Paper presented at the IAEA conference 2006, Singapore

Losing your inhibitions: possible effects on assessment of dynamic, interactive computer items.

Peter Pool, AEU, School of Education, University of Leeds, Leeds UK LS2 9JT
p.c.pool@education.leeds.ac.uk

AEU would like to express its appreciation of the generous support for this paper given by Trumptech (HK), and in particular the assistance of Mr T M Kwok.

Abstract
This paper examines some of the issues that arise in writing mathematics assessment questions for presentation on a computer screen. It questions the assumption that the computer screen functions simply as a replica of a sheet of paper. It seeks to argue that the medium of presentation and expression is in fact an inherent part of the discourse of mathematics and that any move of assessment from one medium to another changes that discourse. There will be new possibilities for cognitive activity and the loss of others. In this way the assessment medium redefines the subject.

Background
World Class Tests began as a UK government initiative to offer challenging questions in mathematics to very able students ('the Gifted and Talented') at ages 9 and 13. These are in the form of questions on both paper and computer. The assessments are not about seeing how much mathematics has been covered - the questions do not require knowledge of mathematical content beyond normal expectations for students at ages 9 and 13, so acceleration through the curriculum brings no great advantage. The questions are about how deeply the mathematics is understood and they offer success to those who can bring insight, perseverance and flexibility of thought to a question.

Information about current World Class Tests is available at:
http://www.worldclassarena.hk (in the UK go to http://www.worldclassarena.org).

Introduction
This paper uses some of the computer items constructed for World Class Tests as a base from which to examine the effects of the change of medium of presentation on the mathematics curriculum. Not all the examples below are from World Class Tests, but they are indicative of the style. Although few of the computer items in World Class Tests are multiple-choice, it does not follow that they are all highly interactive. Experience has shown that assessment material needs to offer a degree of continuity with students' previous experience and the introduction of novel items needs careful handling to ensure that students are not unsettled by the unexpected. In this paper we shall examine, through a number of examples, how information is presented to students and the characteristics of

interactions available when a computer screen is the medium of presentation. This will then lead to a consideration of the implications of the changes for the assessment of the mathematics curriculum by the use of computers.

The effects of the medium

In a general sense we can describe paper as a restrictive medium of assessment in terms of asking questions and an open one in terms of answering them, whereas the computer screen is open in the way questions can be asked but restrictive in the ways they can be answered.

(i) Characteristics of paper

Paper questions are confined to a mixture of printed word and diagrams; the medium is inert and, in a non-literal sense, inflexible. A restriction for the question setter is that there is no control over how students read the information on a sheet of paper. There are many instances of students having their eyes drawn to a diagram or the question line first and then working back up through the question to find the information at the top of the page. Explanations of context have to be accommodated in a limited number of carefully chosen, plain words. By contrast, the response by a student can be very free – a mixture of text and diagrams, connected by lines and arrows, criss-crossing as the thought takes them. No mark-scheme can entirely circumscribe the variety of responses that students can offer to any particular question.

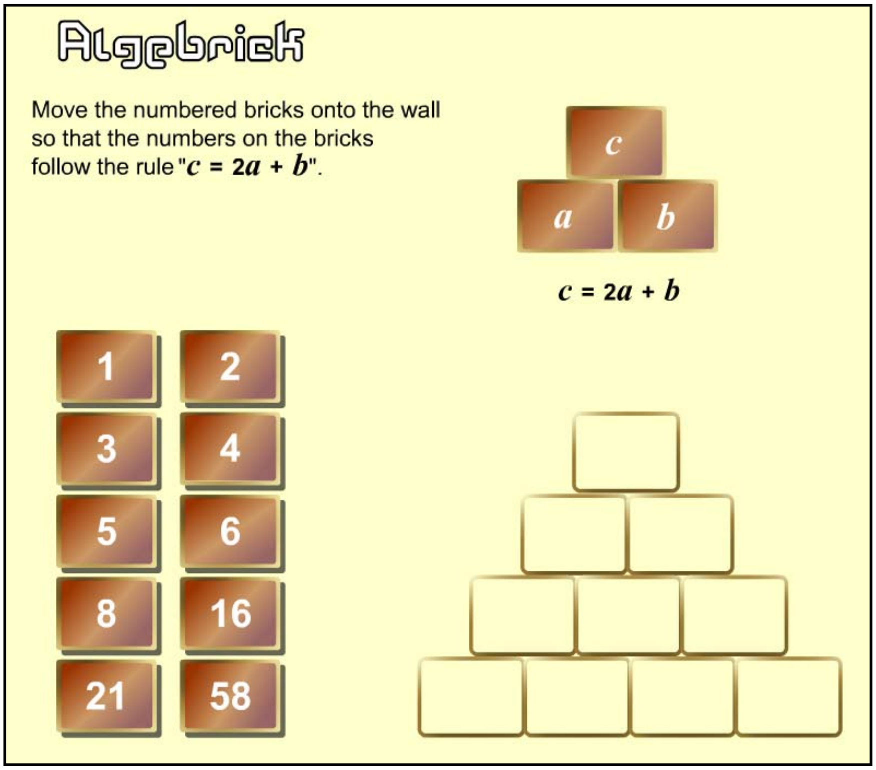
(ii) Characteristics of the computer

A computer can be flexible in its presentation of a question: an animation can communicate what might be inexpressible through text; screen dynamics can offer interactions that are impossible on paper – such as rotating a polyhedron, stretching a polygon, doing a hidden calculation on a number etc. A degree of control can be exercised over the presentation of text by setting the time when, and the way in which, it appears. On the other hand, current software development is such that it tends to be narrow in its range of acceptances of answers – the format of response must be known in advance – an alphanumeric string, a number of mouse clicks, a shape configuration etc. and candidates must express their answers in conformity with such requirements. Paper and computer are not simple alternatives: they are architecturally different.

Examples

Comparisons between computer screen based items and paper ones are inevitable. In particular instances it may be hard to see any difference between a paper question and the 'same' question on a computer screen. This however, is to discount the medium itself. Not all students see the technological change from black and white paper to bright, colourful computer screen in equal terms (Hargreaves, et al. 2004).

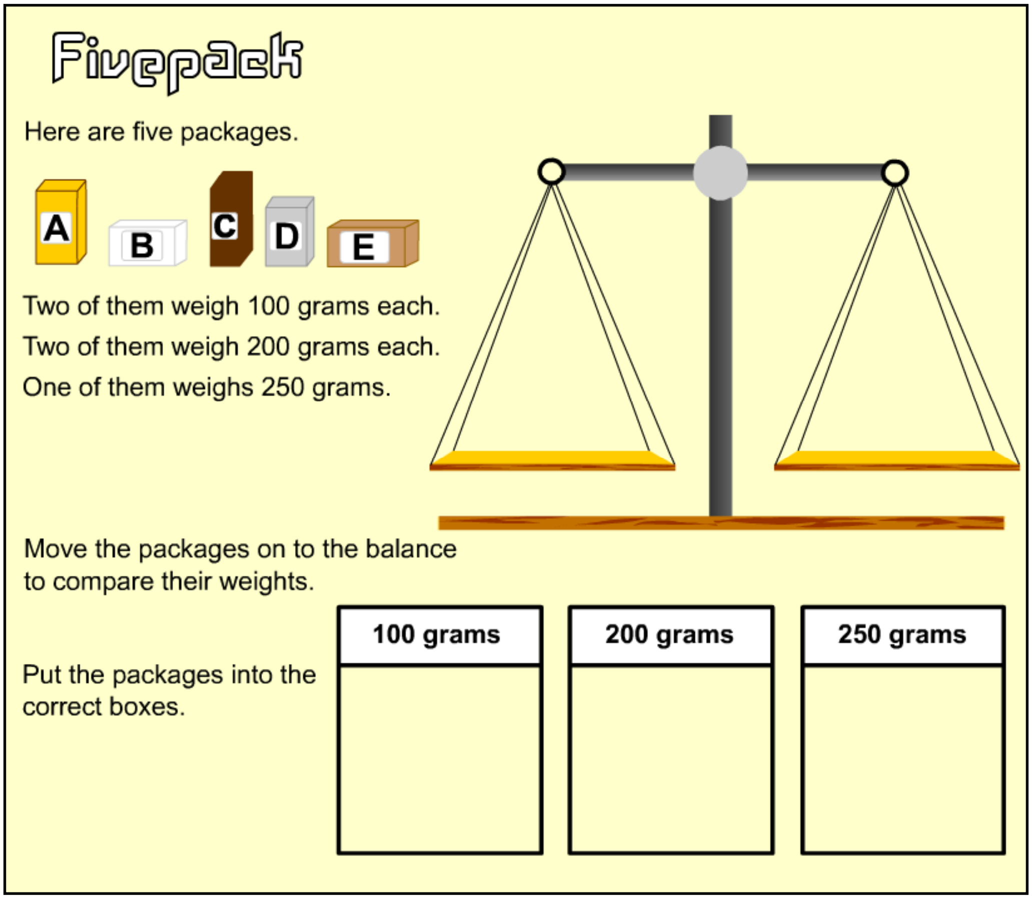
Additionally there are inherent differences between media that may show no apparent effects until particular conditions reveal them. It is possible that even with two apparently 'similar' items the nature of the cognitive activity may be different in the two media and thus the assessment may be of different behaviours, despite the similarity of expressed outcomes. For example, the item **Algebrick**:



On the computer screen (above) number cards can be dragged into (or removed from) the pyramid at will. A reasoned strategy might (correctly) suppose that the smaller numbers must be at the bottom of the pyramid and the larger at the top. There are however increasing possibilities for arrangement as you descend the pyramid. Unless you are very lucky there will need to be some trying out of these possibilities. On paper this is quite tedious, as the diagram continually has to be redrawn or further crowded and obscured with amending annotations. An additional problem on paper is that you must keep track of the given numbers to avoid miscopying / repeating/ omitting any. On screen the trialling is quite quick and easy by means of drag and drop, and miscopying is made impossible by the nature of the interactivity.

A student attempting a paper version of this is likely to resort to a much more reflective, analytical approach to minimise the mechanical effort of copying blocks of numbers (with its attendant risks) and may need to devise a checking procedure to ensure the correct numbers are used. This implies a greater cognitive load on paper compared to the screen version where each arrangement can be tried very conveniently. In either medium we are looking for the same visible outcome and, in a narrow mathematical sense, in both cases, students are solving a question about arranging numbers according to a given rule. But the mathematics cannot be divorced from its medium. In this case the paper version appears to demand considerably more than this particular screen version and so may be considered harder. Neither version is the 'real' question, since each is rooted in its own medium of presentation. This makes any question of which might be the more valid assessment somewhat problematic since they each are assessing different things out of the one mathematical context. Each must be judged on its own merits since the assessment in each case includes the affordances or constraints of the medium. There is no abstract question to be had that is 'pure mathematics'

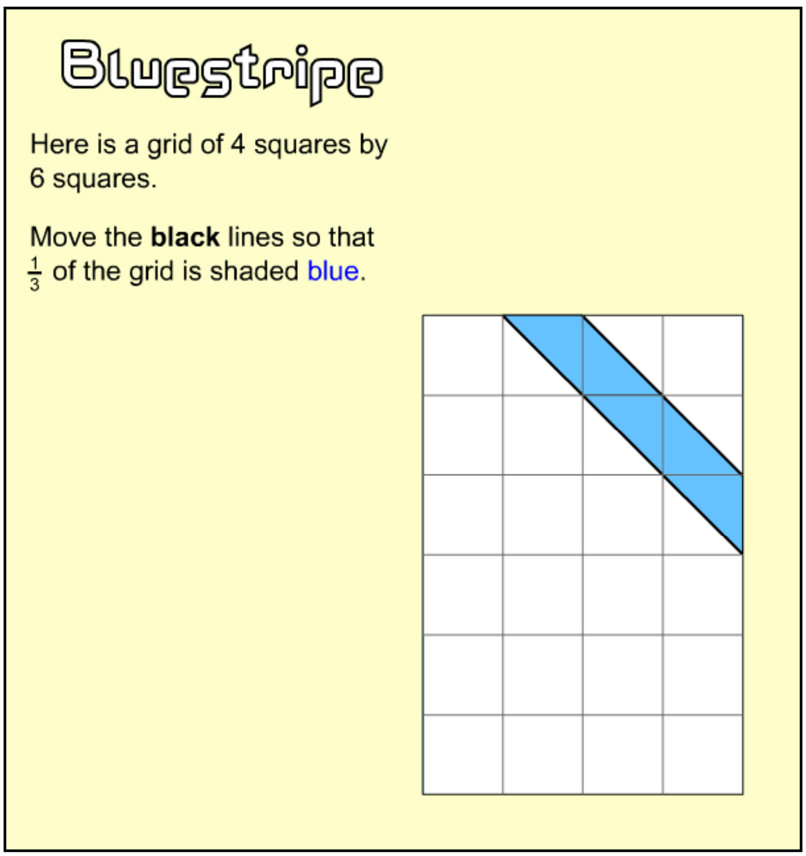
A second example illustrates how information can be presented to students in ways that are quite impossible on paper, and thus it both demands and enables actions that go outside or extend (depending on your point of view) the conventional assessed curriculum. Here is the item **Fivepack**:



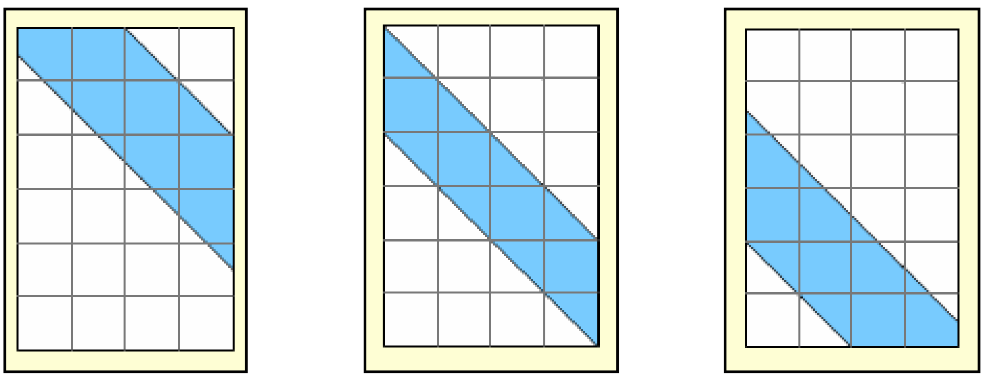
This item offers a controlled simulation of weighing to compare the relative weights of different packages. Packages are dragged onto either side of the balance, which dips or rises to indicate the heavier/ lighter package. Each package can then be dragged into the one of three boxes to indicate its weight according to the given information. It is left to the student to devise a procedure for finding relevant information. Clearly a substantial amount of the kind: 'B is heavier than C', 'A is lighter than E' etc. can be generated (the software, additionally allows more than one package to be placed on either side of the balance). What is required is an ability to make a logical analysis of the information in order to draw conclusions. To this end, it helps to know which information would be worth getting and which would be irrelevant or unnecessary. Too many pieces of information are likely to confuse. The question is an assessment of the ability to be sufficiently analytical in order to obtain necessary information and then organise it.

On paper such a question would need to present the information as prior knowledge (the alternative being to do the example as a practical activity, which of course brings overwhelming organisational problems). This could only be a selection of the possible information that could be obtained and it would necessarily be in some given order. This may make a viable mathematics question, but it is of a distinctly different kind to the screen one above, since it pre-empts decisions about which weighings to make. A further related question that could be asked on paper would be to ask which weighings should be made in order to solve the problem. This turns it into a thought experiment – again an intriguing question to ask but, given the contingency of the weighings, very much harder to think about in the abstract and express coherently than is the case for the computer question as set.

A third example is the item **Bluestripe**. This is a grid of squares with an adjustable shaded band. Each of the two slant edges of the shaded area can be moved parallel to its starting position.



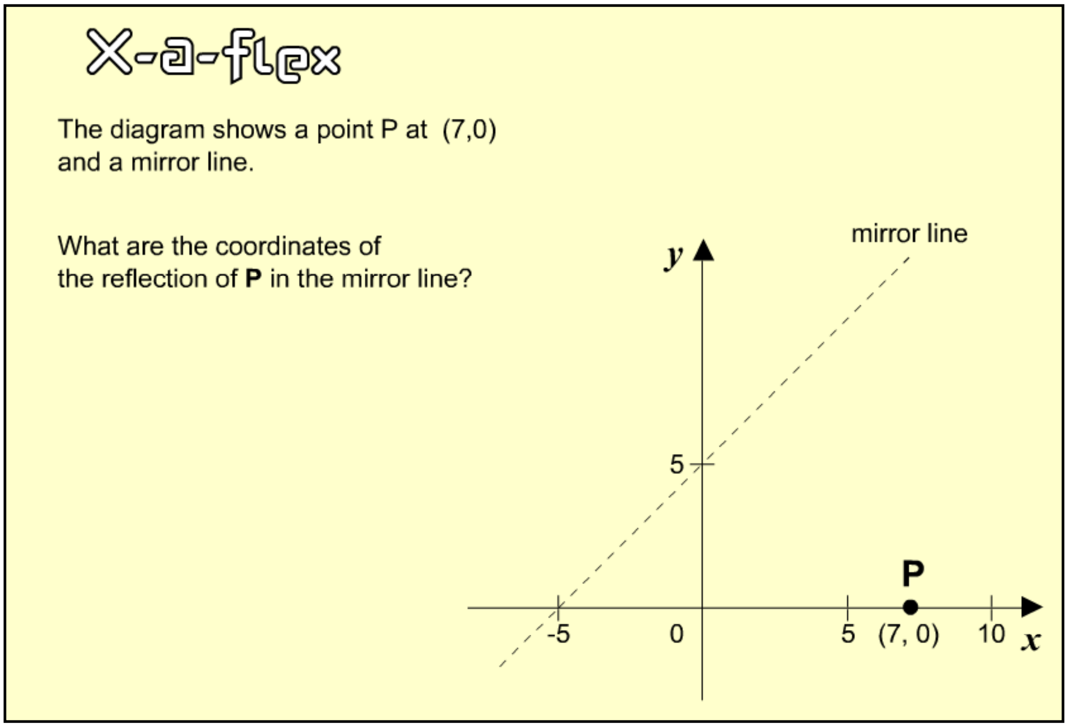
Moving either or both of the parallel sides can cause the shaded area to change its shape from trapezium to pentagon to hexagon to parallelogram. Its area needs to be 8 squares (from the information in the question), but it is difficult to operationalise this fact in any formulaic way over such a wide range of shapes, though whole and half squares are easily countable in individual cases. There are a number of approaches available - from counting squares to using the formula for the area of a parallelogram or trapezium, (knowing that the diagonal of a square is √ 2 x side length). But none of these is likely to be deployed before some exploration has taken place using the interactivity of the diagram. This allows students to see the shapes that are possible, to recognise those that have areas that are easy to calculate, to get a sense of an approximate answer or to notice other aspects that might suggest a way forward. Theoretically, there is an infinite number of possible solutions though most would require precise measurements that are not possible for the student on a computer screen - itself an additional factor for the student to take into account. Three of the more likely solutions are:



In the middle diagram the shaded band can be seen as having four identical vertical parallelogram sections, each one square wide and having a structure of one whole square and two half squares. In the left and right hand diagrams the small shaded part squares can be matched to small white sections to make complete shaded squares.

None of these solutions requires advanced understanding of how to calculate areas of shapes. What is more useful is the ability take advantage of the interactivity to recognise useful features that can be investigated and from which a strategy can be evolved. It is worth noting that it would be almost impossible to ask this question on paper in such a way that a student would be confident she had understood the procedure; the practicalities of then doing the question on paper raise further issues of manageability. In this sense, the mathematics here is ‘new’ in so far as it would not (or could not) be presented in a conventional paper assessed curriculum, though the question itself remains very accessible to anyone who understands conventional mathematics.

A fourth example is **X-a-flex**. This computer item has no dynamic to it (ie no moving parts).



An identical looking question can be presented on paper. Students at age 13 find this a very difficult question and this difficulty tempts many to seek a low level strategy such as sketching a grid in and then locating a point on the other side of the mirror line at the same distance as P, thus the co-ordinates of the image of P are identified. This reveals an understanding of the laws of reflection, but if that were the intention of the question one could provide the grid. In this particular case it is a deeper understanding that is being sought.

On a computer screen sketching a grid is not possible and so a more analytical approach is demanded - perhaps envisaging a line through the point (-5, 0), parallel to the y-axis. The mirror line is at 45°, so the image of P will be along this envisaged line. Since the point P is 12 units along the x-axis from (-5, 0), the reflection of P will be 12 units directly above the point (-5, 0). This will be at (-5, 12).

This is an example of the paper medium allowing methods of solution that undermine the intended mathematics assessment and so bring into question its validity. The reverse is also possible, in a question where endless 'playing' as a result of the dynamic of the screen allows a solution to appear sooner or later; on paper such play is not possible and thoughtful analysis is required.

Discussion

The examples above have all been developed within the framework of a particular software environment with its idiosyncrasies. The particular opportunities or constraints this presents to students in solving mathematics problems cannot necessarily be generalised across the whole medium. Additionally the particular questions that have been presented above and more generally those written for World Class Tests, are of course no more than the product of the imaginations of item writers at a particular time.

It is not the specifics of these particular examples, so much as the general effects they illustrate that are interesting - namely the impossibility of separating the mathematics from the agent and the medium. Following Wertsch (1998), we can say that all action is necessarily mediated: any medium offers (mediational) tools to which the agent turns in order to effect a solution to a mathematics problem. Wertsch offers the example of long multiplication. Someone asked to perform a long multiplication such as 475 x 931 (without a calculator) will instinctively turn to pencil and paper and start to write down rows of digits. Given the mechanical nature of the long multiplication process (the tool the medium offers) we can expect a high success rate. But forbid the use of pencil and paper and this success rate will drop dramatically. In both cases the individual has the same understanding of the multiplication concept as a piece of mathematics, but it is the tool that appears to offer the skill.

This is perhaps illustrated even more vividly when calculators are brought into the discussion. There is a strong body of opinion that would argue that doing a long multiplication on a calculator is no evidence of mathematical skill whatsoever, since 'the calculator has done it for you'. How far is this from the perception that 'the pencil and paper algorithm has done it for you'? And exactly what action might one have to take to claim ' I did it on my own'? Questions about whether the mediational tools invalidate mathematical activity (whether in the classroom or the examination hall) are never far away. Greeno (1998) points out that it is communities of practitioners who share standards that characterise the 'worthwhile' and Watson (2003) adds that it is these standards that constitute the constraints and affordances of a practice. One might note in passing that, historically, mathematics teachers were possibly the last group to buy the now ubiquitous calculators.

Gibson (1979) talks of the affordances that come with mediation: that is to say mediational tools offer empowerment. This is illustrated vividly in **Algebrick** (above) where the facility for trialling possibilities on screen overcomes some of the major obstacles that a paper version would have. But with affordances come constraints: the item **X-a-flex** reveals the restrictions on action when the item appears on a computer screen. The diagram cannot be annotated in any way. In this particular instance the restriction is not seen as negative when considered in terms of desired outcomes. But it is not difficult to imagine questions where this restriction would have a negative effect, for example in the case of being given a diagram and needing to relate an unknown angle to one that is given, by a series of intermediate steps of equating angles or recognising their complements and so forth.

It follows that new mediational tools have to be recognised as such to be useful. Not all children realise of their own volition that folding a piece of paper can solve some symmetry problems (or know how to use a mirror effectively in symmetry). Since the 1960s the time-honoured algorithms for subtraction and long multiplication have been further developed in the belief that making them more meaningful to users is empowering and means that they can be used more effectively.

One can imagine students, new to computer items in mathematics, being faced with the item **Bluestripe** (above) and having no idea as to how to make use of the fact that the boundary lines can be moved, even though they are aware that the solution must come through such action. One might be tempted to see this as evidence of the level of mathematical ability, though this would be contentious. Such judgements take much for granted. What is the relationship between perceived mathematical ability and a knowledge of the mediational tools available? Well, it depends on what you mean by… Some students develop their own mediational tools to meet their needs, but the majority are taught things such as the multiplication algorithm well in advance of any genuine practical need. At some point it may well be decided that mastery of such tools is a sign of competence.

Watson(2003) (quoting Greeno) makes the point:

> *If learning is "improved participation in interactive systems – becoming better attuned to constraints and affordances of activity" then to understand the learning of conventional school mathematics we need to look in detail at the constraints and affordances of mathematical sense-making.*

In this way, we see that the advent of a new medium has significant implications for classroom teaching in order that new tools initially can be mastered and subsequently appropriated (in the sense of becoming a 'natural' resort) by the individual. This process has been illustrated in the past by the advent of the electronic calculator. The use of 'trial and improvement' as a means of solving certain kinds of equation was dramatically eased when the power of hand-held calculators became widely available. So well has this particular mediational tool been appropriated in some cases, that mathematics educators regret its introduction: students see trial and improvement as a near universal tool for solving some of the simplest problems. In other words the mathematical pliers of trial and improvement are used as a rough and ready tool on every nut in sight when use of the correct size spanner would be better.

Popular sentiment may have it that the computer screen is the inevitable replacement for paper. Even if future developments are not quite as straightforward as this and paper is seen to offer, uniquely, the possibility of some valuable mathematical experiences, the dynamic of the screen is seductive and appears to open up new worlds for us – the nearest thing to flying for the hitherto earthbound. If that is the case then we need to be aware of both the affordances and the constraints that come with any medium and that in changing medium we change the mediational tools that are an inseperable part of the practice of mathematics: in that way we are redefining the subject.

References

Gibson, J. J. (1979) *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Greeno, J.: (1994) 'Gibson's affordances'. Psychological Review 101, 336-342.

Hargreaves, M., Shorrocks-Taylor, D., Swinnerton, B., Tait, K, and Threlfall, J. (2004) Computer or paper? That is the question: Does the medium in which assessment questions are presented affect children's performance in mathematics? *Educational Research*, 46(1), pp.29-42

Watson, A. (2003) Affordances, constraints and attunements in mathematical activity *Proceedings of the British Society for Research into Learning Mathematics* 23(2)

Wertsch, J.V. (1998) *Mind as Action*. New York: Oxford University Press.