pedagogical and to economic factors.

Threlfall and Pool observed how learners demonstrated aspects of ICT-specific learning in their approach to some of the dynamic interactive questions in the World Class Tests. The learners "played" with the mathematics in a way which differed significantly from their response to more closed, paper-based activities (2004, 15). So, the authors argue,

> success on some kinds of computer-based assessment items can arise from different skills and abilities than those required for success in related paper and pencil items. (2004, 11)

Threlfall and Pool's description of the exploratory approach taken by learners working on some questions in the World Class Tests of mathematics could offer a route to the assessment of mathematical reasoning and problem-solving (Threlfall and Pool, 2004; Threlfall et al, 2007). However, the skills that learners employ when they explore a piece of mathematics using an interactive program may not be accepted as valid constructs of a mathematics assessment.

> For the questions that do offer exploratory potential, there is… the issue of whether the qualities and skills that are used to answer them are felt to be legitimate criteria for the assessment being considered. (2004, 14)

So, for example, the use of trial and improvement in the solution of simultaneous equations is generally frowned upon in the classroom, and may lead to a loss of marks in a formal test or examination. However, as the authors observe,

> it is… possible that notions of legitimacy will change, and that the use of exploratory approaches and intuitive informal understandings will become accepted as desirable competences, and therefore valid criteria in assessment. In the …. context of World Class Tests, for example, they are seen as qualities that able mathematicians use to resolve challenging problems. (2004, 15)

Not everyone, however, will agree that this change in 'notions of legitimacy' is acceptable. In a discussion of a different selection of questions, taken from the World Class Tests of problem solving, Ridgway and McCusker (2003) argue that 'Too often, work is characterised by guessing, rather than by a systematic attack on the problem' (2003, 326). This rather critical comment on learners' "guessing" may present another perspective on the "exploratory approaches" reported by Threlfall and Pool. Thus a change in pedagogy, with test constructs that focus on reasoning and problem solving skills, may, or may not, be accepted by mathematics educators.

Furthermore, the development of any piece of tool software involves an enormous commitment of time and effort on the part of the developers – and the educational philosophy that drives this effort may not focus heavily on assessment. Even if it were possible to develop an assessment that could be administered, and possibly marked, within a tool software program, the authors of the software might not be willing to accept the pedagogical shift that this would imply.

On the other hand, economic factors may make test developers unwilling to invest in the development of an assessment that depended upon a software tool over which they would have little control. There would always be a danger that the software could become unavailable, or could evolve in a way that made the assessment unusable. Furthermore, there could be copyright issues if the assessment were to be sold commercially – and designing, programming and trialling a computer-based test is an expensive process, so costs would have to be recouped somehow.

There has thus been little impetus for either the creators of the programs or for established test developers to develop and trial tasks to assess learners' mathematical understanding using tool software. The creators may not see assessment as central to the achievement of their objectives, while test developers might not want to risk an undertaking in which they would have to rely on the continued availability of the software on which the assessment depended. But so long as there is no established way in which learners' mathematical achievement using tool software can be recognised and measured, these tools are likely to remain on the periphery of the mathematics curriculum. If the visions of the developers are to become a reality then there needs to be a credible way to assess learners' achievement with reliable, valid, manageable and markable tests.

One possible approach has been trialled in England to assess learners' ability to use tool software effectively in the context, not of mathematics, but of ICT. This involved the development of a 'walled garden' of programs created specifically for the assessment. This set of tools, which was developed for the English national ICT tests for fourteen-year-olds, includes a word processor, a spreadsheet, a presentation tool and a database. In each case the software was designed to be similar, although not identical, to the tools with which the learners were familiar in the classroom.

However, the provision of a similar set of generic software tools for mathematics could present problems. The range would probably have to include a dynamic geometry package, a spreadsheet, a data handling package, and a computer algebra system. Of these, only the spreadsheet has been developed for the ICT tests. Spreadsheets, however, are relatively homogenous, so learners who have had extensive experience with one are not likely to be seriously disadvantaged when they are required to become familiar with another. Tool software that is used to teach most other areas of mathematics, on the other hand, is commonly available in a range of guises. So, for example, *Cabri-Geometre* and *Geometer's Sketchpad* are significantly different (Mackrell 2004; Johnston-Wilder and Pimm 2005), so if a dynamic geometry tool were to be developed specifically for the assessment of mathematics then this would have to bridge the divide between these programs, and also give access to learners who had experience of other tools such as *GeoGebra*.

Furthermore, there could be a danger that teachers might 'teach to the test' by limiting learners' access in the classroom to the software that was available in the assessments. This would prevent learners from becoming familiar with the wider range of tools used in more advanced educational and in commercial and professional settings. This issue has already been raised in relation to the national ICT tests in England. For example, the BBC reported the complaint of one experienced head of a secondary school ICT department that the tests assessed

> "the old spreadsheet, word-processing and presentation applications which I was teaching 15 years ago, plus a bit of e-mail"…But to recoup the cost of developing the test the National Assessment Agency had said that it would continue "in its current form" until 2013. (BBC, 2007)

So a mathematics assessment based on a walled garden of software tools could be quite damaging in the long run. However, a commercial developer might see it as an attractive proposition, especially if, as with the ICT suite, the tools were associated with a high stakes national assessment that would guarantee their take-up by schools.

Moving forward

While there is currently a rapidly growing range of computer-marked, restricted-response tests, both economic and pedagogical factors have tended to discourage the development of tool-based assessments. However, there could be another approach to the assessment of tool-based mathematics. Assessments that start with the mathematics rather than with the programs might be developed, with learners selecting any available tool software for their completion. Each task would require a robust mark scheme to enable teachers to mark it reliably, providing results that would offer valuable insight into the learners' mathematical understanding using whatever tool software they were familiar with. These assessment tasks should not give rise to problems of copyright as no particular software would be specified.

For example, a simple item requiring learners to use any dynamic geometry tool that permits dragging to construct an irregular trapezium would assess their understanding of its defining properties – such as that two of its four sides must be parallel, but the other two need not be. The item would need to be informally trialled with learners using different tools to establish key points in their reasoning, but it seems probable that a generic mark scheme for this simple task might be developed, offering, say, two marks for a completely correct solution, or one mark for a construction that had some, but not all, of the relevant defining features. More significantly, perhaps, this result would provide a useful insight into the learner's understanding of the properties and construction of an irregular trapezium. This test item would not require any programming on the part of the test developer, but it would have to be marked by the teacher. However, marking would involve a simple check to see whether the construction lost any of its defining properties when its vertices were dragged. Similar questions, involving different constructions, could be designed.

There are a number of obstacles to the development of computer-marked assessments of tool-based mathematics – but written tests have always been marked by teachers. Perhaps our next challenge is to develop a set of mathematics tasks to be carried out using any appropriate tool software, but with marking guidelines that will allow for reasonably reliable, moderatable, teacher marking.

References

BBC. 2007. School computer test scrapped [Online.] Available: http://news.bbc.co.uk/1/hi/education/6232207.stm

Clausen-May, T. 2005. Developing digital – opportunities and dangers in the development of electronic test questions. Paper presented at the IAEA Conference ‘Assessment and the Future of Schooling and Learning’, Abuja, Nigeria, 4-9 September [Online.] Available: http://www.nfer.ac.uk/publications/pdfs/confpapers/IAEA2005tandi.PDF

Johnston-Wilder, S., and D. Pimm, 2005. Some technological tools of the trade. In *Teaching secondary mathematics with ICT*, ed. S. Johnston-Wilder and D. Pimm, 18-42. Milton Keynes: Open University Press

Lokar, M. and M. Lokar. 2001. CAS and the Slovene Final External Examination. *International Journal of Computer Algebra in Mathematics Education* 8, 1: 23-44

Mackrell, K., 2004. Cabri 2+ - a review. [Online.] Available: http://www.atm.org.uk/reviews/software/softwarepix/cabri2.pdf

Meagher, M., 2001. Curriculum assessment in the age of computer algebra systems, *International Journal of Computer Algebra in Mathematics Education,* 8, 1, 89-95.

Papert, S. 1980. *Mindstorms: Children, Computers and Powerful Ideas***.** Brighton: Harvester Press

Pimm, D. and S. Johnston-Wilder. 2005. Technology, mathematics and secondary schools. In *Teaching secondary mathematics with ICT*, ed. S. Johnston-Wilder and D. Pimm, 3-17. Milton Keynes: Open University Press

Qualifications and Curriculum Authority 2007. Mathematics: Programme of Study for Key Stage 3 and Attainment Targets. [Online.] Available: http://curriculum.qca.org.uk/subjects/mathematics/index.aspx

Ridgway, J., S. McCusker and D. Pead. 2004. Literature Review of E-Assessment (Futurelab Series Report 10). Bristol: Futurelab. [Online.] Available: http://www.futurelab.org.uk/resources/documents/lit_reviews/Assessment_Review.pdf

Scalise, K. and B. Gifford. 2006. Computer-based assessment in e-learning: a framework for constructing “intermediate constraint”. Questions and tasks for technology platforms. *Journal of Technology, Learning, and Assessment* 4, 6. [Online.] Available: http://escholarship.bc.edu/cgi/viewcontent.cgi?article=1036&context=jtla

Threlfall, J. and P. Pool, 2004. How Might the assessment of mathematics through dynamic interactive computer items be different from that in conventional tests? [Online.] Available: http://www.icme-organisers.dk/tsg27/papers/03_Threlfall-Pool_fullpaper.pdf

Threlfall, 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* 66: 335-348

Weist, L. R. 2001. The role of computers in mathematics teaching and learning. In: Using Information Technology in Mathematics Education, ed. D.J. Tooke and N. Henderson, 41-55 Binghampton, NY: Haworth Press, Inc.

Wright, D. 2005. Graphical calculators: tools for mathematical thinking. In *Teaching secondary mathematics with ICT,* ed. S. Johnston-Wilder and D. Pimm, 145-158. Milton Keynes: Open University Press