Goldsmiths

UNIVERSITY OF LONDON

e-scape portfolio assessment
phase 2 report

January 2007

e-scape portfolio assessment

a research & development project for the Department for Education & Skills (DfES)
and the Qualifications and Curriculum Authority (QCA)

phase 2 report

by the Technology Education Research Unit
Goldsmiths College: University of London

Richard Kimbell
Tony Wheeler
Soo Miller
Alistair Pollitt

January 2007

Acknowledgements

Contents:

executive summary

In 2003, the Technology Education Research Unit at Goldsmiths University of London was asked to undertake research to examine the extent to which - and the ways in which - innovation and teamwork might be more fully recognised and rewarded in assessment processes, particularly within GCSE. The project '*assessing design innovation*' was launched in Jan 2003 and concluded in Dec 2004.

The principal outcome of that project was a developed portfolio assessment system that sat somewhere between a formal examination and a piece of coursework. It was designed to operate in 6 hours - typically 2 mornings - and presented learners with a design task that was to be taken through to a prototype. The outcomes of learners' work during this project were most encouraging. It was possible to demonstrate that different levels of innovation were identifiable in the work and that the best work was highly innovative. Critically, the consensus of teachers and learners was that the portfolio system acted as a dynamic force to drive the activity forward with pace and purpose. The data from the trials of this system is fully reported in the project report (Kimbell et al 2004).

Alongside this development, it is important to note a number of parallel strands of influence that combined to create 'project e-scape'.

Assessment for learning has become a major concern of educators. It places the teacher (rather than any external body) at the heart of the assessment process and presents teachers with large amounts of personalised-learning information to manage. Within this emerging field, we see much value in exploring the use of digital systems to support teachers and learners.

In this digital context***, e-learning*** is a term that has emerged to describe a wide range of digitally enhanced educational experiences; from a straightforward internet search or the completion of a simple screen-based multiple choice question, to full blown multimedia managed learning environments providing access to complete courses. The DfES ***e-learning strategy*** identifies the provision of a centralised e-portfolio as an important priority for reform, second only to the provision of the infrastructure to make it work.

In the context of design & technology alone, **Awarding Bodies** are responsible for the assessment of approx half a million students annually using portfolios in which learners develop a design solution to a task of their own choosing, simultaneously telling the story of their development process. Approx 50% of learners' GCSE marks are allocated on the basis of the quality of their portfolio. The Awarding Bodies responsible for these assessments – particularly at GCSE – are increasingly seeking to exploit the power of digital technologies.

This combination of influences led us at TERU to develop a proposal to QCA/DfES for a digital approach to portfolio assessment. Learning activities in design & technology studios and workshops are increasingly influenced by digital technology, and we believe that the portfolio assessment system that we developed in the DfES "Assessing Design Innovation" project provides a useful model to explore the possibilities of extending digital working in design & technology into **digital assessmen**t of learners' performance. Project e-scape (e-solutions for creative assessment in portfolio environments) was established as a result of discussions with DfES, QCA and Awarding Bodies.

e-scape phase 1 'proof of concept'

Phase 1 of project e-scape ran from January to October 2005 and was designed as a 'proof of concept' to see whether the kinds of technologies that existed at that time could be harnessed towards the creation of effective e-assessment portfolios.

The brief for phase 1 of project e-scape
"QCA intends now to initiate the development of an innovative portfolio-based (or extended task) approach to assessing Design & Technology at GCSE. This will use digital technology extensively, both to capture the student's work and for grading purposes. The purpose of Phase I is to evaluate the feasibility of the approach...'
(QCA Specification June 2004)

The *proof of concept* operated at four levels; technological, pedagogic, manageable, and functional. Each of these four 'proof of concept' deliverables was explored in schools through a series of small-scale trials. We explored the system from both ends. At the classroom activity end, pedagogic priorities and the need to evidence capability dominated our concerns. But at the assessment end we were required to explore the manageability and functionality of an e-portfolio for assessment purposes.

Design studios and workshops are typically not full of desktop or laptop computers – which are often down the corridor in a different room – often an IT suite. Since we were concerned to explore the **dynamic creation** of an e-portfolio (rather than a sanitised one created through 2nd hand re-telling of the real story) we chose to use peripheral digital tools (e.g. digital pens, cameras, PDAs) that were capable of being used in the workshop/studio setting.

Specifically, the activity we were seeking to enhance was the 6-hour 'light fantastic' activity developed for the *assessing design innovation* project. This activity was capable of subdivision into a series of component parts, and – for the purposes of exploration with digital peripherals – we divided the activity into a series of 'work-parcels' – some focussed on supporting learners' designing and some on supporting teachers' assessment. We undertook a series of school-based trials between Jan and May 2005, with learners from year 6 to year 12. The second area of work concerned the *technical systems* that would need to be in place for the learners to be able to develop their solution to the task in a *web-space* - accessible to the learners themselves, and their teachers, and (ultimately) to examination board assessors.

The outcome of this phase 1 proof of concept was a body of digital work from learners and evidence from the associated e-assessment trials of that work. DfES, QCA and the Awarding Bodies were persuaded of the concept, and we were invited to take the project to the next stage.

e-scape phase 2 working prototype

Phase 2 of project e-scape (November 2005 – Jan 2007) has involved the creation of a working prototype system. The system was to facilitate the designing activity in the classroom – using peripheral digital tools – and to allow all learners' work to be tracked and logged in a website for subsequent assessment by Awarding Bodies. The system was to be tested through a national pilot (principally with yr 10 learners) in the summer term of 2006.

Developing the system

The prototype was designed in association with two technology partners Handheld Learning and TAG Learning. HHL are specialists in the use of peripheral digital tools and specifically PDAs. Phase 1 of the project had shown that PDAs would be good tools to focus on, principally because of their multi-functionality – for capturing drawings, photographs, writing and speech. TAG Learning have a strong track record in web-based portfolio creation for schools. For the phase 2 prototype we brought these two partners together and invited them to devise a system that would allow the handheld technology to 'speak' directly to the website, enabling us to track – in real time – the evolution of learners' portfolios in design workshops in schools.

Whilst the technology partners were working on the system, in TERU we worked on the classroom activity and the protocols that would render it manageable as a design activity worked largely through digital means. Whilst we worked from the 6 hr activity structure of the paper-based version in *assessing design innovation*, we modified it substantially to capitalise on the potential that is made available through the PDA. The drawing and writing

tools can replicate paper based drawing and writing, and the camera removed the need for a separate one that we had previously used. But the speech tool was entirely new. For the 1st time we could capture the authentic voice of the learner at points through the activity. And throughout, we retained the importance of real materials – as

learners struggle to create their own prototype solutions to the design task.
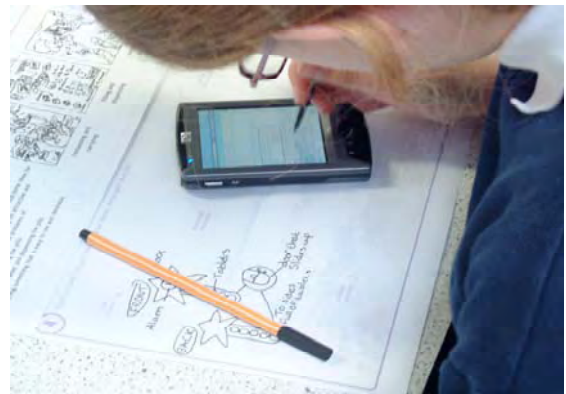
By May 2006 we had a working e-scape portfolio system, and all the resources needed to make is operate as an assessment system in schools. We launched the national pilot in 14 schools across the country through June and July 2006 and as a result accumulated 250 e-portfolios of year 10 learners' performance in the website. The system worked remarkably smoothly and we are grateful for all the support and enthusiasm from teachers and learners.

Making assessments

Assessing web-based portfolios can be done using the same systems as are conventionally used for assessing paper-based work, by allocating scores for elements within a rubric. But having all the work in a website opens the possibility of using a quite different approach. In association with assessment specialists we used an approach of 'comparative pairs' judgements that was developed originally in the 1920s. Essentially this involves a judge looking at two portfolios and deciding (only) which – on the basis of agreed criteria - is the

better / stronger piece of work. Then looking at another pair, and another pair, and another pair. Many paired judgements are made, enabling each piece of work to be compared to many other pieces, and the process is undertaken by a team of judges so that each piece is seen by many different judges (see sect 2.12). The combined effect of all this is two-fold.

First, using Rasch analysis, the mass of wins and losses for individual portfolios can be transformed into a rank order of all the portfolios. Those at the top are those that win every comparison and those at the bottom have lost every time. In the middle are those that win half the time and lose half the time.
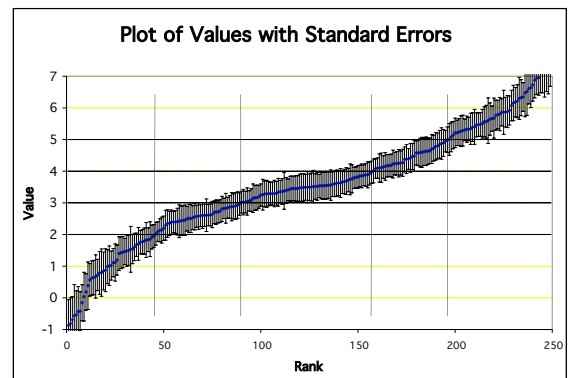
Second, since in our case each portfolio was judged at least 17 times (sometimes significantly more) and by 7 judges, the judging process renders highly reliable results. The standard error attaching to the placement of individuals within the rank order is significantly lower than would be the case in conventional portfolio assessment. (See sect 2.13)

The judging process (including training of judges and 3 round of judging) was undertaken in September and October 2006 and the resulting data used for analysis of learners' performance. Whilst the pilot was principally designed to test the system itself, it was necessary to test learners within the system and the resulting data has proved interesting (see sect 2.17) concerning for example the relationships in performance between designing on paper and designing digitally; between girls and boys performance; and in relation to different 'general ability' groups. Overall however, we were concerned to gauge teachers' and learners' reaction to the activity system in schools (see sect 2.7) and the judges' reaction to the assessment system (see sect 2.14). The power of this dynamic real-time e-portfolio system is best captured in section 2.16, where we illustrate performance at various levels using the data (drawings, photos, notes, and sound files) direct from the website.

Conclusions

The successful conclusion of phase 2 of project e-scape raises many issues of importance for the future of e-learning and e-assessment. We summarise these below in relation to the four categories of research question that launched project e-scape.

Concerning **technological** challenges;
The key point is that the whole system is driven by a remote server dynamically sending and receiving data to and from hand-held digital tools, putting the teacher in control of the sequences of the task and automatically building an evidence trail in the web portfolio.

Concerning **pedagogic** challenges;
The key point is that everything we did for the purposes of collecting evidence for assessment **also** helped to scaffold the progress of the activity and the performance of learners.

Concerning the **manageability** challenges;

The key point is the infusion of technology into activity. Real-time activity in studios, workshops, playing fields, theatres, science labs and the like, is typically not aligned with digital power. That power typically sits in splendid isolation in the shimmering purity of IT suites. In e-scape we have shown how the technology can get down and dirty and unleash its digital power where it is really needed. And in the national pilot we demonstrated that it was manageable.

Finally, concerning the **functionality** of the assessment system;
The key point is that performance assessment is notoriously difficult, and at both ends of the system. It is difficult to manage the performance itself in ways that assure equity to all learners and it is difficult to ensure reliability in the assessment. Within e-scape we have created a system that operates right across this terrain. Learners and teachers report that it worked well at the performance end, and the data shows that it produced reliable statistics at the assessment end. The prototype has done enough to demonstrate that it is a functional system for assessment.

Next steps

The two major innovations in e-scape have been (i) to create a system in which hand-held digital tools link dynamically to a website to create portfolios, and (ii) enabling 'comparative-pairs' judging for reliable assessment. These two innovations are at the centre of a new proposal that is currently being negotiated between DfES, QCA, Becta, Awarding Bodies and TERU. In this next step we propose to move forward from the prototype system, creating an authoring tool that allows teachers to create many different kinds of structured coursework activity – with steps of their choosing and timings for them as best fits their school. These activities might be in design & technology, but equally might be in geography or some other curriculum area. At the same time we propose to work alongside Awarding Bodies to explore how such real time e-portfolios can be seamlessly integrated into the data management processes that lead to the award of qualifications – and specifically GCSE at age 16.

Discussions are advanced for this new project – *e-scape phase 3* – to run from Feb 2007 to March 2009, by which time a scalable national system will be operational.

e-scape phase 1

1.1 Context

The story that underpins this project brings together a number of strands of educational debate:

a. design & technology
b. assessment for learning
c. e-learning
d. e-learning in design & technology
e. portfolios - what they are and what they aren't
f. e-assessment & Awarding Body innovation

a) design & technology

In 1999, the latest version the National Curriculum was published, including the most recent formulation for design & technology. One of the welcome additions to each of the subject areas for NC2000 was the articulation of 'importance' statements, in which the vision of subjects is encapsulated. The Statement for design & technology reads as follows:

The importance of design and technology

Design and technology prepares learners to participate in tomorrow's rapidly changing technologies. They learn to think and intervene creatively to improve quality of life. The subject calls for learners to become autonomous and creative problem solvers, as individuals and members of a team. They must look for needs, wants and opportunities and respond to them by developing a range of ideas and making products and systems. They combine practical skills with an understanding of aesthetics, social and environmental issues, function and industrial practices. As they do so, they reflect on and evaluate present and past design and technology, its uses and effects. Through design and technology, all learners can become discriminating and informed users of products, and become innovators.
(DfEE 1999)

At the time of publication, the DfEE, in concert with the Design & Technology Association (DATA) established a Strategy Group for design & technology, charged with the task of steering the subject through the following years. The group undertook a number of development tasks, including an externally commissioned review of the literature concerning the impact of Design & Technology and a review of new technologies that might be encouraged to support the growth of design & technology in the immediate future. One task - undertaken by members of the group itself - was to review the internal coherence of design & technology as presented in NC2000, with particular regard to the match between the vision statement, the Programmes of Study (PoS) and the Attainment Target (AT).

It was noted that the vision statement encapsulated the need for creativity, innovation and teamwork in design & technology.
- 'intervene creatively'
- 'creative problem solvers'
- 'members of a team'
- 'become innovators'

It was also noted that whilst the PoS are less clear on these points, there is at least an implicit recognition of their importance and the scope or flexibility to interpret these imperatives into school curricula. However it was noted that the Attainment Target is starkly bereft of any reference to, or recognition of, these key factors.

Beyond NC requirements, related problems were evident with GCSE assessments, partly through the syllabus specifications themselves (which lack reference to innovation, creativity and teamwork), and partly, inadvertently, through the impact of 'league-tables'. Teachers, departments and schools are now almost as dependent upon the GCSE results as are their learners, and a typical response in schools is that teachers impose ever-more rigid formulas on learner project portfolios to guarantee success. The concern of the DfES Strategy Group was that as GCSE project work portfolios become more formulaic, innovative learners may be penalised by comparison with well organised, rule-following learners. This has had the result that - in relation to the design & technology vision statement - the wrong learners (or at least some of the wrong learners) are rewarded with the best grades in GCSE assessments.

b) assessment for learning

Whilst assessment takes many forms and has many purposes (e.g. summative, diagnostic, evaluative, formative), a particular focus has recently been placed on formative assessment. This is not assessment for award and certification purposes, but is rather concerned with assessment to improve and enrich the learning environment. The QCA view of Assessment for Learning provides the following checklist for effective practice:

  • *sharing learning goals with learners*
  • *helping learners know and recognise the standards to aim for*
  • *providing feedback that helps learners to identify how to improve*
  • *believing that every learner can improve in comparison with previous achievements*
  • *both the teacher and learners reviewing and reflecting on learners' performance and progress*
  • *learners learning self-assessment techniques to discover areas they need to improve*
  • *recognising that both motivation and self-esteem, crucial for effective learning and progress, can be increased by effective assessment techniques.*
  (QCA May 2005)

Black et al (2003) informed this debate with the launch of their book "Assessment for Learning: putting it into practice", and the Tomlinson Report "14-19: Extending opportunities, raising standards" took the debate further. The argument from both is essentially that too much time, effort and expense is tied up in external assessments and that more attention should be devoted to the kinds of assessment that are classroom and teacher based, and that are designed to inform the processes of learning and teaching. Assessment arises in almost every exchange between teacher and learner, and operates as a feedback device, informing the teacher of any misunderstandings, or half understandings that stand in the way of learners' progress.

"formative assessment can occur many times in every lesson"
(Black P & Harrison C 2004)

Seen in this way, formative assessment, or *assessment for learning* helps the teacher to shape their next intervention and it casts the debate on assessment into a personalised form – customised towards the needs of individual learners and their progress. Phrases such as 'individualised assessment' and 'learner-centred assessment' thereby become key parts of the lexicon. As OFSTED has pointed out:

"Regular feedback by teachers to learners, with a note of their strengths and weaknesses, significantly enhances learners' progress."
(OFSTED   May 2003)

But achieving these benefits increases substantially the requirement for interaction with learners and presents teachers with a significant increase in the amount of information they have to manage. Within this emerging field, we believe that there is enormous potential in the use of digital systems to support teachers and learners with appropriate tools to manage these rich and complex data. And there is a paradox here that is worth noting. Assessment reform is being led by groups with limited understanding and experience of digital systems, whilst digital developments for the classroom are pressed forward by groups with limited understanding and experience of learning and assessment.

We believe that sympathetically designed digital systems could provide both a framework of support to better understand these processes of assessment as integral to learning, and at the same time provide flexible tools to manage and implement this (largely) new emphasis.

c) e-learning

The present government has embarked on a major programme to digitise many of the activities and services it offers, driven by (among other things) the promise of greater control, improved efficiencies, cost savings and better standards of service. This focus on developing new ICT systems straddles many aspects of government from (e.g.) taxation, registration, legislation, communication, health and education.

These initiatives have developed as largely isolated programmes and we have now reached a point where it has become clear that there is a pressing need to, and significant additional benefits to be gained from, joining these systems up. An obvious common denominator to facilitate a more connected approach is the individual citizen and recent e-government proposals anticipate binding existing systems together through new bridging services such as personalised e-learning systems and e-identity cards. E-learning is a term that has emerged to describe a wide range of digitally enhanced educational experiences; from a straightforward internet search or the completion of a simple screen-based multiple choice question, to full blown multimedia managed learning environments providing access to complete courses.

With the new focus on joining up e-services, e-learning has gained an additional, longitudinal dimension through the proposal to provide "personal online learning spaces". Interestingly, this requirement is identified not just by the DfES but comes as part of an overarching policy direction from the Prime Minster's Strategy Unit. In a document entitled

"Connecting the UK: the Digital Strategy", action1 is defined as "Transforming Learning with ICT" and describes the need for everyone to have an electronic portfolio for lifelong learning:

Over time we should see the technology join up better across institutions, so that this is available to learners to build on wherever they go – to further learning, or to work-based learning. And in the future it will be more than simply a storage space - a digital site that is personalised, that remembers what the learner is interested in and suggests relevant web sites, or alerts them to courses and learning opportunities that fit their needs. We will encourage all organisations to support a personal online learning space for their learners that can develop eventually into an electronic portfolio for lifelong learning. (Prime Minister's Strategy Unit. 2005)

Developing a similar theme, the DfES e-learning strategy identifies the provision of a centralised e-portfolio as an important priority for reform, second only to the provision of the infrastructure to make it work:

Our second priority extends this personalised support to learners, helping with all stages of education and with progression to the next stage. We will encourage every institution to offer a personal online learning space to store coursework, course resources, results, and achievements. We will work towards developing a personal identifier for each learner, so that education organisations can support an individual's progression more effectively. Together, these facilities will become an electronic portfolio, making it simpler for learners to build their record of achievement throughout their lifelong learning. (DFES e-strategy 2005)

It is important to recognise however, that these centralised, regulated developments arise in the somewhat more anarchic and dynamic world in which young people live. Here technology is integrated into a wide range of social, cultural and productive aspects of young people's lives to the point where digital technologies have become a ubiquitous element of learners' experiences outside the classroom. One example of this phenomenon is provided by research conducted by MORI for the Nestlé Social Research Programme into the role of mobile phones in young people's lives.

Access to:	Year 9-10	Post 16 in full time education
Mobile	95%	99%
Internet	85%	95%
e-mail	70%	90%
(Haste H. 2005)		

Moreover it is not just that they have access to the technology, they also use it; with 9 out of 10 texting at least once a day and over 25% taking photos daily.

d) e-learning in design and technology

Developing effective approaches to e-learning (embedding ICT) within curriculum subjects has proved to be a significant challenge, and DfES is currently working on a number of programmes to promote more effective and widespread integration of ICT within subject teaching and learning. Design and technology has shown that this integration is possible, and statistics from the annual DfES survey of ICT in Schools reflect increasing use and positive effects.

Use of ICT in areas of the curriculum		
Secondary	d&t	English
substantial	62%	19%
some	35%	69%
little/none	3%	12%

Positive effect of ICT in areas of the curriculum
Secondary	d&t	English
substantial	*64%*	*24%*
some	*32%*	*63%*
little/none	*4%*	*13%*
(DfES 2003)

The statistics in this DfES survey suggest that design & technology makes the best use of ICT when compared to other secondary subjects, and this is reinforced in the OFSTED report of 2004.

"Secondary design and technology (D&T) departments continue to make widespread and effective use of ICT in their teaching."
(OFSTED 2004)

This report goes on to note the range of ICT related activities that are typical in design & technology

Increasingly, learners are developing competencies in:
• using the internet to carry out investigations
• recording ideas and information using attractive graphics
• simulating and modelling ideas as they develop solutions to problems
• using computers and related machinery to design and make products to high levels of sophistication
• using computers to control systems.
(ibid)

We note however, that this list – pleasing though it might be – tends to place the focus for learners use of ICT in design & technology onto *doing* and *recording* activities; 'to control' 'to simulate' to manufacture'. There is little here that suggests the ICT is being used formatively to generate, initiate, stimulate, and develop learners' ideas. Nor is there much scope in this list for acknowledging any ICT role in relation to learners' reflecting, reviewing, critiquing and evaluating their ideas. These are the designerly, intellectual qualities that lie at the heart of learner portfolios in design & technology.

e) portfolios - what they are and what they aren't

The concept of a ' portfolio' is problematic, arising in part from the fact that the term portfolio means very different things to different people. The potential for different interpretations is increased by the use of portfolios as an assessment tool, and complicated yet further in the context of e-learning, where 'e-portfolio assessment' has become a minefield of misunderstanding and confusion.

As a starting point, we recognise that there are many purposes to which portfolios might be applied. These have been articulated by IMS Global Learning (developing specifications for e-learning environments) in the following terms.

• Assessment portfolios
• Presentation portfolios
• Learning portfolios
• Personal development portfolios
• multiple owner project portfolios
• working portfolio
(IMS Sept 2004)

For the purposes of this project we believe it would be helpful to clarify our understanding of what a portfolio is and how it works in design & technology. Whilst d&t portfolios have been

refined over the years and attuned in particular to the priorities of assessment, nonetheless, the essence of a d&t portfolio involves a mix of what the IMS lists as an *assessment* portfolio, a *learning* portfolio and a *working* portfolio.

Through custom & practice in design & technology it is possible to observe several forms of what a portfolio might be.

i. The most common meanings of 'portfolio' defines it as something akin to a box-file into which the learner (or perhaps the learner's teacher) can place work to demonstrate that certain operations, or skills, or processes have been experienced. Viewed in assessment terms, the learner's portfolio becomes a collection of evidence that is then judged against some rubric to arrive at a mark or a level. A portfolio of this kind is conceived as little more than a **container** for evidence.

Translated into the e-portfolio world, it is possible to conceive of many ways in which the evidence being 'contained' could be enhanced through the application of database or spreadsheet systems, which might even be designed to automate the process of containment, standardising, streamlining and potentially removing the need for human interaction.

ii. A somewhat more sophisticated view of portfolio arises from process-rich areas of the curriculum, where teachers encourage learners to document the story of a developing project or experience. This results in learners **reporting** what they have done at various points in the process.

In this kind of 'presenting' or 'reporting' e-portfolio, it is not unusual for learners to use linear digital presentation technologies - e.g. PowerPoint - to give a blow-by-blow account of where they have been in the project - and how they finally got to the end.

However, whilst these two accounts might be seen as part of the picture, neither of them captures the dynamic capability dimension that informs our view of a design & technology portfolio.

The central problem - in both cases - is that the portfolio construction is conceived as a second-hand activity. First you do the activity - whatever it is - and then afterwards you construct a portfolio that somehow documents it. The portfolio is a backward-looking reflection on the experience.

iii)   A third and far richer view of the concept of the portfolio is evidenced in schools (particularly in design & technology) where teachers have embraced the challenge of linking learning and working concepts of the portfolio to the more commonplace assessment portfolio.

In this rich form, the portfolio is transformed into an entity that is integrated into and grows dynamically with the project - and in the process it shapes and pushes forward the project. The best analogy is neither a container nor a reported story, but is rather a **dialogue**. The designer is having a conversation with him/herself through the medium of the portfolio. So it has ideas that pop up but may appear to go nowhere - and it has good ideas that emerge

from somewhere and grow into part solutions - and it has thoughts arising from others comments and reflections on the ideas. Any of these thoughts and ideas may arise from procedural prompts that are deliberately located in the activity to lubricate the dialogue. Looking in on this form of portfolio is closer to looking inside the head of the learner – revealing more of what they are thinking and feeling, and witnessing the live real-time struggle to resolve the issues that surround and make up the task. Importantly, this dynamic version of the portfolio does not place an unreal post-activity burden on learners to reconstruct a sanitised account of the process. Creative learners are particularly resistant to what they see as such unnecessary and unconnected tasks, and this significantly accounts for their underperformance in portfolio assessments that demand such post-hoc story telling.

But real-time dynamic portfolios are not tidy, nor is it possible to present them in a pre-determined PowerPoint template. It is more like a designers sketchbook - full of notes and jotting, sketches, ideas, thoughts, images, recordings and clippings. These manifestations are not random - but are tuned to the challenge of resolving the task in hand. And the point of the portfolio is that the process of working on it shapes and develops the activity and the emerging solution.

Our three categories of portfolio are somewhat dissimilar to those identified by Ridgway, McCusker and Pead for Nesta Futurelab in their literature review of e-portfolios.

> *There are three distinct uses for portfolios:*
> *• The first is to provide a repository for learner work;*
> *• the second is to provide a stimulus for reflective activity – which might involve reflection by the learner, and critical and creative input from peers and tutors;*
> *• the third is as showcase, which might be selected by the learner to represent their 'best work' (as in an artist's portfolio) or to show that the learner has satisfied some externally defined criteria, as in some teacher accreditation systems (e.g. Schulman 1998). (Nesta-Futurelab 2005)*

Whilst their 1st category is the same as ours, their 3rd seems to be little more than an extension of this – allowing for the repository to contain work selected over time and used – inter alia - for assessment purposes. It is a container with some display potential. Furthermore, whilst their 2nd category contains some elements of dialogue potential, it does not capture the dynamic creative essence of portfolios as we see them operating in design & technology.

These disagreements demonstrate the thorny territory that is conjured-up merely by the use of the term e-portfolio. We are very conscious of this issues and it demonstrates the absolute necessity of being very clear about what is proposed within phase 2 of project e-scape.

f. e-assessment and Awarding Body innovation

In design & technology alone, approx half a million learners are assessed annually using portfolios of the kind we have described as a 'dialogue', with learners developing a design solution to a task of their own choosing, and simultaneously telling the story of their development process. Approx 50% of their GCSE marks are allocated on the basis of the quality of their portfolio.

Awarding Bodies responsible for these assessments – particularly at GCSE – are increasingly seeking to exploit the power of digital technologies. And there are at least two 'drivers' for these initiatives;

Awarding bodies have faced the challenge of learners using commercial software systems (particularly CAD/CAM) as part of their product development work, and increasingly teachers have sought to obtain permission to submit this work digitally. Whatever view one takes of what the portfolio is, it seems logical that if the work is being done digitally, it seems somewhat perverse - and inauthentic - to then print it all out as though the work had been done on paper.

Quite apart from the issue of authenticity, there is a practical issue. Awarding Bodies can see the advantage of submitting such work digitally (e.g. on a disc or via a secure website) simply because of the reduced labour, resource (e.g. paper) and costs (e.g. postage) involved.

As regulator of the activities of Awarding Bodies, the Qualifications and Curriculum Authority (QCA) developed its own strategy for addressing e-assessment. In 2004, QCA's 5-year objectives were as shown below, though it should be noted that, as events have turned out, this has proved an over-optimistic schedule:

> *by 2009:*
> • *all new qualifications should include an option for on-screen assessment*
> • *all awarding bodies should be set up to accept and assess e-portfolios*
> • *most GCSEs, AS and A2 examinations should be available on-screen*
> • *National Curriculum Tests should be available on-screen*
> • *on-demand assessments will begin to be a feature of GCSEs*
> • *10 new qualifications, designed for electronic delivery and assessment, should be developed, accredited and live*

Towards this objective, the following timeline will apply:

> *by 2005*
> • *Field trials successfully completed by awarding bodies in at least two subjects*
> • *75% of basic and key skills tests delivered on-screen*
>
> *by 2006*
> • *A code of practice, plus audit and other regulatory criteria, is developed*
> • *AQA, OCR and Edexcel offer live GCSE exams in two subjects each*
> • *Pilot of at least one qualification, specifically designed for e-assessment*
>
> *by 2007*
> • *10% of GCSE examinations administered on-screen*
>
> *by 2008*
> • *On-demand testing introduced for GCSEs in at least two subjects*
>
> *by 2009*
> • *e-assessment becomes increasingly routine*
> *(QCA 2004)*

Beyond QCA and the Awarding Bodies however, it should be noted that the importance of e-portfolios within this strategy has been underlined by OFSTED in their recommendation concerning the development of ICT in schools. They make clear that at the school level there is a need to:

"develop electronic portfolios of learners' work alongside the use of web- or intranet-based applications that enable assessed work to be easily accessed by teachers, learners and parents"
(OFSTED 2004 [ii])

1.2   Starting points

DfES project 'assessing design innovation'

The problem described in 1(a) above - of learner GCSE portfolios in design & technology becoming formulaic and teacher controlled – may be interpreted in relation to the three kinds of portfolio outlined in 1(e) above. The problem being addressed by the Design & Technology Strategy Group was that assessment pressures – linked to the publication of league tables – have distorted the nature of the d&t portfolio. Essentially - in order to ensure success for learners, teachers have increasingly shifted from the dynamic 'dialogue' notion of portfolio to the more passive 'reporting' form that is easier to control and present neatly. Teachers have felt obliged to control the portfolio to maximise learners' opportunity for getting marks.

The Strategy Group recommended that research be undertaken to examine the extent to which - and the ways in which - innovation and teamwork might be more fully recognised and rewarded in assessment processes, particularly within GCSE. The Technology Education Research Unit (TERU) at Goldsmiths was asked to undertake the work and develop a system of assessment that would measure and reward design innovators. The project was launched in Jan 2003 and concluded in Dec 2004. The thrust of our work arising from this brief has been to reinvigorate a view of portfolio assessment that transforms it back into dynamic dialogue mode.

The principal outcome of the project was a developed portfolio assessment system that sat somewhere between a formal examination and a piece of coursework. It was designed to operate in 6 hours - typically 2 mornings - and presented learners with a design task that was to be taken through to a prototype.

The following structure is characteristic of the activities developed. The task ('light fantastic') centres on re-design of a light-bulb packaging box, so that, once the bulb is taken out for use, the package/box can be transformed into a lighting feature - either by itself or in association with other 'liberated' light-bulb package/boxes.

(i)   read the task to the group and (through brief Q&A) establish what is involved
(ii)   explore a series of 'idea-objects' on an 'inspiration table' and in a handling collection designed to promote ideas about how boxes / packages / containers might transform into other forms and functions.
(iii)   put down first ideas in a designated box in the booklet
(iv)   working in groups of 3, learners swap their booklets and each team-mate adds ideas to the original
(v)   team-mates swap again so that each team member has the ideas of the other two members
(vi)   booklets return to their 'owner' and team members discuss the ideas generated
(vii)   the teacher introduces the modelling/resource kit that can be used throughout the 2 mornings
(viii) learners develop their ideas in the booklet - and/or through modelling with the resources
(ix)   learners stop to reflect on the user of the end product and on the
          context of use, before continuing with development

(x)   at intervals, learners are asked to pause and row a dice - with questions on each face. The questions focus on procedural understanding e.g. "how would you ideas change if you had to make 100?' and learners answer the questions in their booklet

(xi)   photographs are used at approx 1 hr intervals to develop a visual story line to illustrate the evolution of models & prototypes

(xii)   at the end of the 1st morning, learners - and their team members reflect on the strengths and weaknesses of their evolving ideas

(xiii)   the 2nd morning starts with a celebration of the work emerging from day 1. This is based on post-it labels that highlight learners' thoughts about the qualities in their ideas

(xiv)   further prototype development

(xv)   regular hourly photos and pauses for reflective thought on strengths and weaknesses

(xvi)   final team reflections, when (in turn) team members review each others' ideas and progress

(xvii)   individually, learners then 'fast-forward' their idea illustrating what the product will look like when completely finished and set-up in context

(xviii)   learners finally review their work from start to finish.

All the learners' work was structured into an A4 workbook that folded out to become an A2 sheet. The activity was designed to be administered by teachers in ordinary design & technology facilities. The workbooks were carefully designed to unfold throughout the activity, ensuring that learners always had sight of the instructions for the sub task they were currently working on *at the same time* as being able to see the work they had just completed.

The illustrations below show two learner booklets. The photo story lines demonstrate the progress of the ideas from inception to final prototype. In the 1st case it is clear that the strength of this idea emerges predominantly through the medium of 3D modelling. The 2nd case however illustrates a learner who is equally comfortable with developing ideas through drawings and with 3D modelling. In both cases, the booklet allows us to 'see' their very different modes of working in operation.
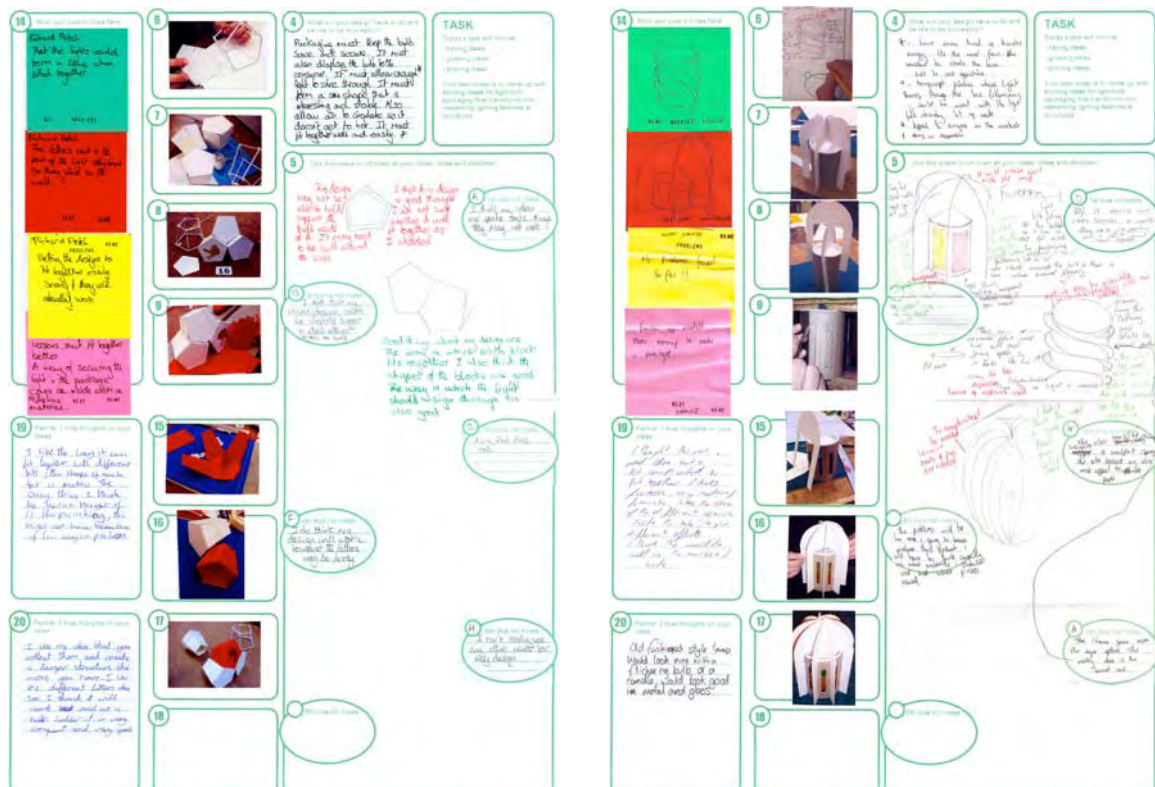
The concept underlying the 1st piece of work was named by the learner 'your name in lights', and was for light-bulb packaging to become pentagonal and tapering, allowing 'used' boxes to build into a spherical lighting feature that – when illuminated by a light at the centre - projected letters around the wall. In this case the learner developed his prototype using a combination of graphic modelling and 3D modelling, supported by a considerable amount of reflective comments and critique.

The outcome of learners' work during this project was most encouraging. It was possible to demonstrate that different levels of innovation were identifiable in the work and that the best work was highly innovative. Critically, the consensus of teachers and learners was that the portfolio system acted as a dynamic force to drive the activity forward with pace and purpose. A second round of trialling was undertaken in association with the four Awarding Bodies for England and Wales. This involved 8 schools and approx 300 learners, all of whom did two activities. The data from this trial is fully reported in the project report (Kimbell et al 2004).

In the process of working on this project, we were able to identify other features of the portfolio – or of the setting within which it works – that significantly impact on its effectiveness. And the key one is the learning and teaching culture created by the teacher in the workshops and studios in which learners operate. This culture in turn influences each of the following features:

• motivation
For learners to be fully engaged and performing at their best requires levels of motivation that – in design & technology at GCSE level – must be maintained over an extended period (typically 6 months). Our 6-hour activity was equally dependent upon generating enthusiasm for the task and we used a number of techniques to generate and maintain it.

• ownership
Who is the portfolio seen to belong to? Is it the learner's, or the teacher's, or the department's, or the GCSE Awarding Body's? Learners' sense of ownership of the work is typically a pre-requisite for fully engaged performance.

• environment
For dynamic creative work to be generated by learners, the environment must be one in which the working atmosphere in conducive to those values. In terms of our project, this required teachers to be open not just to learners' ideas but also very flexible in how they encouraged learners to express and develop them.

• Ideas
At the heart of dynamic creative portfolios are *ideas*. We were explicit in encouraging learners to have ideas, grow their ideas and prove their ideas. Equally we encouraged teachers to facilities these features of learners' performance.

Each of these four will be seen to have an e-equivalent within the e-scape project.

digital enhancement

It was during the development of the activities for this previous project (assessing design innovation) that we became aware of the potential for digital enrichment of the activity. Learners increasingly use digital technologies as part of their work in design & technology. They use digital photography to record their designing and manufacturing processes. They increasingly use the internet for information searches; computer aided design (CAD) systems for design development work; and - in some cases - this extends to computer aided manufacturing (CAM). Also, they increasingly access, complete and store their work on school networks and intranets that allow access from their home computers. This extends the working environment beyond school workshops and studios and allows them time-unlimited access to their work. It also broadens the tool set that is available to them to envision, manipulate and develop their ideas, and in the process it raises important cultural issues associated with the origins of ideas, the ownership of work, team-work and plagiarism.

These thoughts led us to develop a proposal to QCA/DfES for a digital approach to portfolio assessment. Learning activities in design & technology studios and workshops are increasingly influenced by digital technology, and the portfolio assessment system that we had developed in the previous project *"Assessing Design Innovation"* provided a useful model to explore the possibilities of extending digital working in design & technology into digital assessment of learners' performance.

This development involved introducing new technologies into the classroom, as well as extending the range of existing technologies into the domain of assessment. The expanded use of these digital technologies into the realm of assessment will have some serious impacts on current approaches to teaching and learning. We are absolutely committed to undertaking these developments without compromise to the underlying concepts of design & technology as expressed in the 'importance of design & technology' statement in Curriculum 2000. Indeed we believe that the work may contribute to taking forward our collective understanding of the power of design & technology as a learning vehicle.

'peripheral' digital technologies

One of the problems surrounding the use of digital technologies in schools is that teachers tend towards the assumption that this needs to take place in a computer suite, rich in desktop or laptop machines where learners work with a keyboard and screen.

Our starting point is very different.

We start from assumptions about the nature of design & technology – the circumstances of which are almost always workshops and studios. Two of the constants of these typical design & technology spaces are that
• they are full of materials, apparatus, machinery, and specialist work-spaces
• they are associated with the detritus of manufacturing

They therefore make challenging locations for computers, keyboards and screens. First there is not enough space; second the space is not clean (glue, paint, flour & water, sawdust) and third learners themselves get oily or painty or gluey or floury fingers that are not then ideally suited to keyboard use.

For all these reasons we do not believe that digital enhancement of the designing activity will involve computers, keyboards and screens. At least we do not believe that these tools will be at the leading edge of activity. Rather we think that peripheral, back-pocket technologies will be more appropriate: mini digital cameras, digital pens, digital PDAs.

At least at the 'input' level these technologies enable activities in workshops and studios to go ahead almost as normal. The computing power does not take up too much space and (because they can be pocketed) they are not too sensitive to the clutter of the working space and our trials in schools showed that the use of hand-held technology was indeed manageable and effective.

Interestingly, learners at KS4 now (almost universally) have access to mobile phones, a significant proportion of which have digital cameras as a built-in feature. As the telecoms companies race to differentiate their systems through enhanced features, the current distinction between handheld PDAs and mobile handsets is disappearing as the two previously unconnected technology strands merge. While 'smart' phones, with all the features of a PDA, are currently not marketed to learners, camera phones are becoming more ubiquitous and other 'smart' features will increasingly work their way onto phones for children. This trend will be all the quicker if it is seen (or marketed) as providing valuable tools for learning, thereby justifying additional parental expenditure.

In short, we are witnessing the growth of 3rd generation computing. Mainframe computer technologies of the 1960s and 70s gradually faded with the emergence of 2nd generation 'desktop' computers. These completely transformed our working relationship with computers – providing us with far greater interactivity, apparently unmediated by the programmers whose services had formerly been essential. We could 'drive' our own 2nd generation computers in the 1980s and 90s. As the technologies shrank, the growth of laptop computers particularly in the final decade of the 20th C did not materially change our relationship to computers. They operated merely as slightly (very slightly) more mobile versions of the desktop. But the new 3rd generation of computers is radically different. They are FAR more mobile, are equally powerful, and can now genuinely be regarded as 'back-pocket' computers. As such, they are in the process of transforming – once again - our working relationship with computers. The transition to 3rd generation mobile technologies will be just as dramatic as was the transition from the 1st to the 2nd generation. In the contexts of learning, teaching, curriculum and schools, these transformations will be profound. We believe that the e-scape project will provide us with many insights into the educational implications of this 3rd generation.

1.3   brief for e-scape - phase 1

The brief for phase 1 of project e-scape can be summarised as follows:

> "QCA intends now to initiate the development of an innovative portfolio-based (or extended task) approach to assessing Design & Technology at GCSE. This will use digital technology extensively, both to capture the learner's work and for grading purposes. The purpose of Phase I is to evaluate the feasibility of the approach...'
> (QCA Specification June 2004)

Phase 1 of the project (Nov 04-Jun 05) has been - in several senses - a "proof of concept" phase, to explore the feasibility of the concept outlined above. This proof of concept operates at four levels:

i) technological
Concerning the extent to which existing technologies can be adapted for assessment purposes within the portfolio system as currently designed for the DfES "Assessing Design Innovation" project. This will include the applicability of other international work in this area and of any relevant system standards.

ii) pedagogic
Concerning the extent to which the use (for assessment purposes) of such a system can support and enrich the learning experience of design & technology

iii) manageable
Concerning issues of making such assessments do-able in 'normal' d&t classrooms / studios / workshops
• the training / cpd implications for teachers and schools
• the scalability of the system (including security issues) for national implementation

iv) functional
Concerning the factors that an assessment system based on such technologies needs to address;
• the reliability & validity of assessments in this form
• the comparability of data from such e-assessments in d&t with non e-assessments

Each of these four 'proof of concept' deliverables was explored in schools through a series of small-scale trials. The research report (Kimbell et al 2004) – covering the four 'proof of concept' factors – was the required 'deliverable' for phase 1 of the e-scape project. We summarise in section 1.5 of this report the main findings from phase 1 and then translate those findings into a detailed specification of what a working system might be like. This specification – in section 1.6 of this report – then becomes our working template for developing the prototype in phase 2 of project e-scape.

1.4 e-scape methodology – phase 1

The work for project e-scape was divided - broadly - into two areas of concern. The first was with the ways in which digital technologies might be used *to enhance learners' designing*. This was the priority concern of the research team at the outset, since we were determined to ensure that any digital systems introduced into the designing activity should operate as an *enhancement* to the activity - rather than a distraction or a distortion. Accordingly we worked with schools - some of which had been involved in the 'assessing design innovation' project - and explored a range of technologies with learners. These trials are outlined in section (b) below.

The second area of work concerned the *technical systems* that would need to be in place for the learners to be able to develop their solution to the task in a *webspace* - accessible to

the learners themselves, and their teachers, and (ultimately) to examination board assessors.

e-scape work-parcels

As explained above in 2(c), we had in mind to start our explorations with a range of 'peripheral' digital technologies – typically hand-held – that we might use to enhance the designing activity.

Specifically, the activity we were seeking to enhance was the 6-hour 'light fantastic' activity developed for the *assessing design innovation* project.

This activity was capable of subdivision into a series of component parts, and – for the purposes of exploration with digital peripherals – we divided the activity into the following 'work-parcels'.

(i) to support learners' designing

- contextualising; task setting; handling collection
  (to contextualise and get the activity up-and-running)
- early ideas
  (to express early ideas enriching them with support from design teams)
- design-talk
  (to allow discussion to enrich the designing activity)
- photo story line
  (to record [hourly] the evolution of modelling processes)
- design bot
  (to prompt development through questions & task-related information)
- project genie
  (to connect all the above into a coherent interface)

These work-parcels were developed iteratively. Initially we worked with a new technology - and sometimes with the supplier of a new technology - until we had developed it to the point where we felt it might be useful to support learners' designing. At that point we arranged a school trial - often just so we could see what happened. We were frequently unsure about what learners would do with the products and systems, and we were continually astonished at their ability to assimilate the new technologies and make purposeful use of them. We outline some of these experiences in section (b) below.

The second area of work - to support teachers' assessment – was also developed into a series of work-parcels.

(ii) to support teachers' assessment

- collect & compile files
  (to bring together files from different hard/software systems)
- data transfer and web access
  (to make them accessible in a web-space)
- present and share for assessment
  (to present them as a coherent portfolio output for sharing/assessing)

The challenge here was somewhat different, and therefore our methodology was different. We did not focus these work-parcels towards school trials, in part because schools are just not equipped with the technology systems to do what needs doing. Our approach here was to engage in a series of meetings with leading-edge systems developers – and to a lesser extent Awarding Bodies – to discuss the possibilities for developing systems that might be able to achieve what we increasingly saw as necessary. These discussions are outlined in section (c) below.

school trials   (Jan-May 2005)

A series of trials was undertaken in the following schools and the details of these trials are reported in the e-scape phase 1 report.

Saltash Community College (year 12 AS group)
Leasowes School (year 10 group)
Cabot School (year 12/13 group)
Invicta Grammar School (year 10 group)
Pedmore Primary School (year 5 group)

The focus of these trials was on the following issues:

 - to explore the impact of the technology and associated communications systems on normal working practices in design & technology

• virtual handling collections of images and movie clips was created for the PDA

• a new workbook was created and linked it to the use of digital pens and PDAs.

• beaming files between learners and printers.

• design-talk' to capture conversation between learners about their work.

• photo-story-line, using PDAs as cameras

• peer evaluation by learners to identify the strengths and weaknesses

• whether *year 10 learners* reacted in similar ways to the year 12 group in Saltash
• whether the *larger group* of year 10 learners drew out any additional issues

• to explore a semi-intelligent, natural language, digital assistant.

• to explore the impact with mature users – albeit that they were the youngest learners to try the activity.

In addition to undertaking the formal activity trial with the year 12 AS group at Saltash, we also provided the group with a PDA each from Easter to the end of their AS project work. The purpose of the exercise from our point of view was to see what happened when learners have regular and free access to the technology. The assumption was that they would move beyond regarding them as toys – to be experimented with – and begin to use them more naturally as tools to support their designing. We also encouraged them to explore the value of using them in their other curriculum subjects and in their extra-curricular activities.

Each learner produced a report for us, focusing on their use of the PDA over this extended period.

systems development discussions

Quite apart from our school trials of the e-scape approach in the classroom, we undertook a series of discussions with technology-based companies. The focus of these discussions has been to explore the systems by which 'hand held' technologies in the classroom might be linked (through data transfer) to web-based portfolios and subsequently viewed (remotely) for assessment by Awarding Bodies.

While there is no system at that time that offered the dynamic integration and presentation features we required for this project, there were a number of e-portfolio platforms that provided the core data management systems necessary to drive the system we envisaged.

TAG Learning had significant experience in two of the key technology aspects of this project:
• handheld digital peripherals devices
• web based, contributory, moderated, portfolio assessment systems.

Qinetic had a useful Show-N-talk system, operate the SIMS assessment manager and a system of portable electronic school registration.

Extended systems operated the OneBridge system, and with the Social Services had developed a remote voice to text system based on web-enabled, hand-held (Zire) machines linked to a remote server using DragonTalk.

As part of these developments, we also worked with **Dudley LEA** who operated hand-held systems (Zire) linked to Show-n-talk and OneBridge. and these discussions were helpful particularly in relation to scalable implementation issues. Finally, we had meetings with **Awarding Bodies** to explore with them the technologies that they see as emerging in their systems.

The central feature of our requirements in relation to any system that we might adopt was *connectivity;* the capability to beam data automatically from classroom-based, hand-held technologies into pre-designated web-spaces. We explored several possibilities (e.g. using USB, IR [infra-red], Bluetooth and Wireless systems) and identified our priorities for phase 2

development. In the web-spaces we explored a number of presentation options, including morphing, panoramas, zooming (SimpleViewer, Postcard viewer), galleries (Flickr), and albums (i-photo).

These connectivity and presentation tools needed an additional feasibility study to resolve outstanding technical issues, and we undertook this during the summer of 2005, as an extension to phase 1 of the project.

1.5 phase 1 findings

The work outlined in section 1.4 resulted in a set of findings reported here under the four headings from the brief:
- technological findings
- pedagogic findings
- manageability findings
- functional findings

i) technological findings
Concerning the extent to which existing technologies can be adapted for assessment purposes within the portfolio system as currently designed for the DfES "Assessing Design Innovation" project. This will include the applicability of other international work in this area and of any relevant system standards.

digital cameras
We explored the potential of a number of different types of camera:
- Kodak EasyShare system   (single function)
- Mobile Phone cameras   (multiple function)
- PDA with built in camera   (multiple function)

Our initial trials reflected the Mori survey findings that many learners already had digital cameras in their mobile phones. In this first phase learners reported that these cameras were typically of a low resolution and not capable of capturing sufficient detail in close up to adequately record their modelling. By phase 2 of the project learners' attitudes had changed and some commented that the cameras in their phone were better than those we provided in the PDA. Also they were much more familiar with transferring and sharing data from their phones using Bluetooth.

Putting learners in control of the recording process freed teachers to concentrate on other aspects of facilitating the assessment task. It also engaged learners in the important process of selecting appropriate evidence from the wide range of photographic material they were able to collect. Combining the camera with the integrated features of a PDA meant that it was also possible to annotate and sketch over photos to convey additional meaning.

digital pens
We explored the potential of the following types of digital pen:
- Logitech iO (V1.0 and V2.0)
- Nokia (bluetooth)

• Pegasus Note Taker

To avoid the bottleneck of post-activity digitisation, we wanted to develop a system in which learners were in control and that digitisation happened continually as they worked. We considered the use of tablet PCs and a range of scanning devises, but rejected these because they all proved impractical in the context of a busy workshop environment. They were typically expensive, delicate, complicated to use and all took up far too much working space. At the outset of the project we believed digital pens would provide the most effective route to digitising the notes and sketches produced by learners during the 6 hour e-scape task. To move to a full pilot implementation with the pens, we identified a number of additional hardware requirements.

Beyond the technical management of the hardware there was also a software requirement to provide automatic systems that securely collected, collated and managed the data from the pens, tagging and formatting it and sending it to individual learner web workspaces. In addition we considered a range of other enhancements to the basic digital pen that would make them more suitable as d&t assessment tools, for example:
• fingerprint recognition system
• interchangeable mark-making tools
• integral microphone, camera and printer

As we 'imaged and modelled' these 'nice-to-have' future technical enhancements we realised that we were merely re-inventing a PDA and that many of our requirements were already available in products like the Palm Zire and at much the same cost as a digital pen.

digital sound (speech-to-text)
In our search for an effective system to capture, edit and display learner's 'design talk' we considered the following technologies:
• DragonTalk 'Naturally Speaking' software with head sets and a 3 way switch
• Olympus (digital voice recorder)
• Palm PDA (Voice recorder)
• Qinetiq Social Services system

All the speech-to-text systems we trialled were too demanding for the technology we were likely to have in place for phase 2. The team noted that it would be at least 1 to 2 years before the processing power of a PDA was sufficient to handle natural language in this way. We also explored the potential of creating a conversational 'chatterbot' to support learners working on the 6hr design task.

PDAs (personal digital assistants)
Although a wide range of PDAs was available, there were few that combined the features and ease of use required for this project. Initially we identified and compared 2 models:
• Palm Zire 72
• HP i-Pac rx3715

Both devices offered similar functionality and largely for reasons of cost we chose to pilot with the Palm Zire 72. We had not intended (or budgeted) to commission any software development at this stage of the project, however in order to judge the ability of the small

PDA screen to deliver instructions and act as a contextualizing device, we commissioned Handheld Learning to build a Virtual Handling collection for the Palm. This was a rich media application, which displayed, images, text, sound and video. Work on developing this mini application led to the more integrated proposal for a project management "genie".

Data transfer from the PDA to a secure web based system was a critical aspect of the system and we explored a number of routes to achieve this. Initially we explored the systems provided with the Palm, including; USB, IrDA, Bluetooth, SD card and wireless.

The market for digital equipment moves very quickly. In the time leading up to Phase 2 of the project Palm stopped making the Zire 72 PDA. We worked with our technical partners and selected an alternative devise, the HP iPaq RX3175 which although more expensive had a better specification (particularly the bundled software packages). This product was also withdrawn before we had secured funding for phase 2 and we were forced to make a speculative purchase to ensure we had the technology we required for the phase 2 pilot.

We should note at this point that whilst the phase 2 pilot went ahead with PDAs as the basic data entry tool, it would have been possible to use alternative devices e.g. laptops, tablets, digital cameras etc. The choice of PDAs was driven by a combination of concerns, principally;
• the need for the hardware to be right inside the action, typically in the d&t workshop
• the necessity of having one machine for each learner throughout the activity
• the need for learners to respond in many ways (drawing, writing, with photos and speech)
• the related (organisational) need **not** to have many different bits of technology
• the need not to dominate the workspace with bulky hardware
• the need for it to be robust in challenging workshop environments
• the need for it to be as cheap as possible.

So, whilst we recognise that future scenarios for e-portfolio creation may involve many types of hardware our choice for phase 2 was the PDA. However, as we point out later in this report, the ambition would be that the e-portfolio system should be agnostic about whatever hardware is used to create the work in the classroom/studio/workshop.

Web based e-portfolio systems
The MAPS (managed assessment portfolio) system from TAG Learning provides web based ICT portfolio assessment facilities to over 60,000 learners. The system is hosted on a remote server and provides critical aspects of technical functionality such as; user management, secure access, remote set up, file storage, tagging, back up, virus protection, help and support. In partnership with TAG we identified the technical requirements necessary to manage a web-based portfolio for the 6hr design & technology task. While many of the basic portfolio functions were available from the existing MAPS modules, critically there were a number of additional elements required. In particular we needed an effective link between the classroom-based PDA collection device and the web-based portfolio system.

Display technologies
Assessing design & technology capability on-screen from portfolios of digital evidence was a new endeavour. From analysis of the assessment processes carried out with 'real' scripts by

markers on previous projects, the two key display functions appeared to be 'comparison' and 'scale'. While there were no bespoke assessment systems available to achieve this, the team considered a range of display technologies that could help to augment the assessment process.

ii) pedagogic findings
Concerning the extent to which the use (for assessment purposes) of such a system can support and enrich the learning experience of design & technology

In this area, the findings from the phase 1 trials were quite unequivocal. It was clear that the use of peripheral digital tools offered the opportunity for considerable enhancement of the teaching and learning environment. The following examples illustrate some of this potential:
• the use of the PDA as a device on which preliminary design ideas can be generated was extended hugely by the potential for 'beaming' work between learners.
• the PDA enabled learners to build a digital scrapbook to enrich projects that were underway.
• the facility to take regular photos of emerging models and prototypes was welcomed by learners and teachers alike.
 • the design talk feature was widely welcomed by teachers who were keen to accentuate the peer-reflection / peer-review potential of the system.

iii) manageability findings
Concerning issues of making such assessments do-able in 'normal' d&t classrooms / studios / workshops, the training / cpd implications for teachers and schools and the scalability of the system (including security issues) for national implementation

The issues of manageability were at the forefront of our thinking and these phase 1 trials indicated a number of important considerations:
• was the hardware/software manageable for learners?
• did the hardware raise theft, damage and loss issues?
• was the activity manageable on hand-held technologies?
• could teachers manage the activity?
• was the outcome of learners' work assessable?
• was the system scalable for national assessment purposes?

We were also concerned to establish how familiar learners needed to be with the kit and how much training teachers needed to be comfortable with the system?

iv) functionality findings
Concerning the factors that an assessment system based on such technologies needs to address such as the reliability & validity of assessments in this form and the comparability of data from such e-assessments in d&t with equivalent paper-based assessments.

The three big issues here were – validity, reliability, and comparability – and from the phase 1 trials we were in a position to comment principally on the first.

Validity takes several forms, but for the purposes of our work may be summarised as the extent to which the activity, as we have designed it, represents 'good' design & technology.

28

A standard approach to the problem of deciding on validity, is to appoint an expert panel of design & technology specialists, and invite them to make a judgment as to whether the activity is – in their view – a good example or model of design & technology. In relation to this procedure, there are three kinds of evidence that may be considered.

First, the activity is a direct development from the booklet-based activity that was initially developed in the previous project; 'assessing design innovation'. The assessment activities that were devised in that project originated with the research team but were then shared with the principal subject officers of the four Awarding Bodies. These Bodies were sufficiently impressed with the work to each 'volunteer' two of their principal design & technology moderators to work with us in developing more activities of the same kind. These then formed the basis of the extended pilot, and were warmly received in the 8 schools in which they were administered.

Second, the e-scape activities have been developed directly from these former activities. We conducted one trial of the 6hr kind that was a complete reflection of the previous project (even using one of the same tasks), and we additionally conducted a series of shorter trial activities to test parts of the process. All of these trials were done in booklets that – whilst being modifications of the original – were recognisably the same format and had a number of identical features. There is prima facie evidence that the resulting activities might be regarded as good models of design & technology merely through association with their predecessors.

However, the third arm of the validity case lies with the teachers who have undertaken these e-scape activities. These teachers have been clear in their view that whilst the tools have changed (from analogue to digital) and that this has resulted in a very different form of representation of designing, the activity itself remains true to design & technology.

Reliability is a more straightforward concept, and may be thought of as repeatability, in the sense of whether an assessment made by one teacher (using our system) is repeatable by another teacher. Do they arrive at the same result? This may also be extended to examine repeatability between those concerned with the process. Do teachers in the classroom make the same judgments as Awarding Body officers given the same evidence?

The comparability question is a straightforward challenge, asking whether a learner achieves a similar level of performance when assessed through the e-scape system as he or she does when assessed through a paper-based equivalent activity. Reliability and comparability are both matters that we examine later in this report.

These findings; technical, pedagogic. manageable and functional – taken together – informed the development of a specification for the e-scape prototype system that we built and trialled in phase 2 of this project.

1.6   specifying for system for e-scape phase 2

The purpose of the e-scape phase 2 prototype is to create a system where the individual components, explored in principle in phase 1 of the project (and judged to support d&t capability and make it available for assessment) are built into a working prototype. The

elements must work together sufficiently well in the field to ensure we can put enough learners through the system to collect sufficient data to answer key research questions for the various stake holders, such that they are confident to move forward to a commercial, scalable implementation in phase 3.

Throughout this process, data integrity and security was a particularly important issue. Specifically we needed to minimise the possibility of data loss during the 6hr activity where learners were working on potentially volatile handheld systems. We also needed to ensure that the interfaces we developed (both on and between the technologies) were intuitive and easy for teachers (and learners) to understand and operate. We used the outline specification below to guide our development.

In preparation for phase 2 we worked closely with our technology partners TAG Learning and Handheld Learning to develop the specification for a working system. While we had a clear idea of what we wanted the e-scape system to do in principle we needed lots of help to convert this into a working prototype that functioned technically as well as pedagogically. We are grateful for the support of our two technology partners without whom the project would not have been possible.

TAG Learning has a long and successful history of working in partnership with a range of organisations to develop innovative web based portfolio systems across a wide range of active learning and assessment contexts. Their MAPS e-portfolio assessment system provided the foundation for the research team to image and model how the web components of the e-scape system might function. TAG developers worked closely with the e-scape team during the first phase of the project and through an iterative specification process modified the components of their core system to describe the functionality we required.

Handheld Learning are the leading organisation promoting mobile technology in education in the UK. As well as developing PDA hardware and software they have initiated and supported a wide range of projects in schools and local authorities. In the first phase of the project Handheld Learning developed a virtual handling collection for the Palm PDA that presented prompts to encourage learners to think differently about the objects presented. Handheld Learning developers then worked closely with the e-scape team to convert the stages of the paper-based activity into a format that could be delivered within the confines of a PDA screen.

The development process was further complicated by the need for the web and PDA systems to talk seamlessly to each other throughout the activity. TAG and Handheld Learning collaborated closely to make this happen sharing protocols and aspects of their individual systems in a way that transcended normal commercial expectation, and that guaranteed the system worked.

The system modelled as closely as possible what had been shown to work for design & technology in the paper-based environment. No changes were made for purely technical reasons. The changes that were made were for reasons of pedagogy or manageability. As far as possible we avoided making demands on school ICT systems, over which we had little or no control.

key aspects of the system

During the first phase of the project we explored, developed and trialled a range of individual digital enhancements to the original paper based framework for assessing design innovation. In order to consolidate these into a working system, we described the following system requirements:

1) access before (to prepare for the test)
    - including teacher, learner and Awarding Body support systems
2) setting up
    - hardware, data, connectivity requirements and teacher support
3) the 6 hour activity
    - security, instructions, resources, tools, data integrity
4) data transfer
    - between learners, learners and teacher, to web-space, to assessor/moderator
5) access after the test (to develop, mark and moderate)
    - log in, web-based workspace for teachers/ assessors/ moderators/ Awarding Body
6) administration roles
    - school administrator (examinations officer?)/ awarding body administrator
7) Collecting evidence/feedback
    - during trials and as part of the published system

Standards

The e-learning, e-portfolio, e-assessment territory is informed by many sets of technical standards (some international) that seek to 'standardise' that territory – either in terms of systems input, user protocols, and/or output processes. Wherever possible and appropriate we will take due note of these standards. However it should be noted that the challenge for the e-scape phase 2 prototype was to create something that did not exist anywhere in the world, and accordingly there were no standards that entirely circumscribed our work. The e-scape project will inform international standards as much as be led by them.

1.7 emerging research questions for phase 2

The development of the prototype system in phase 2 of the e-scape project was based on the specification outlined in section 1.6 above, and following its development, a school pilot was conducted to explore the efficacy and the effects of the e-scape, e-portfolio, e-assessment system. A number of research questions informed this work.

In the construction of the prototype system, we were guided principally by *technical* questions, concerning for example;
   • the connectivity between hand-held devices in the classroom and web-spaces
   • the possibility of pre-defining this web-space so as to construct a virtual booklet
   • the security of access to this virtual booklet through user-names / passwords
   • how robust is the system?
   • what is involved in managing and maintaining the system?

However, the process of developing the prototype was also informed by pedagogic, manageability and functional assessment questions, for example:

pedagogic: how will the construction and appearance of the virtual booklet impact upon the questions and sub-activities that need to be built into the activity? How is the designing activity changed by the system? What backwash effects would teachers anticipate in relation to KS3 practices?

manageability: how often will the PDA need to be sync-ed to the web-space? How long does the process take and can a class of (say) 24 learners manage this process simultaneously? How do-able is the digital activity in normal studios/workshops? How much cpd/training do teachers need to prepare for this mode of assessment?

functional: how does the assessment process change when viewing the virtual booklet in the web-site as opposed to real paper-based booklets? Does the system enable valid and reliable assessments?

The work of phase 1 is more fully described and analysed in the phase 1 research report (see App 4.1)

e-scape phase 2

Following a series of negotiations with DfES, QCA, Awarding Bodies and TERU, phase 2 of the e-scape project was launched in November 2005 and ran to the end of January 2007. The story of phase 2 is told through the following section 2.1 to 2.19.

2.1 task trials

Before we could be in a position to trial the e-scape system, we had to trial the design task that we had developed for the e-scape tests.

The escape task, entitled '*the pill organiser project'* involves learners in a product design activity, developing a container/dispenser for pills. Learners have to identify the user group (maybe a 6yr old on a school trip, an activity-sports enthusiast or an elderly lady living alone) and think about all the issues involved:

> how many pills?
> taken how often?
> how to remember to take them?
> how to keep them secure?
> how to make the container/dispenser desirable? etc etc

Learners develop their design solutions using the basic approach developed in our previous DfES/QCA project '*assessing design innovation'*. The activity runs for 6 hours through 2 mornings in schools, and includes a significant amount of 'soft material' modelling activity to support design development.

The task is resourced by five things:
• the booklet – with associated images (one for each learner)
• the handling collection of idea objects (one for each group of 3 learners)

• 'client' or 'user' cards – profiling particular users and their pill requirements
• the central 'inspiration' collection
• the central modelling kit

Additionally the activity is managed through the administrator 'script' and the e-scape application on the PDA.

All of these had to be trialled to make sure that the task would provide opportunities for learners to do their best and most imaginative designing, and to check that it was do-able across the specialist material territories of d&t.

We undertook four trials:
• with a Goldsmiths ITE learner group
• with a year 10 group in Bulmershe School
• with a year 10 group in Invicta School
• with another year 10 group in Invicta School

Throughout this process we developed not only the test activity and script, but also the resource lists. It is interesting how the availability of particular resources influences the emerging designs. At Bulmershe, despite some interesting ideas emerging, there was a 'boxy' feel to much of the work. When (in later trials) we supplied more fabrics and (particularly) some plasticine, the variety of responses blossomed. We concluded that:
• sheet materials (paper/card) best enable '*boxy'* forms
• strip materials (dowel rod/straws/wire) best enable *skeletal* forms
• fluid materials (plasticine/clay) best enable *organic* forms
• and also that textile materials (fabrics) often link to and operate across these types

By the end of the task-trialling we were confident that we had a task (and a set of resources) that learners could have a good run at, and show us what they could do.

2.2 e-scape system components

The escape system operates at several levels:
i. The school workshop/studio activity
The e-scape activity is self-contained in the sense that the research team take ALL necessary materials and equipment to the school and remove it thereafter. We recognise the demands we make on teachers and schools and we endeavour to make the experience as manageable as possible for teachers and as stimulating as possible for learners. Studio components include:
• group handling and inspiration collections for each group
• modelling kit with range of rapid prototyping materials
• learner workbooks and administrators script

ii. The PDA

The central transformation of **e-scape** (moving on from the '*innovating assessment'* project) is that the paper-based work has been, where appropriate, replaced by work on hand-held computers – PDAs – in this case the HP3715. The strength of this tool lies in its pocket size, so it does not dominate the space and its multi-functionality. It can be used for:

• *drawing*,

• taking *photos*

• recording *voice memos*

• *note-writing*

  (transcribed handwriting or virtual keyboarding)

Rather than write a complete suite of applications from scratch, the development team at Handheld Learning created a software wrapper (it maybe helpful to think of this as digital sellotape) which tied together the existing painting, camera, text and audio tools. The resulting e-scape PDA system presented the activity as 20 linked screens each of which presented:

• the instructions for each sub task

• the time remaining

• direct links to the specific tools required (i.e. text and audio recording)

Each PDA is personalised and registered to a single learner, with the learner's name written on the back, and to avoid confusion during the activity whenever the PDA is switched on it confirms the learner name on the screen.

Moreover the PDA has Bluetooth and wi-fi capability, enabling us to link a class-set of them into a local area network. It therefore has interesting potential not only as an
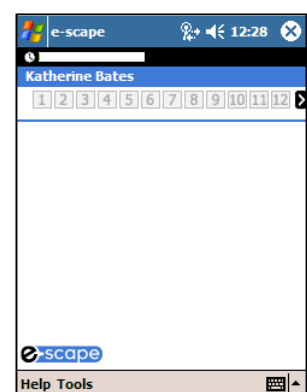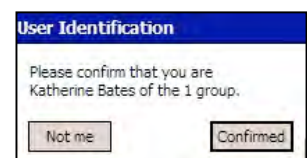
assessment tool, but also as a collaborative (group-based) learning tool.

iii. The local area network

This is run from a laptop computer managed by the teacher/researcher administering the activity. The PDAs all have the e-scape software loaded into them in the form of 22 linked screens through which the activity progresses. These screens operate in a similar way to the boxes in the '*innovating assessment*' booklet. The administrator controls the activity through the laptop. Once logged in, and just as they could with the paper version, learners are able scroll through all 22 screens and see the complete extent of the activity on their PDA. The screens are greyed out and it is not possible to work in them until the teacher/administrator switches them on. As s/he activates "box 1" on the administrator interface, the signal is sent (via wireless router) to all the PDAs and they all "come live" with box 1.

Learners can draw/make notes on the screen, and at the end of the allotted time for box 1 a warning window pops up prompting learners to 'tap-here' to save there work. The material they have produced is then sent back through the router and stored in the laptop – as well as in the memory on each individual PDA.

At this point the box switches from edit mode to view mode, so the learners can review what they have done but not change it. This means that everything learners add to the PDA is time stamped and, unlike the paper version of the test where we had to juggle with different coloured pens to establish when something happened, we can be sure that material in box 1 was produced in the first 5 minutes of the activity and not added later. The 'fast back' facility was added at the very end of the task to allow learners to go back and review everything they had done and add digital post-it review notes of what they might have done differently.

iv. Uploading data

Information passes to and from the set of PDAs throughout the activity, opening the next activity box, exchanging files between group members, and collecting, collating and storing work in the web portfolio on the portable server. It was necessary to create this virtual network in order to overcome shortcomings in many school wireless networks. The router and portable server were simple to set up and as the capabilities of PDAs continue to grow, the suggestion from our technology partners is that it would be possible in future to replace the portable computer with a teacher's PDA, and run class sets of PDAs in future e-scape activities from this.
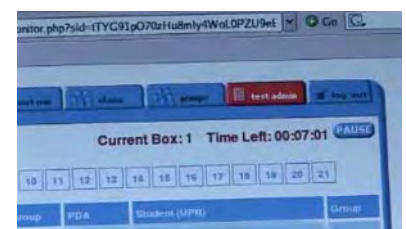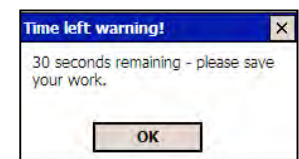
By the end of the activity, learners will have created a mass of design data spread through the 22 linked screens. The digital evidence trail of their design development process includes drawings, notes, photos, and sound files. All these data are held (temporarily) on the PDAs, and on the laptop. We also have a routine to save all the data from a school/centre on a 500 mb memory-stick attached to the laptop. These "belt & braces" techniques are designed to secure the data in case of a problem with the primary system.

This primary system involves linking the laptop to an internet-connected computer so that all data is up-loaded directly into a secure web-space. Throughout the activity the portable server system 'sniffs' for an external wireless connection to the internet (it could also be connected through a cable system), if it detects one and makes a connection the locally cached material can be instantly uploaded to the web based portfolio system.

v. Reviewing and assessing the portfolios online

As soon as the data has been uploaded from the portable server it is displayed in a secure web space, which can be accessed, with the appropriate password and username, via any internet connected PC. This means that individual portfolios can be viewed simultaneously by the learner, or the teacher or any number of assessors. Because of the small screen size, the PDA presented the work as a series of individual boxes, accessed from a number bar at the top of the screen. For the web based portfolio TAG Learning arranged thumbnails of the 22 boxes on a single screen, to give an overview of the whole portfolio, with the option to click on each to zoom in and out and examine each stage in more detail.
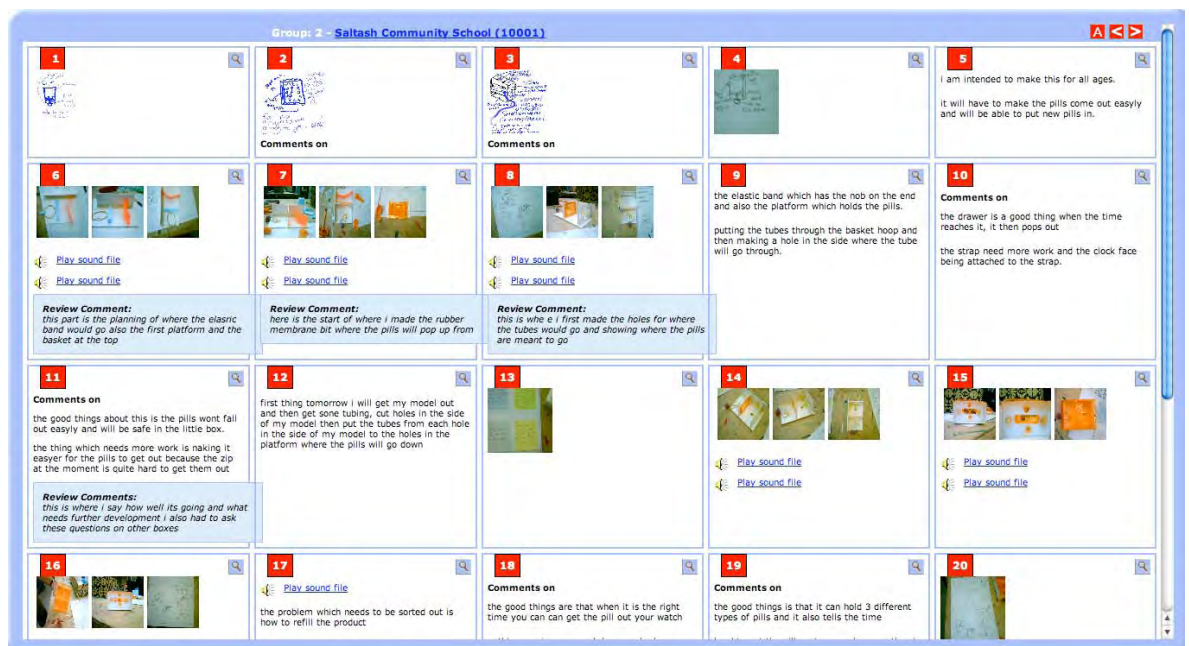
TAG also included a marker training section on the website where the research team could upload exemplars, mark schemes and other materials to help the assessors mark the

portfolios effectively. There was also a facility for the markers to upload their marksheets once they were completed.

2.3 system trials (May 2006)

The beta version of the e-scape application was delivered to TERU on 18th April and we did a complete 'walk-through' of the activity from start to finish, partly to check how it worked and partly to check that it did what we wanted it to do. We immediately identified some modifications that we requested prior to formal trials.

Trial 1 Uplands Community School

This was the first ever attempt to make the system work – and it showed. The school, the teachers and the learners were quite fantastic and worked their socks off to help us make the system operate. But the 1st morning was so disjointed (system crashes – lost data – PDA freezes) that we decided to use the 2nd morning just as a re-run of the 1st morning – but with a different task.

This trial illustrated what a challenge it is to get box 1 / 2 / 3 / and 4 completed properly. Box 1-3 is where learners initiate their ideas and then swap them around the group – for team supportive comments and drawings. Box 4 then allows them to review all these comments and ideas before moving on. In short, the whole potential of e-scape is being put seriously to the test in the 1st 20 minutes of the activity. After that it eases somewhat.

The 1st trial demonstrated that we needed to be far more careful about our management of that opening 20 minutes. Re-running it all on the 2nd morning began to show us how to do that and by the end of the 2nd morning we had got to the middle of the activity. We then used some spare time to show the learners what they had done. We projected their work back for them and they were not only impressed and fascinated by it, but made the point that it really would have helped to see (at their initial training session) this 'big picture' of how it is all supposed to work.

Trial 2 Camborne Technology School

After a one-week interval, during which TAG put right as many of the glitches as we could identify, the team de-camped to Cornwall for a week of trials. The Camborne trial was altogether a smoother affair. The system held up throughout and we only had a small number of PDA freezes. This success was due in equal measure to TAGs sorting of glitches in the software, and our re-designing of some of the classroom protocols at the start and end of boxes. Once we knew (from the 1st trial) where problems might lurk, we could design protocols to overcome them. As a result, in the 2nd trial we sailed through the challenge of box 1,2,3,4 and on each morning we completed the allotted sections of the activity. So by the end of the 2nd morning we had – for the 1st time – run the activity right through to the end.

As the activity went right through this time, we were dealing with far more of the photography and sound files – both of which caused great interest with the learners. Whilst recording the sound files initially caused some embarrassment, they soon learned the process and it rapidly became just a normal part of the activity.

Since we were developing and refining these protocols during the 2nd trial, it is not surprising that we still had some 'lost' data from individuals in the Camborne trial. None-the-less we got to the end of the activity and most of the learners up-loaded all their data.

Trial 3 Saltash.net Community school
In this trial we focussed closely on the timings necessary for familiarising learners with the tools in the PDA. Our early discussion with developers had suggested that learners would need *weeks* to become familiar with the tools. We subsequently debated leaving the PDAs in the school for the preceding *week* – so that learners could become familiar with them. In the end – for logistic reasons - none of this was possible and we were rather working with a familiarisation time of *hours* on the days preceding the activity. This general move to shortening the familiarisation time-scale was also based on the reception of the learners themselves to this new technology. They adopted the PDAs as an almost natural extension of their mobile phones, having very few difficulties with it. We found that the familiarisation routines had rather to be focussed on the e-scape application interface – which was completely new to them.

In the end, we tried to do the whole familiarisation session in 1 hr, and concluded that it really was not possible to cover everything in that time. The trials showed that ninety minutes (often a double period or an afternoon) proved adequate, and we accordingly contacted all the main e-scape schools to find out the exact timing of session – both for training and for the activity – to make sure that we get the necessary time.

The Saltash trial went reasonably smoothly. We put in place all the protocols, and by now we were becoming more comfortable with the set-up and management of the system. Throughout the two mornings of the activity, very little data went missing, and the system worked well.
Overall, the trials did what they were designed to do. They enabled us to tweak the application and taught us how to manage it in the classroom / studio / workshop setting. Throughout the process we were really impressed by the schools, the teachers and the learners – without whom we would not have been able to take this important step forward.

2.4   training the teachers (May 2006)

It had been clear from the outset that we would need to train the teachers in the use of e-scape system. Initially we had hoped to do the training and thereafter to expect the teachers to run the activity in the schools – freeing the research team to become participant-observers of the activity. This would have enabled us to adopt more of a *researcher* role during the activities. In the event, it became clear from the trials, that operating the system as a novice was a fairly demanding task. This arose from the fact that this 1st generation of the e-scape application (and its surrounding hardware) was a complex melding of existing systems and applications. The e-scape application operated as digital sellotape linking pre-existing bits and pieces into a single entity. Inevitably there were some cracks in the system and it took a couple of runs through to become confident in administering it. But having done it two or three times it becomes far more manageable. Accordingly we decided that the teachers could not be expected to be the front line administrators from the start of the activity and that we would have to take that role. Gradually however – through the two days of the activity - we hoped that teachers would feel able to take over the management of the activity.

Despite this change of position, we still felt it well worthwhile to train the teachers in advance and accordingly we scheduled two one-day training sessions, one in Birmingham on Tuesday 23 May and the other in Newcastle on Wednesday 24 May. The locations, schools, and attendees for these training sessions are shown below.

Tues 23rd May   The Bond Company. Birmingham

Nancledra C.P. School	Pauline Hannigan
Alexandra Park School	Ross McGill
Redruth School	Mike Laramy
Meole Brace School	Stephen Cox
Ashfield School	Joanna Hayes & Dave Rogers
Edensor Technology College	Nick Bradbury
TERU Goldsmiths	Kay Stables
Edexcel Awarding Body	Dale Hinch

Wed 24th may   Blackfriars Cafe Bar. Newcastle upon Tyne

Duchess's High School	Craig Watson & Diane Murphy
Dixons Academy	Maria Eccles
Coquet High School	David Coils
Hirst High School	Bob Miller & Mark Raper

One teacher (Fiona Mather from Bedlington High school, Northumberland) was unable to make either of those dates due to other school examinations, so we arranged for a private training session for that teacher in her school on the day after the Newcastle training.

The training day programme was based on a number of features, including;
• the structure and design of the 6 hr activity (why it's like it is)
• the e-scape kit (hard & software + task & resources)
• the experience of the trials (+ showing the website with existing work)
• e-scape training for learners in their schools
• setting up the school work-space and running the activity

At the end of the training sessions, the teachers reported that they understood the system and were confident that they could set up the facilities appropriately in their schools.

2.5   the national pilot (June / July 2006)

We had numerous communications with the schools taking part in the national pilot, mostly by letter or email. Having got permission from the head and made contact with the department staff responsible, we began a drip-feed communication link with them – partly to keep e-scape active on their radar and also to accustom them to communications with the TERU office and administrator.

The training sessions were run in the week before half term in the summer term, and the e-scape pilot was planned to start immediately after that half term. To run the national pilot we had acquired four sets of kit;
• four class sets of PDAs (24 per set - for 7 groups of 3 learners with 3 spare)
• four administration laptops running the local network
• four wireless routers

• four sets of handling collections
• four sets of modelling materials

This enabled us to run the e-scape activity in four sites simultaneously, provided that we also had four administrators. In addition to Tony Wheeler and Richard Kimbell, four colleagues who had been involved in earlier stages of the project were trained to help facilitate the national pilot: Soo Miller, Tristram Sheppard, Kay Stables and Ruth Wright.

Eleven schools took part in the pilot, and since we could run four at a time, they were clustered as follows with the activity being run by the TERU team member.

Round 1 (June 5th-9th)   TERU
    Dixons City Academy – Maria Eccles   (Tony)
    Nancledra CP School – Pauline Hannigan   (Soo)
    Edensor Technology College – Nick Bradbury   (Tristram)
    Redruth School – Mike Laramy   (Richard)

Round 2 (June 19th-23rd)
    Bedlington Community High School – Fiona Mather   (Tristram)
    Hirst High Technology College – Bob Miller   (Ruth)
    Coquet High School – Steve Thompson   (Richard)
    Duchess's High School – Craig Watson   (Tony)

Round 3 (July 10th-14th)
    Alexandra Park School – Ross McGill   (Kay)
    Meole Brace School – Stephen Cox   (Tony)
    Ashfield School – Joanne Hayes   (Richard)

The details of the pilot are shown in App 3.3)

The pilot week in any school had the following broad framework. We undertook a training session with the learner group to familiarise them with the *e-scape* application and the PDA and specifically with the sketching, text entry, voice recording and camera functions that are important in the e-scape activity. This training typically took a double period (typically 1.5 hrs) on the afternoon preceding the 1st day of the activity. Thereafter the activity itself ran for two consecutive mornings providing a total of approx 6 hours of activity. The activity was run by one of the research team. The lead teacher in the school was present throughout the activity to observe and assist as necessary.

setting up the system

On arrival at the school, the activity administrators' first task was to set up the laptop, wireless router and the PDAs in such a way that the PDAs each communicated with the laptop via the router. This was done in preparation for the learner-training afternoon session, since we decided that the training should involve more than just the PDA toolkit. It was also important to familiarise learners with the e-scape application interface. Accordingly, the training involved running the group through the first 6 boxes (sub tasks) of the activity. By that time they had used the drawing tool, the writing tool, the camera and the voice memo tool, all in the context of the *e-scape* application.

Having completed the learner training / familiarisation, the administrator had to strip out from the system all the learners' work completed during the training and prepare it for the activity the following morning. To achieve this it is important to recognise that the preparation of the hardware involved a serious level of personalisation. The laptop was programmed with the class names and details of the up-coming centre, and each of the PDAs was allocated to one of the individual learners. So when the learner switched on the PDA for the 1st time it would immediately come live with the interface '*please confirm that you are Sam Walker of group 2'* (there were 7 groups of 3 learners in each class). Since the PDAs each had sticky labels with the same names, there was rarely any difficulty with this arrangement.

Where difficulties did arise it was typically because of absence. If Sam Walker was absent, then one of the three reserves took her place. This involved re-assigning one of these replacements into group 2 in place of Sam Walker and up-dating the 'register' in the administration laptop. This happened in several schools and sometimes with more than one learner. We sought to create the *final* working group at the training session in the afternoon preceding the 1st morning of the activity. It was rare (but not unknown) for this updating to be further up-staged by another absence on the morning of the activity itself. The easiest way to resolve this (with the time pressure of getting the activity started) was simply to ask the new substitute to adopt the name programmed into the system. In addition to setting up the system, the studio/workshop space had to be prepared with the handling collections and the general resources of the activity.

2.6   an overview of the activity

Once the system was set up, the activity typically ran extremely smoothly. Our previous experience with *Assessing Design Innovation* and all the trialling of the task that we had undertaken in early 2006 ensured that the activity provided a rich and engaging experience for learners. They were typically enthusiastic about using the PDA as a design tool and they invariably responded to the 6 hours of activity with wholehearted commitment.

The arrangements for the start of the activity involve a classroom set up with seven working tables each with a group of three learners. The groups had a handling collection to explore at the start of the activity (in the yellow/black box) and this contained approx 15 objects that were intended to stimulate thoughts and ideas about storing and dispensing pills. None of the objects were associated directly with pills but rather with the *concepts* of storing and dispensing. Once the task had been established, the activity started with five minutes of free exploration of the handling collection.

The learners also had their PDA, their booklet, (see App 2.1) a set of fine-liner pens for use with the booklet, the user profile cards, and a set of wooden 'pills' (twice full size). Elsewhere in the room was a table or work surface set aside for the additional handling

collection of 'big' objects. These were either objects that would not fit into the box or objects that we did not have in sufficient quantity to put in every handling collection box.

The final elements of the room set-up was the table or work-surface set aside for the modelling resources (see App 3.1). Following the explorations described above we made sure that this contained sheet, strip, fluid, and textile materials as well a collections of ways to cut form and fix them together. There was also an assortment of plastic balls, springs, elastic bands and other odds and ends of gadgetry.

To appreciate the set-up of the room and the activity, readers will find it helpful to refer to the CD enclosed with this report. This is presented in a series of chapters and we would refer readers in particular to the two early chapters on task setting and handling collections. Additionally readers may wish to refer to the Teachers TV website where three programmes about the escape project can be found (see also section 2.8 below)

(http://www.teachers.tv/video/3306>http://www.teachers.tv/video/3306)

Following the exploration of the handling collections, learners are asked (in box 1) to put down their 1st ideas for how they might develop a new pill dispenser. They do this (for 7 minutes only) on the PDA.

At this point (for boxes 2 and 3) the system moves learners' work around the group so that learner x's work appears on learner y's PDA. Their task is to extend and enrich the first ideas of their teammates.

When, in box 4, the work returns to the originator, learners are asked to review all the thoughts and ideas that have been offered by their team-mates and to consolidate what they think are the best ideas for moving their product forward. They do this in the paper booklet whilst reviewing the contributions on the screens of the PDA.

We then ask learners to reflect on these ideas, specifically in terms of who the *users* will be, and what their specific requirements might be. They do this as text in box 5 on the PDA, using a keyboard tool and/or the transcriber.

42

We then introduce learners to the modelling resources. This is an important step since they are typically not familiar with the relatively 'free' notion of material modelling that is central to e-scape. We accentuate that models can be of many kinds, and that early conceptual models might look crude and certainly unfinished but that is the point of them … their value is in helping learners to sort out their ideas.

Learners then have a choice of continuing to develop their ideas by through sketches and notes in the booklet, or by working – in modelling mode – with the modelling resources.

Approximately one hour into the activity we ask learners to take a set of photos of their work (booklet or model or both) with their PDA. Having taken them we ask them to answer two questions; one focussed on what's going well with their work, and the other on what is not going so well and that needs further development. These questions are posed through the PDA and we ask learners to record their answers as 30 sec sound bites or voice-memos.
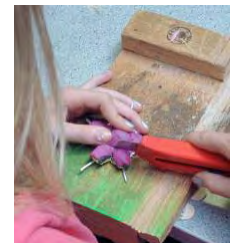
These voice memos were completely new to all the learners involved in the project. It was therefore important that they learned the procedures for doing them and became comfortable with them. The photo-sound-bite routine was repeated 6 times through the course of the activity and learners became skilled at composing and reporting their thoughts. In the early stages however it was important that they could listen to their 1st attempts and – if they were no content with them – they could erase them and do a 2nd 'take'. The activity then proceeded as a series of sketching / modelling phases iterating with the 'pauses for thought' provided by the photo / voice-memo sequence. Typically the modelling became progressively more sophisticated with the ideas behind being subjected to ever-closer scrutiny in the sound bites. The ideas illustrated below are representative of the sample as a whole.
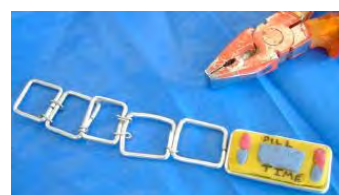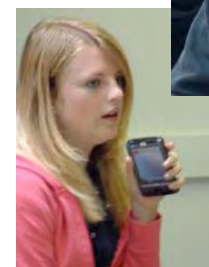
At the end of the 1st morning and then again towards the end of the activity as a whole, we asked team-mates to be part of the review process – advising each other on the strengths and weaknesses of their work as seen through the eyes of their team-mates.

These were text-based activities in the PDA and were taken very seriously by the teams. The photo here illustrates a profoundly deaf young man working with his signer as he reviews the work of his teammates.

The penultimate step of the activity involves learners making a 'pitch' to the managing and marketing directors of a company that might be interested to develop their product concept into a real product. They have 30 seconds – in a lift as the directors are going up to the 6th floor – in which to summarise the originality and potential of their product concept and to persuade them to buy the idea.

Finally, learners review their whole process of design & development – reflecting back (with the benefit of hindsight) on the various steps they have taken. What might they have done differently if they had known then what they know now?

As the teachers became familiar with the workings of the system they too typically lost their nervousness and started to take a part in the management of the activity '"I'll do the next bit" and "Can I have a go at running it". It would have been too much to expect teachers to run it from the start, but by the end of the two mornings it was commonplace for them to be comfortable with it.

Following the activity, all the hardware was returned to TAG who stripped off all the learners' work and mounted it in the website. This could have been done remotely from the schools, but to save administrators' time it was decided simply to back up the data set using a USB stick and return the kit to TAG. All the non-digital learner outcomes resulting from the activity (models, booklets, questionnaires) were transported to and stored in the TERU offices.

2.7   the response in schools

In addition to the performance data arising from the activity, we have collected two other sets of evaluative data to compile an authentic account of the e-scape activity and the response of the schools.
• the learner questionnaire (completed at the end of the activity)
• teacher feedback (after the event and reflecting on it)

We present a summary of these data below.

• *the learner questionnaire*
The questionnaire was completed by 256 learners and can be seen in full in Appendix 2.3

The opening part of the questionnaire asks learners about their previous experience of using PDAs, and 84% said that they had little or no previous experience of using them. When asked how much more time would be needed to get used to working with them, 82% reported that they would need little or no more time. These data confirm the informal impression created in the activity that learners very rapidly got to grips with the device and its associated software. Whilst teachers were somewhat more nervous of them, the learners adopted them as merely an extension of the mobile phones that they all use ubiquitously. This is confirmed by learners' response to the statement 'it was easy to learn to use the PDA'. By a massive majority of 127:1 (99%) learners agreed or strongly agreed with the statement. Equally their response to the statement 'it was fun using the PDA' showed a 50:1 (98%) majority agreeing or strongly agreeing.

The 2nd part of the questionnaire asks about particular features of the PDA and learners reactions to them as part of the activity.

96% agreed or strongly agreed that it was easy to follow the activity through the e-scape interface.
94% agreed or strongly agreed that it was good for making the photo story lines.
92% agreed or strongly agreed that it was good tool for designing.
90% agreed or strongly agreed that the text tool was good for explaining ideas.
89% agreed or strongly agreed that they were able to show what they could do.

Within all of these sections of the data there is no significant gender variance.
The only significant gender effect is observable in response to the statement *'the voice memos were good for explaining my ideas'*. 50 boys but only 24 girls *strongly* agreed with this statement, whilst one boy but 14 girls *strongly* disagreed. In response to the voice-memos therefore, at the extremes of the data there is a clear effect that suggests girls are less likely to appreciate it. The less extreme data (agree or disagree with the statement) is gender balanced, and overall 70% of learners agreed or strongly agreed that the voice memos were good for explaining their ideas. We believe that the identifiable gender effect at the extremes is related to the embarrassment / discomfort that some learners felt in talking about their work in the public arena of the activity and the working groups.

Some learners moved away into corners to record their thoughts, partly (perhaps) from embarrassment and partly (perhaps) because it offered a quieter setting in which to think about what to say. In any event, whilst 81% of the boys agreed or strongly agreed that the voice memos were good for explaining ideas, the figure for girls was 64%, with a further 13% of them *strongly* disagreeing.
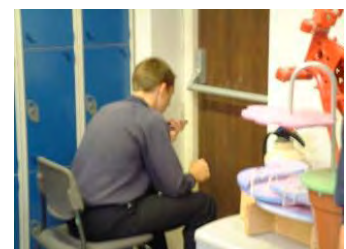
The final part of the questionnaire is a free-response section inviting learners to tell us the three best things (thumbs up) and the three worst things (thumbs down) about working with the PDA. Their responses have been analysed into *categories* of comment, and we show

below the top four categories in each case. Naturally – since these are free-response comments – the percentage mentioning any particular issue is smaller than those from the fixed elements of the questionnaire discussed above.

Thumbs up

category	*mentioned by..*
quick / easy to use	48 i.e. 19% of learners
photos / camera	41 i.e. 16% of learners
fun / different	38 i.e. 15% of learners
voice recording	20 i.e. 8% of learners (16 boys / 4 girls)

Thumbs down

category	*mentioned by..*
slow / crashed / it went to sleep	37 ie 15% of learners
sketching awkward on screen	31 i.e. 12% of learners
transcribing / stylus too small	31 i.e. 12 % of learners
voice recording	27 i.e. 11% of learners (6 boys 21 girls)

It is interesting to note that the top category of 'thumbs-down' comments is purely technical; clear evidence of technical failure with the PDA during the activity. And we should note that the *technical* difficulties giving rise to these comments are rapidly being ironed out through the emergence of more stable platforms and devices. By contrast two of the top three 'thumbs-up' categories are about the emotional response of learners to operating on the PDA – its fun / easy / different. It'll be even easier, quicker and more fun as the stability of the devices improves.

• teacher feedback
Comments were sought from teachers on several issues and the following comments were typical of those returned to us.

Concerning the **task** (pill organiser)
• I liked it because it was a new idea for the children – probably something they had not even considered before
• The pills task was unfamiliar to our learners ... I think this helped them approach the problem with fresh minds and with less pre-conceived ideas.

Concerning the **activity**:
• The activity structure works quite well and maintains pace and focus, I have been trying a similar approach with some of my key stage three groups with reasonable success (RR)
• I was very pleased how the children stayed on task even thought they must have been flagging by the end. I think this was due how the task was structured as well as the eagerness to do all things digital.

*Concerning the **e-scape system***
• I was amazed how quickly the children grasped the technology and were in no way over awed by it (I shouldn't have been!)

46

• *I was particularly impressed with how they used the voice recordings and took them so seriously. I feel this has tremendous potential for capturing their thoughts and feelings about their work as they are doing it. (Nan)*
• *They found the novelty and ease of use of the PDA's a positive motivator*

One of the teachers had written a brief report to his head-teacher, describing the event and the reaction of his learners. With his permission we re-produce it here to indicate the general response of the group and of the department.

Review of the e-scape project

Outline of the project
At the start of June twenty-one year 10 learners took part in a national pilot project aimed at developing innovation and team-working capability using an e-portfolio examination system. The main purpose of the pilot was to test a software system that uses hand held PC's to allow learners to develop their design & technology projects digitally, to submit them via a secure website and for the examination authority to download the work for assessment. This was apparently a world first for using hand held PC's in this way.

Learners took part in four three-hour sessions over two weeks. They were given two specific design tasks (innovation challenges) that required them to develop their ideas and produce a three-dimensional model of their solution. For the first challenge all recording of ideas was through the hand held PC, sketching ideas or making notes directly onto the screen using a stylus, making voice recordings to evaluate their ideas and annotate photographs, they took of their models. All information from the PCs was automatically stored on a central PC for assessment and evaluative purposes. The same learners then took part in a second challenge, this time recording ideas in a workbook. The results of the two approaches will be compared in the final analysis of the results.

Benefits for learners in taking part
o *Learners taking part gained from the focus on creativity and innovation; this will help them with selection and initial development of their GCSE coursework project.*
o *Encourages generation of ideas through modelling, this benefits those who are put off by sketching ideas, because of their perceived lack of drawing skills.*
o *They have learnt new techniques for sharing ideas and working with others to evaluate and develop their design ideas, this is a key feature of assessment for learning and merits further development across design technology.*
o *The different approach to using ICT for collecting and recording information will provide them with a concrete example of how computers can change working practices.*

Feedback from learners:
All learners found the PDA's easy to use
95% of the learners thought that the structure of the activity was very helpful.
   "I liked the way the sheets were laid out, it really helped"
85% felt that the blocks of time were helpful
The three best things about the activity ...
   "The long amounts of time for modelling…I could get lots done"
   "The timed periods as it made me work hard all the time"
90% of the learner found the modelling resources were helpful
   "The materials we were given helped"
All learners found the collaborative group work very helpful
   "Group input for ideas helped me come up with my final idea"
   "we were in groups and shared ideas and comments"

Benefits for technology department and the school as a whole
Apart from the use of the hand held PC's the project included a number of useful and interesting features in the delivery of the project which can easily be incorporated into both key stage 3 and 4; these include:

o   *the use of handling collections of three dimensional artefacts that the learners can study and derive ideas from.*
o   *the use of modelling as the main strategy for developing ideas.*
o   *the use of "post-it" notes to conduct a review of work at the half way stage, using some of the ideas and principles of six hats thinking.*
o   *the regular incorporation of photographs of learners' work as it progresses.*
o   *the use of partners for peer assessment and in sharing and contributing to each others ideas*

Mike Laramy: design & technology department

2.8   Key Stage 2 (year 4/5) trial

In e-scape phase 1 we took the opportunity of trialling the 'light fantastic' activity in several schools at Key Stage 2. The ease with which young learners adapted to the use of the technology encouraged us to include at least one school in the national pilot for phase 2 and we report the results here. The school is Nancledra Primary School in Cornwall.

The activity and methodology remained the same as in the KS4 trials. All children completed the task and evolved a working prototype within the same time frame. The design brief and workbook were identical without any modification to language, size of font etc and the sequence of the activity ran in exactly the same way as we had used for year 10 learners.

The children with whom we worked were in Years 4 and 5 and the brief was complex for such young children. Having discussed it with them however they seemed immediately to understand it and began to address what they could manage. Inevitably some adjustment was made in the language used within the classroom context and to the timings within the sequence of the activity.

Principal findings from the KS2 trial
The trial with all the associated technologies appeared to work as successfully with KS2 children as it did with KS3&4 learners. Indeed the age of these young learners belied their mature use of the technologies. Whilst many of the uses of the technology were new to them they adapted to them effortlessly. As examples, two Year 5 children discovered for themselves elements of the drawing package (circles, smudge, airbrush and erasers) prior to being shown, and a year 4 child found out how to use capitals in the text tool and how to

playback and delete voice recordings without any adult assistance. These discoveries were soon shared with others.

Voice memo

The children were apparently very comfortable with the multi-functionality of the PDA with photographic, voice, writing and sketching capability. But it was most noticeable that they approached the voice recording activities with particular attention. Whilst they were not at all self conscious about talking into their PDA we can detect a difference in capability with these young learners. Whilst older learners responded by articulating sophisticated ideas such as '*easy to manufacture, stylish and attractive, target audience'* etc. younger learners tended to focus on their own experience of the struggle to create their working prototype *'very nearly finished - everything is working perfectly – I really want to test it out', 'its looking good'.*

The voice memo particularly enhanced the learning experience within the realms of appraising ideas. Learners appeared more in control of their reflections through this review process since they did not need to write anything. They could articulate the thoughts in their head and talk to the PDA as a surrogate teacher, explaining what they were doing, and perhaps why it hadn't gone according to plan.

'I don't know about the spring, it might not be long enough - I don't have the mechanism I wanted, it was way too complicated'.' I need to fix up the teeth but I don't know whether they will work because they're not even'. ' I somehow need to put the pills in but it is a bit of a problem, because as you put your wrist in they all fall out'.

Screen drawing

The sketching on the PDA tended (for most learners) to flatten the difference in technical quality of drawing/sketching between the younger and older learners. And the ideas expressed are impressively comparable across the age ranges.

Year 5 response …..   and a year 10 response

The small screen created some difficulties when the young learners needed to see the whole of their teammates work at once, since they did not seem so comfortable with scrolling the screen. It also posed a problem for some children when writing with the stylus - normally their writing/drawings would be significantly. The consequence of these combined difficulties can be seen below in an example of reflections on a teammates idea.

The relentless pace of the activity probably exacerbated this difficulty, as did the unreliability of the 'transcriber' tool when used in the context of emergent writers. The transcriber did not reliably convert handwriting that was not well formed.

Generating Ideas
The influence of the 'handling collections' was evident in the generation of ideas, which ranged across a mass of product types: pens, slides, necklaces, bracelets, key rings and headphones. And whilst young learners' proposals for the

mechanisms which were incorporated into their design might not have been quite as sophisticated as their older counterparts, they nevertheless considered mechanisms that swivelled, sprung, slid, that were propelled by fans and magnets and that allowed tablets to be removed one at a time.

Modelling and making
The activity required the use of the same soft modelling resources. These young learners' normal exposure to designing activities revolve around using 2D paper and pencil sketches and their concept of modelling would normally be confined to trying out how part of a product might work. Here we presented the challenge of creating a prototype in totality and there was some confusion about the difference between the prototype of the product and the product itself. It is only fair to report that this distinction was not always clear to year 10 groups, but it was marked with year 5. But we should note that this difficulty was

50

not about the PDA or the digital system more widely. Rather it was about the concept of design teaching and learning.

Collaboration

These young learners worked well both independently and particularly collaboratively; continually offering ideas and assistance to each other. In addition to the in-built collaboration of box 1/2/3 and the later team reflections, the depth of thought reflected in their spontaneous discussion showed a real awareness of the subtleties of the task:
 'What happens if she has a heart condition?'
 'How will she know when to take the tablet?'

Development styles

Just as with the year 10 groups, many styles of development were evident. For some, ideas popped up but appeared to go nowhere, for others good ideas emerged and grew into part solutions, some followed their ideas tenaciously, others saw complexities and made trade-offs, and some abandoned their risky ideas totally because they could not manage the technicality of making it work. Most of the group developed one idea and stuck with it throughout, only a few offering many alternative possibilities. This was evident in previous KS 2 trials of the paper-based activity (light fantastic), suggesting that the new technology did not impact fundamentally on their ways of working.

Overall evaluation

Despite the intensive 6 hours over which the activity spanned it proved completely 'doable' with KS2 learners. They were quite comfortable – and accomplished - when designing, sharing and discussing, modelling, collaborating and reflecting throughout the activity. The teacher – along with the head and other staff in the school - was equally enthusiastic about the potential of the PDA to enrich design learning & teaching. And it is perhaps appropriate to leave the final comments on the Nancledra year 5 experience to her and her class.

"They liked the task and were very excited by the PDAs"

"I was very pleased how the children stayed on task even thought they must have been flagging by the end. I think this was due to how the task was structured as well as the eagerness to do all things digital."

"I was amazed how quickly the children grasped the technology and were in no way overawed by it."

"I was particularly impressed with how they used the voice recordings and took them so seriously. I feel this has tremendous potential for capturing their thoughts and feelings about their work as they are doing it."

And she reports the final comments from the class:
"When can we do that again?"

2.9   Teachers TV

Through QCA, we were contacted by a Teachers TV company (Evans-Woolfe) with whom Richard Kimbell had worked on previous projects. They had heard of the trials and were interested to film them as part of a design & technology series that they had been commissioned to produce. They agreed to our conditions concerning the filming and we agreed to allow them to film the 2nd Cornwall trial in Saltash.net Community School. The school were enthusiastic about this and all appropriate protocols were followed.

The outcome has been two teachers TV films (15 minutes each); one focused on the e-scape activity over 2 days ('The future's hand-held') and the other is based on a series of discussions about how the activity worked and what its significance might be for the future ('The issues'). This 2nd film includes reactions and discussions with the learners, the teacher, another teacher from a different department, and the TERU team. Both of these films were broadcast for the 1st time on Monday 8th Jan 2007, the 1st at 5pm immediately followed by the 2nd at 5.15pm.

We were very pleased with the way in which Chris Guiver, camera-man/editor of Evans Woolfe, filmed and edited the e-scape activity into the two films. It was filmed very sensitively, without any significant interruption to the flow of activity. The learners just got on with their work and ignored the filming. As a result, Guiver was able to capture the authentic feel of the unfolding activity, with the learners responding straightforwardly to any questions he posed to them.

Having watched the two films, we were aware of the vast quantity of film that had been captured in the 2 days in Saltash but that Guiver had not been able to use in the relatively short TV programmes. Accordingly we commissioned a further programme from Evans Woolfe; a research-based documentary in which we describe and illustrate (through the extensive film shot in Saltash) the structure of the activity and the rationale for it. This programme runs to approx 90 minutes and is built around 8 'chapters' which vary in length, the shortest being approx 6 minutes and the longest approx 12 minutes. We are delighted with this DVD – since each of the chapters captures a subtle combination of rationale and exemplification. We include a copy of the DVD with this report

As we were completing the additional filming for this documentary, we were contacted by another production company, Real Life Media Productions Ltd, who wanted to film e-scape for a quite different Teachers TV programme. The focus of the programme was not on design & technology but on e-assessment. RLMP gained permission from Evans Woolfe to use enough of their film to portray the classroom/studio activity in Saltash and then we filmed more in the TERU offices – essentially debating issues of e-assessment.

The two Evans-Woolfe films were first screened on 8th Jan 2007
- *'the future's hand-held'* at 5pm
http://www.teachers.tv/video/3306
- '*new technology; the issues'* at 5.15pm

http://www.teachers.tv/video/3307

The Real-Life-Media film was first broadcast the following day,12.30pm 9th Jan 2007
• *e-assessment – where next?* (approx 30 minutes).
http://www.teachers.tv/video/5431

Once broadcast, the films are available for viewing and downloading through the Teachers TV website
http://www.teachers.tv/subject

Whilst the filming of these programmes has represented additional un-planned work for the team, we believe that it has been well worthwhile. First of course the programmes provide excellent dissemination to schools about the approach we have been developing and some of its potential and consequences in the future. But second, we have been provided with free access to a professional film crew who could record the whole e-scape process in detail. This has been a fantastically valuable resource, and the Evans Woolfe documentary in particular stands as a terrific reference product for the project.

2.10   the paper test

One of the research questions that we have been interested in with this project concerns the performance of learners when they design digitally **compared with** their performance when designing on paper. Accordingly we asked all schools if they would (in addition to e-scape) be prepared to run a paper-based test of the kind that we used in the previous project for *'assessing design innovation'*. In the end, the logistic challenges of making sure the e-scape pilot ran smoothly somewhat undermined our plans for the paper activity. For a number of reasons, we were forced to reduce the number of schools in which we could run the paper activity. The plan we settled on therefore emerged as follows.

During e-scape week 1, we took the paper-test resources with us, and after completing the escape activity we left them the paper resources to run that activity at some point of their choosing *within the following two weeks*. We then re-visited all the e-scape wk 1 schools, collected all the work and transported all the resources to the e-scape wk 3 schools. They then had two weeks to complete the paper activity *before we turned up to run the digital e-scape activity*.

This plan eliminated the Northumberland/Newcastle schools from the paper test. And we recognise that there may be a difference between those schools that did the e-scape activity *first* and those that did it *second*. It has been interesting to see the extent to which the experience of the paper activity (before or after the digital one) affected the performance of those involved in both tests.

The paper-based test activity is the 'light fantastic' activity developed for *Assessing Design Innovation*. (See App 2.4 and 2.5) It has been fully reported in the research report of that project and little more need be said here. The form of the e-scape activity was evolved directly from this Light Fantastic structure with differences only where the new technology provided some new and important opportunities (e.g. with the voice memos). Otherwise, the timings, structures, and resources for the activity are the same.

2.11 the e-scape web-site

At the end of the trialling process (May 06) and the national pilot (June/July 06) we have fourteen schools of data in the e-scape website. One of these was from the very 1st trial (Uplands Community College) and, as reported above, that activity had been so disrupted by technical failure that all the portfolios were incomplete and we concluded that they could not be used for assessment purposes. Accordingly, whilst we retain the data for research purposes, the learner performances will not be included in any analyses. Another school of the 14 is Nancledra County Primary School in Cornwall, and was part of the research simply to investigate the impact of the e-scape system with young learners – in that case year 5.

We therefore have a total of 12 schools for assessment at age 15 and a total of 249 portfolios. Considering that the maximum possible number of portfolios was 252 (21 learners in each of 12 school), we consider that being only 3 short (for absence and other technical reasons) is a substantial achievement and we are immensely grateful for all the efforts of the schools, our administrator colleagues, Handheld learning and TAG Learning. All the learner portfolios are in the website at the following address. http://212.100.251.115/e-scape/

Because of the nature of the data on the site – with real names of schools and learners and including some photos and their voice files - access to the site is restricted to the research team, Ian Williams at QCA and research staff at AQA and Edexcel.

The website is organised through three layers of interface. Initially access is through the moderator log-in screen demanding personal access codes; then through a 'schools' page and finally to a class list that identifies names, UPN and group numbers etc.

Double clicking on the school will open the class list for that school, and double clicking on a learner name on that list will open their
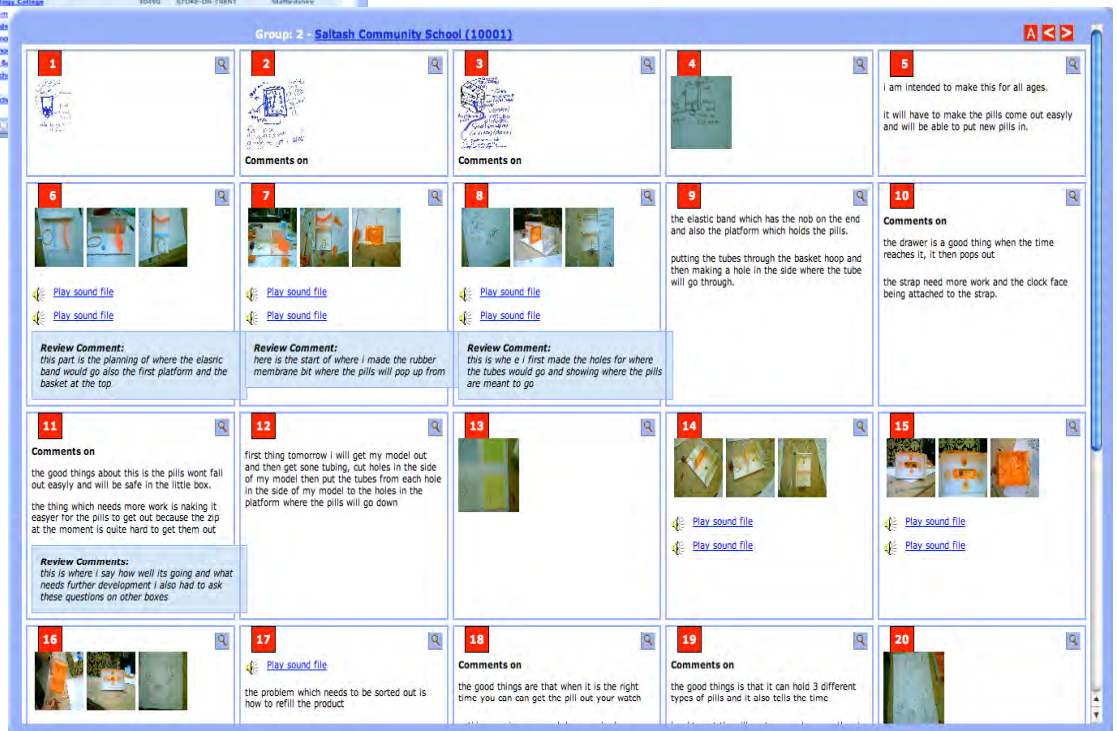
The portfolio is structured through the 23 sub-tasks of the 6-hour activity, with response modes of various kinds (drawing, writing, photographing and speaking) and with both individual and team-based purposes. Like the paper-portfolios that were the precursor to *e-scape*, these web-screens provide a very real picture of the learners' evolving prototype and their thoughts about it along the way.

This snapshot of box 8 illustrates the richness of these data. The three photographs show the drawing up to that moment and two photos of the model - from different angles. Clicking on the magnifying glass brings the images to full-screen size. The two sound files are the authentic recorded voice of the learner responding to two questions – 'what is working well?' and 'what needs further development?' – and together these provide a real insight into their understanding of their work. It is important to note that this 'photo and sound file' routine recurs throughout the activity – essentially once an hour for the 6 hours. At least three significant things result from this. First, many of them get better – more articulate – in describing their work and the circumstances surrounding it. Second, the routine – taken together – leaves a real-time visual/audio evidence trail that is quite unique in the assessment of performance in design & technology. Third, learners' approach to the task is enriched as they are more motivated, braver (take more risks), and think more deeply about what they are doing.

Finally, the review comment (below the sound files) is a reflection by the learner made at the *very* end of the 6 hours of activity. Looking back over their whole work, we invite them to think about what they might have done differently if they had known then what they know now. Sometimes these meta-cognitive responses are descriptive – as in this case – and sometimes they are deeply analytic of their own performance.

We are not aware of any equivalent system of real-time, dynamic, e-portfolio assessment for any subject in any country. We believe this to be a world first. The 249 rich portfolios that inhabit the website have now become the focus of our work in the project. And the 1st challenge was to develop an approach for the assessment process.

2.12   an approach to assessment for e-scape

At the outset of the project we had assumed that the assessment process for e-scape would follow broadly the same principles that we had used for *Assessing Design Innovation*. In short, we had developed a set of criteria for assessing performance initially on a holistic scale (1-12) then in relation to four broad categories of design-innovation performance:
• Having ideas
• Growing ideas
• Optimising ideas
• Proving ideas
The complete assessment form is reproduced in Appendix 2.6

However, during some early discussions of our web-based portfolios, our attention was drawn to the work of Alistair Pollitt – formerly the research director for the University of Cambridge Local Examinations Syndicate. He had been advocating an alternative form of assessment that was particularly appropriate for *performance* disciplines because it is based

on judges making *whole judgments* about pieces of work rather than allocating marks to criteria and them adding up the result.

We have long argued for the pre-eminence of holistic assessment in design & technology and in 1991 had demonstrated through APU data that such holistic assessment could be very reliable with suitably trained assessors. It was for this reason that the holistic judgment was the 1st judgment to be made with the form discussed above. The resulting assessment was not to be seen as the arithmetic sum of the 4 part-judgments. Rather the four part-judgments were used to illuminate the nature of the performance that had already been judged holistically.

Pollitt's approach is based on a theory of judgment derived initially from Thurstone (1927) and based on what we might call 'differentiated pairs'. The assessment process involves making a single judgment about which of two portfolios is the better, and – if enough judgment of this kind are made (and by enough judges) a rank order of performance can result. After a series of discussions with QCA, Edexcel, AQA, Pollitt and the TERU team, we decided on a trial of the system using our existing archive of paper-based performance from the *'Assessing Design Innovation'* archive.

The case for the approach – and the outcome of that assessment trial – have been written up by Pollitt, and we include here his report to us.

Grading learners' work

The fundamental requirement of summative examinations is to judge the overall quality of learners' work in some educational domain on a standard ordinal scale. Usually, in Britain, the scale is then divided into bands to indicate relative grades of performance. In 1792 a Cambridge tutor, William Farish, introduced the idea of attaching numbers to the various parts of a learner's performance to indicate their quality and aggregating these to give an overall mark. Two years later he was elected Professor of Chemistry and the tradition of marking examinations was established. Over time it spread from its origin in the sciences to be used in every subject.

But while scoring and aggregation seems to suit those examinations composed of many small questions examiners in some other subjects, where they want to assess created objects or performances, have often used the method only reluctantly. In the assessment of 'writing', for example, the debate between advocates of analytic and holistic approaches has never been resolved; it appears that examiners of Art at GCSE and A Level apply the method 'backwards', assessing the overall result first and then applying 'suitable' numbers as marks. An examination that assesses design ability belongs to this group where marking may be inappropriate.

*Given the Fundamental Requirement, it is not obviously necessary that exams should be marked. The requirement is to find some way to **judge** the learners' performances in order to create the scale that is needed, and marking items to add up their scores is just the way we seem to have chosen to do this. An alternative method does exist, in which the examiners are asked to make holistic judgments of the quality of learners' work.*

In the words of a recent book on the psychology of judgment, "There is no absolute judgment. All judgments are comparisons of one thing with another" (Laming, 2004). In other words, all judgments are relative. Since 1995 almost all (non-statistical) studies of examination comparability in England & Wales have used a method of relative judgment (Pollitt, 1994), in which examiners are asked to compare pairs of 'scripts' from different exam syllabuses, simply reporting which is the 'better' of the two.

These studies include:
A Level or AS
Geography, Mathematics, Chemistry, Biology, Accounting, Psychology, Sociology, English, History, Media Studies
Vocational Certificate of Education/A Level
Health & Social Care, Business Studies
GCSE - International and UK
French, Dutch, German, Afrikaans, English, Mathematics,

World Class Arena - Mathematics

Cambridge Proficiency in English – ESOL
Speaking - Oral interview

Key Stage 3
English

Whether it is used for studying comparability or simply for grading a single examination, the method is based on the psychophysical research of Louis Thurstone, and specifically on his Law of Comparative Judgment (Thurstone, 1927). The essential principle in this law is that, whenever a judge compares two performances (using their own personal 'standard' or internalised criteria) the judge's personal standard cancels out. The greater the true difference between the quality of the two performances the more likely it is that the better one will win each time they are compared. Thus a large set of comparisons does more than just generate a rank order; the relative frequency of success of one performance against another also indicates how far apart they are in quality.

Statistical analysis of a matrix of comparative judgments of 'scripts' can construct a measurement scale expressing the relative value of the performances. The result of comparisons of this kind is objective relative measurement, on a scale with a constant unit. Furthermore, if a few scripts that have already been agreed to represent grade boundaries – perhaps from a previous sitting of the examination – are included in the comparisons, the whole process of marking, grading and comparability of standards can be replaced by the collection and analysis of paired comparative judgments.

The trial exercise

To confirm that Thurstone's method is applicable here, a preliminary exercise was carried out, using 20 scripts which had previously been marked in the traditional way. Members of the research team each judged about 40 pairs of these scripts, in each case reporting just which of the two better met the assessment criteria. The results are summarised below, with the numbers representing the scripts (in a random order).

RASCH ANALYSIS using the PAIRED COMPARISON model
******* Goldsmith's-Edexcel : Pairs test *******
Plot of Parameter Estimates

```
| 3|      -08- -17-
| |
| |
| 2|      -03- -15-
| |      -14- -20-
| |      -02- -09-
| 1|       -16-
| |      -07- -11-
| |      -04- -05-
| 0|       -19-
| |
| |       -10-
|-1|
| |       -13-
| |
|-2|
| |
| |
|-3|
| |
| |       -06-
|-4|      -01-
| |      -12-
| |      -18-
|-5|
```

The scale runs from a low of –5 to a high of +3; the average of the 20 scripts' parameters is 0.00. To interpret this scale it needs to be "anchored", as mentioned above, with a few scripts that have already been graded. The scale reliability was estimated to be 0.92, at least as high as would be expected in a GCSE marking study.

To confirm the results of the analysis the scripts' parameters were plotted against the marks previously assigned to them.

$$y = -6.419 + 2.33x - .249x^2 + .01x^3$$
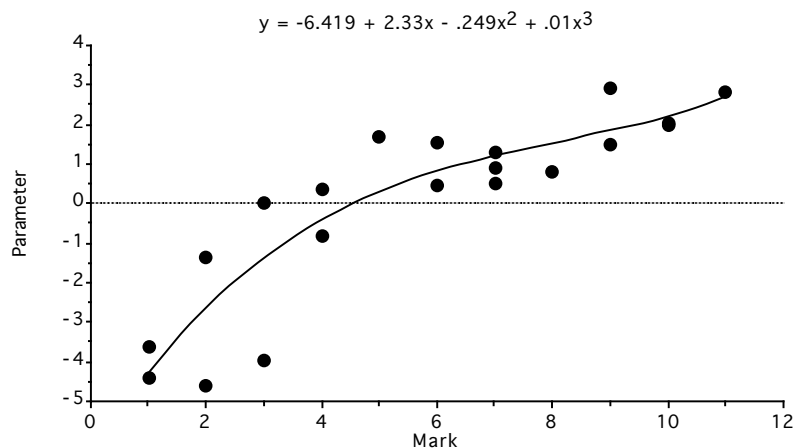
As expected, there is a strong but non-linear relationship between the parameters and the marks. (The relationship is expected to be non-linear because the mark scale is bounded, with a minimum of 0 and a maximum of12, while the parameter scale runs from -∞ to +∞.) The value of R^2 was 0.81, corresponding to a correlation of 0.90 between two linear variables, as high as could be expected in a case like this.

We should make it clear that the trial outlined above led to the creation of data sets that were separately analysed by Alistair Pollitt (using his system) and Malcolm Hayes (using the Edexcel system). Reassuringly, both sets of analysis produced the same results. The outcome of this trial was very interesting. Not only was the resulting rank-order virtually identical to that which we had derived from the conventional assessment process, but moreover the six 'judges' who had been involved all felt that the holistic comparative judgment process was both easier and intuitively more appropriate than trying to allocate numbers to parts of the portfolios and then summing these numbers.

A final consideration was based on purely practical matters of assessment manageability. The issue arises from the question … 'If comparative pairs is so clever, why aren't we using the system already for large scale assessment?' The answer to this question is all about manageability, and it was clearly exposed during the trial process outlined above. We had 6 judges and 20 pieces of work and we all sat at a big round table. Judge 1 was looking at script 15 and 5; judge 2 was looking at 17 and 2, and so on. Soon more than one judge needed the same piece of work and just had to wait till the other judge was finished with it. And by then the scripts had all got jumbled up in the middle of the table and did not come easily to hand. If this was a problem with just 20 scripts, imagine the challenge of a 100 or 1,000 or 50,000 scripts. The distribution process alone makes the process of repeated comparative judgments (by different judges) quite unmanageable. But the situation changes dramatically when all the portfolios are in a website. There, *every piece of work* is available *all the time*, *simultaneously,* and for *any of the judges.*

Accordingly, following the success of the trial, and after further debate, we agreed that the approach to assessment with the e-scape portfolios should be by using Pollitt-style comparative pairs judgment, and we then set about designing the system. It is important to recognize that the process has never been used before for 'front-line' assessment. Its use has been restricted (for manageability reasons) to inter-board research studies of comparability, and such studies are based on only a handful of scripts.

Pollitt proposed 16 as the basic number of comparisons that need to be made for any script; i.e. each portfolio is compared with 16 other portfolios. Moreover some of the judgments would need to be repeated by different judges. After a good deal of debate we settled on a 3-cycle approach using the 7 judges that had agreed to take part in the process.

• Round 1 would involve each judge in making 140 comparisons. The outcome of the resulting 980 or so pairs (i.e. involving 1,960 viewings of scripts), would be an approximate rank-order based on about 8 comparisons per script. Some of these judgments would be very easy as some of the comparisons would be excellent work compared with poor work.

• Round 2 pairings (another 140 each) would be drawn up using the approximate rank-order from round 1. In this round we would no longer encounter those big differences, but rather the pairs selection would be focused on closer judgments to refine the rank-order.

• Finally round 3 would be used to target notional grade boundaries. We decided that we should model the Awarding Body awarding process, using round 3 to firm up the data at the boundaries between notional grades.

To make the pairs comparisons somewhat more manageable for the judges, a system of 'chained-pairing' was adopted. So, for example, judge 1 might be asked to undertake the following pairings:
25 : 210
210 : 77
77 : 125
125 : 48

Having got into the work of the 1st pair of scripts, and having made a judgment about which is better, I put away No 25 and just open No 77 to compare with No 210 with which I am already familiar.

The judging system was developed through August and early September 2006 and we ran a training session in TERU (8th September) for all the judges to become familiar with the website and how to navigate around the work within it. Each judge was issues unique access codes for the website and during the training we examined several pairs, shared our thoughts both about the work and about the process of arriving at a decision.

The first round of judging was undertaken in the last two weeks of September. Pollitt undertook the resulting analysis and returned new pairings for the team. Round two then took place in the first two weeks of October. At this point, having analysed the resulting data, Pollitt suggested an alternative strategy for round 3. Instead of doing more pairs, he asked whether we would be able to examine (say) 6 pieces of work and place them in a rank order. We believed that this would be possible and accordingly round 3 involved (for each judge) a few more pairs and then two sets of 6 to be ranked. Each ranking of 6 pieces provides an equivalent amount of data to 15 paired comparisons, and we were confident that each set of 6 would not take the same length of time that 15 pairs takes.

2.13   e-scape performance analysis

The judging was finally completed at the end of October. Pollitt undertook the analysis and prepared a report for us on the resulting rank order and distribution of performance. We include that report here in full.

Paired-comparison scale creation for *e-scape*: a report from Alistair Pollitt: **8th November 2006**

Based on 249 portfolios: 12 schools: 7 judges: 3 rounds of judging:

Analysis of the judgment data

This exercise was a simplified, and rather crude, simulation of how an operational system would work. Instead of being continuously interactive, it consisted of three discrete rounds of judgments. In the first round 8 comparisons were made for each portfolio against other portfolios chosen randomly. In the second round a further 8 comparisons were made for most portfolios; the pairs for these comparisons were matched closely, using the results of an analysis of the first round data. The data from this round were added to the first round data and the whole set analysed again. The outcome of this was a scale that looked satisfactory for awarding purposes. A smaller third round of comparisons was added to improve the measurement quality around the putative grade boundaries, and to explore a possible improvement for future applications.

The principle behind the analysis method is that, when a portfolio A is compared to a portfolio B, the probability that A will be judged better than B depends on how much better A 'truly' is. If A and B are not very different in 'true' quality the chance of A winning will be close to 50%, but if A is much better than B its chance of winning will also be much higher. Given a set of results of comparisons between lots of portfolio pairs, the analysis finds the value for each portfolio that best explains all these results.

The set of values, one for each portfolio, establish a measurement scale. As with temperature scales or the A Level Grade/UMS scale, this scale can be transformed arithmetically for convenience. In the report below, five notional grades have been defined; the scale has been transformed so that all portfolios with values below 2 are Grade 1, those above 2 but below 3 are Grade 2, and so on. There are, of course, other ways in which the grade boundaries could be set, but these boundaries illustrate what the method could produce.

Before presenting the results, it is worth looking at the analytical report. It begins by reporting the percentage success for each portfolio ("object"):

Wins and Losses for each object

```
-----------------------------------------------------------------
Object |          Wins  Losses  Comparisons     %
-----------------------------------------------------------------
   1   |       1    10     12        22        45.5
   2   |       2    10     26        36        27.8
   3   |       3    10      6        16        62.5
   4   |       4    13      7        20        65.0
   5   |       5     6     10        16        37.5
   6   |       6    12      4        16        75.0
  . . .
 247   |     247     9      7        16        56.2
 248   |     248    12      6        18        66.7
 249   |     249    18      2        20        90.0
-----------------------------------------------------------------
     Total number of comparisons = 2322
```

These percentage scores are the starting values for estimating appropriate values for each portfolio: because some of them were mostly compared to 'better' portfolios and others to 'poorer' we cannot use the percentages directly. After correcting for these differences in comparators, the program reports a list of values:

Estimates of object parameters
```
--------------------------------------------------------
Number      |  Value   S.Error | Object Name
--------------------------------------------------------

  1         | -0.716   0.484   | 1
  2         | -5.543   0.398   | 2
  3         | -0.026   0.717   | 3
  4         |  0.987   0.499   | 4
  5         | -2.953   0.706   | 5
  6         |  2.838   0.664   | 6
. .
 247        | -0.335   0.586   | 247
 248        |  0.094   0.569   | 248
 249        |  4.788   0.776   | 249
```

At this stage the average value for all portfolios is, by definition, 0.00. After a simple rescaling, as described above, the report is:

Estimates of object parameters
```
--------------------------------------------------------
Number      |  Value   S.Error | Object Name
--------------------------------------------------------

  1         |  3.02    0.35    | 1
  2         | -0.42    0.62    | 2
  3         |  3.52    0.51    | 3
  4         |  4.24    0.36    | 4
  5         |  1.43    0.50    | 5
  6         |  5.56    0.47    | 6
. . .
 247        |  3.30    0.42    | 247
 248        |  3.60    0.41    | 248
 249        |  6.95    0.55    | 249
```

These are the final values – equivalent to the final marks – for each portfolio. The first digit indicates the grade (although anything less than 2 is considered Grade 1, and anything greater than 5 is a Grade 5). Note that portfolio 2 was judged to be extremely poor, and portfolio 249 extremely good.

In addition to Value or Grade, these reports give a 'Standard Error' for each portfolio. Traditional marking fails to report these, although assessment specialists are all aware that the assessment process *should* acknowledge the amount of uncertainty in any exam result. It's hard to generalise, but the standard errors for GCSEs are probably between 0.5 and 1.0 grades for most subjects. For this analysis the average standard error is 0.46 grades, which seems acceptable: if we exclude the 30 best and 30 worst portfolios the average standard error is 0.41 grades. Of course, there are just five grades in this exercise. However, increasing this to nine grades would mostly involve splitting the two extreme grades – Grade 5 would give A* and A, Grade 1 would give F, G and N/U – and the middle three grades would only be expanded to four, giving us an average standard error still only about 0.55 grades for most candidates.

The complete set of results can best be seen in the diagram below. In it, the portfolios have been sorted into order and are shown with their standard errors. In formal statistical terms 68% of the portfolios' "true" values will lie within one standard error of the reported value. (This is the basis for Paul Black's much-quoted remark that "They give the wrong grade or level about 30 per cent of the time.")

Vertical lines are drawn through the grade boundaries to show how many learners would fall into each grade. Note that Grades 2-4 are, by definition, equal size, and that this leads to more learners lying in the central grade than the ones either side of it.

Plot of Values with Standard Errors
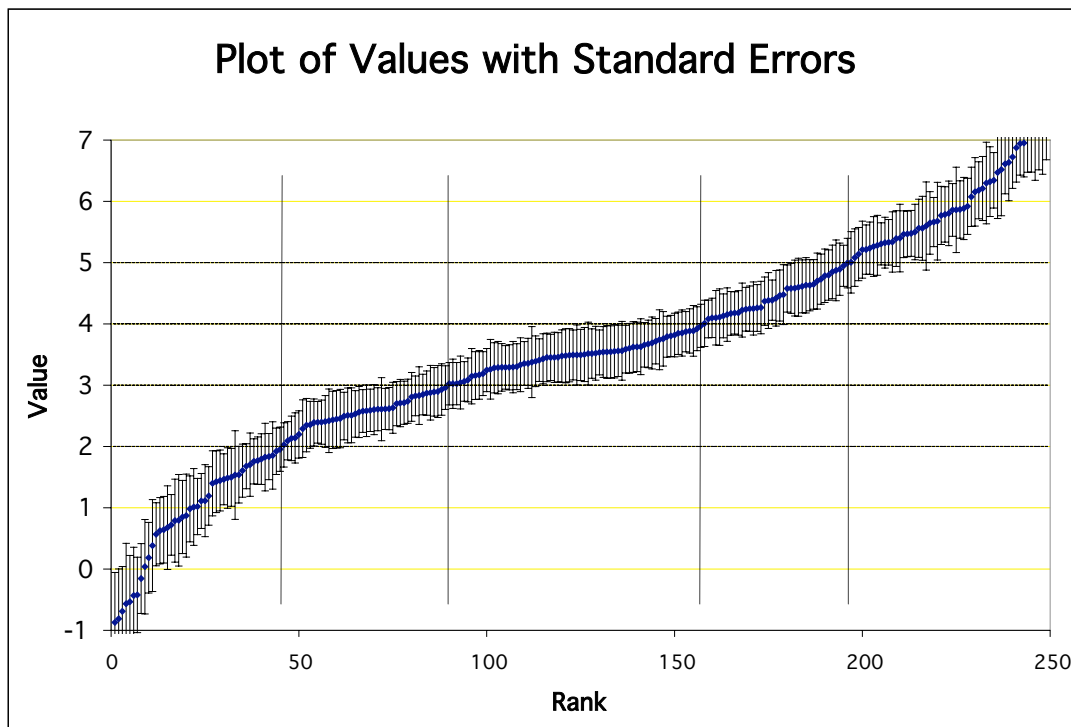
The diagram also shows the effect of Round 3. This concentrated on the portfolios close to the grade boundaries at the end of Round 2. Each of them was put into 2-5 extra comparisons, in addition to the 16 (usually) they had already been in; as a result their standard errors are smaller than average. The effect is not large, but only because Round 3 was a rather small exercise. In an operational system more of these 'borderline' judgments would be made so that as many as possible of the portfolios would be assigned with high confidence to the appropriate grade. Note that this focus on borderline portfolios is not possible in a mark-based system, unless we go back to the procedures that used to be applied – rather unsystematically – after the award was complete. In this system 'borderlining' can be applied routinely, and as much or as little as is desired.

The analysis of the judgments also gives a traditional indication of the quality of the measurement process:

Summary of the Scale properties

Standard deviation of Object parameters	=	2.458
rms estimation error	=	0.646
Separation coefficient (sd/se)	=	3.807
Reliability coefficient (like alpha)	=	0.931

Assuming these represent a population 6 sd's wide, and that bands 3 se's apart are distinguishable, then there are up to 7.95 reliably distinct bands of objects.

The key figure here is the reliability coefficient of 0.93. This figure allows for unreliability between markers *as well as* for lack of internal consistency within the examination – most traditional reliability coefficients only allow for one of these. Only a few current GCSEs are likely to be as reliable as this if we consider both sources of unreliability.

Finally, the report checks the consistency of each judge and each portfolio. The report on the judges is:

Summary of Fit Statistics for Judges

Number	Count	MnResid	UWMnSq	UW-z	WMnSq
1	307	0.25	0.82	0.02	0.80
2	324	0.30	0.70	-0.25	0.83
3	324	0.28	1.18	0.48	0.99
4	326	0.29	0.70	-0.98	0.86
5	333	0.29	0.82	-0.26	0.85
6	323	0.32	0.84	-0.01	0.94
7	325	0.26	0.60	-0.55	0.79
Mean:		0.31	0.83	-0.12	0.89
S.D.:		0.07	0.17	0.48	0.08

The column 'WmnSq' (weighted mean square) is the most important. A figure greater than the mean plus twice the S.D. – in this case 1.05 – would indicate cause for concern, a judge who was not behaving in the same way as the others. None of these judges fails the test, but Judge 3 should be monitored for a while in future to make sure that he does not drift further away from the others.

The similar report for the portfolios finishes:

Summary of Fit Statistics for Objects

Number	Value	MnResid	UWMnSq	UW-z	WMnSq
. . .					
246	4.68	0.78	0.47	0.71	0.83
247	-0.34	0.59	1.62	0.86	1.53
248	0.09	0.57	0.54	-0.10	0.72
249	4.79	0.78	0.66	0.73	1.05
Mean:	-0.00	0.63	0.80	0.41	0.85
S.D.:	2.46	0.15	0.74	1.13	0.23

It shows that a few of them ought to be checked (at least 2 of the 249). The criterion would be 0.85+2*0.23, or 1.31; portfolio number 247 exceeds this, suggesting that there is something about it that is unusual enough to warrant a further look – perhaps different judges valued them in different ways.

There are three nuggets in this report to which we would - in particular – draw the attention of readers, quite apart from the performance scale itself.

First the *reliability* of the resulting scale.
"The key figure here is the reliability coefficient of 0.93. This figure allows for unreliability between markers as well as for lack of internal consistency within the examination – most traditional reliability coefficients only allow for one of these. Only a few current GCSEs are likely to be as reliable as this if we consider both sources of unreliability." (see Pollitt above)

But this reliability is hardly surprising. Each piece of work has been compared with many others, and the judgments had been made by many judges. Any idiosyncratic judgments were soon outweighed by the weight of opinion of the team. The process is almost inevitably more reliable than current GCSE practices, where much of the work is assessed by the teacher alone, or at best by the teacher and one external moderator.

Second it is important to note the consistency of the judges. In this comparative pairs approach, the analysis automatically produces a reading of the judging team, specifically concerning their consensuality. The system notes how often – and by how much – my judgments are at variance with the other judges and in the end produces a mean score for the whole sample. If I am more than two Standard Deviations from that score, then I am a cause for concern.

"None of these judges fails the test"   (see Pollitt above)

Third, the system also automatically produces data on the consensuality of judgments applied to individual portfolios. Reference to the 'plot of values' (above) shows some portfolios with much longer standard error 'tails' than others. These are the portfolios over which there was a considerable amount of disagreement within the judging team. In the process, the system automatically highlights the pieces of work that need closer attention.

*"It shows that a few of them ought to be checked (at least 2 of the 249). The criterion would be 0.85+2*0.23, or 1.31; portfolio number 247 exceeds this, suggesting that there is something about it*

that is unusual enough to warrant a further look – perhaps different judges valued them in different ways." (see Pollitt above)

These three are all automatic virtues of the comparative pairs judging process.

2.14 the response of the judging team

The judges responded to the process of undertaking the judgments, and we captured these thoughts in a feedback form.

We were interested initially in the time that the process takes, and there was a degree of conformity on the matter. Round one was generally agreed to be the toughest and the early pairs (say the 1st 20 or 30) took as long as 10 minutes per pair to decide, but gradually we got quicker. This speeding-up resulted in part from being more skilled in working our way around the portfolio, and in part from the fact that the pairings inevitably threw up repeats. Having got properly inside a piece of work at the 1st time of asking, it took only a much briefer scan 2nd time around to remind us of its qualities. By the end of the 140 pairs we were typically doing each pair in 2 minutes. On average, for the whole sample, we can estimate 4.5 minutes per pair, amounting to approx 10 hrs for each judge. The 1st round therefore used 70 hrs of judging time to produce a rank order for 249 portfolios. Put another way, each portfolio took approx 17 minutes to locate into the rank order.

The pairs in the 2nd round were closer together but the relative difficulty of these round 2 decisions was offset by the familiarity (by now) with much of the work. Generally round 2 was quicker than round 1. Round three was a very limited scrutiny of the grade boundaries and was relatively easy and quick.

The wider feedback from judges about the process of undertaking the three rounds; about their attitudes to it and about the things that were easy/hard about it are summarised below.

Concerning the concept of comparative pairs judging

- *..it made you look at the work in a different way, not in the context of marks or abstract qualities, but forced you to consider it in the round*
- *Collaborative (lots of markers), and pairs decisions .. ameliorates the guilt re getting it wrong re destiny of learner. Pairs decisions much easier than holding lots of portfolios in head. Pairs decisions lends itself well to holistic 'marking'.*

Concerning the difficult things about using the system

- *The first (round) because each set of judgments was new for a large proportion of the round and also the first round was like a personal training exercise, where one familiarised oneself with what one considered to be valuable in a candidates work.*
- *..valuing different sorts of poor performance (i.e. gave up on an idea with potential and went backwards against poor plodders throughout)*
- *Where two pieces of work were very close it was sometimes difficult to find a clear piece of evidence to separate them.*
- *Being prepared to 'reject' a good portfolio in comparison to an even better one, and conversely having to 'reward' a poor portfolio over an even worse one.*

Concerning the easy things about using the system

- *Everything really, except what I found most difficult - see above!*
- *Where one was a 'wow' and one was 'going nowhere'*
- *The final round and also as soon as the candidates work started to repeat.*

Concerning the process of making judgments

• if it was not immediately obvious which was better I looked for 3 characteristics (growth, grip and goodness) First by comparing boxes 1, 4, and 20 in more detail, if still no clear winner, then box 6, 16 and 17, and finally looking at all the reflective comments in detail.
• Worked through all in each script from start to end using > button, including listening to all sound files. Didn't usually need to look back at the earlier of pair. Did this even if it appeared at first holistic scan to be pretty obvious – (a) was very interesting to see and hear all and (b) felt might miss a deciding gem if didn't and (c) I think paid off when got to the repeats section. (RW)

Concerning comparisons with traditional marking
• Not having to allocate a series of smaller numerical awards (e.g., 3 out of 5, 8 out of 18, etc), adding them up and going back to adjust them to get the grade you first thought of.
• In comparison to marking previously unseen paper-based 'long-project' portfolios, or even day-to-day projects / homework this is so much faster and feels much fairer because less likely to miss vital bits of evidence – plus the process *is recorded (as real, not as retrospective concoction) – plus the voice bites often help.*
• I would argue that the judging criteria could be created by consensus early on in the process, but do not have to be in place right at the beginning. This is a very different approach to a conventional marking system which is dependant on an agreed mark scheme being in place from the outset. Also that non-consensual judgments are cancelled out. Thus the system is self regulating.
• I think it forces you to have different sorts of conversations with the portfolios. If you have one standardised marksheet, then you are forced to have the same conversation with every piece of work (Which may or may not be appropriate for you or the work) I think each comparison has the potential to be a different sort of conversation, some are very obvious, short and precise. Others are more demanding, deeper and longer in order to build up the confidence to toss a coin!

There is far more feedback data from the judges than we have reported here, and these snippets are included merely to illuminate the process from the inside. Other sections of the feedback data (particularly concerning the wider potential of the system in the future) will appear later in the report under the 'issues arising' from the project.

2.15   assessing 'Light Fantastic'

It was part of our research design to have as many as possible of the e-scape sample additionally undertake a pencil/paper activity. The point of the exercise was to compare their performance when designing on paper with their performance when designing digitally. We have explained earlier the logistic problems that resulted in only being able to do this with a sample of 4 schools (84 learners), but none-the-less the plan went ahead and we collected all the work from the four schools.

The marking procedure was a blend of the original approach to assessing 'light fantastic' as part of the previous project (*Assessing Design Innovation*), and the approach we have used for e-scape.

In a nutshell, we used the original assessment rubric (see Appendix 2.6), but only as far as step 1. We divided the scripts into 2 sets and Tony Wheeler marked one and Richard Kimbell the other. This was done partly be rank-ordering the pieces in our respective sets and then deciding where to latch the rank order onto the quantitative scale. Since this was done separately, we had in effect created 2 latched rank orders (TWs and RKs) and we then spent a considerable time moderating the two to reconcile any disagreement *either* about the ranking *or* the latching onto the scale (or both). In this way we created a simple holistic mark on a scale of 1-12 for each of the 84 'light fantastic' pieces of work.

2.16   illustrating performance

The research team at TERU have explored the nature of performance in design and technology in a number of previous projects over the past 25 years. Comprehensive explanations of the principles behind this work can be found in *The Assessment of Performance in Design & Technology* (Kimbell et al 1991) and more recently *Assessing Design Innovation* (Kimbell et al 2004)

These prior projects have described the wide range of **procedural competences** necessary to make effective progress towards the resolution of design challenges. Furthermore they describe two characteristic qualities of performance within this overall process, a **reflective** ability, allowing us to think around the task, and an **active** ability, allowing us to take action in response to the task. These two qualities are linked though the critical quality of **appraisal** and together the three qualities account for the iterative process of 'to-ing and fro-ing' between thought and action. The interaction of these three qualities formed the basis of a holistic assessment framework that has underpinned much of our subsequent work.

Assessing Design Innovation, the precursor to the ***e-scape*** project, extended the team's original procedural assessment framework to focus more specifically on innovation (design **ideas)** and **group work,** two qualities that have been significantly undervalued in most assessment systems for design and technology in school. Ideas, rather than outcomes, lie at the heart of the creative process, and we developed this to embrace *having* ideas (sparkiness), *growing* ideas (modelling) and *proving* ideas (criticality).

The key purpose of the e-scape project has been to explore the potential of digital enhancements to the existing paper-based assessment systems developed in previous projects. Even though the technology has changed, the tools and format of both the evidence we can collect of learners' performance, and the processes we use to judge it, we have worked throughout to the same procedural and assessment principles. Rather than reiterate the qualities of performance here, we feel it is more important to describe some of the key differences and similarities between performance in the paper and digital activities. The following section is divided into comparisons between the four main media types available on the PDA, sketching, photography, text entry and audio notes.

Sketching using the PDA
Given that in total we only allowed about 15 minutes for learners to familiarise themselves with the drawing tools on the PDA, their early ideas sketches in box 1 were rich, varied, and overall appeared much more dynamic than the equivalent performance we have observed in box 1 in previous responses to these timed activities on paper. This was in part due to the availability of a range of easy to use tools in the PDA paint package, but would also seem to indicate learners already have a high level of familiarity with, and ability to transfer skills from, other digital imaging tools.

Rich example of box 1 from a paper test
• Typically learners worked in black pencil on the paper test, the use of colour was rare.
• Most combined a mixture of notes and sketches, with some learners choosing just to write and others just to draw.
• Notes, where they were used were always on the same plane.
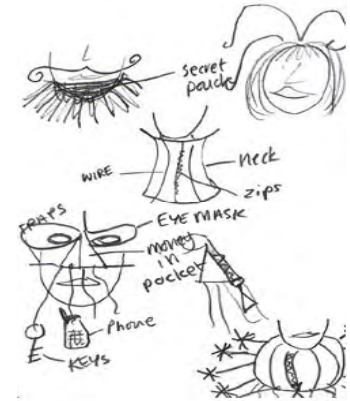
Rich example of box 1 from PDA test
• Typically learners worked with a black line tool but on the PDA most included some colour, either as washes or coloured lines.
• As in the paper test most combined a mixture of notes and sketches, with some learners choosing just to write and others just to draw.
• Notes and sketches were often added on different planes (as in this example) as learners rotated the devise and worked around the screen.
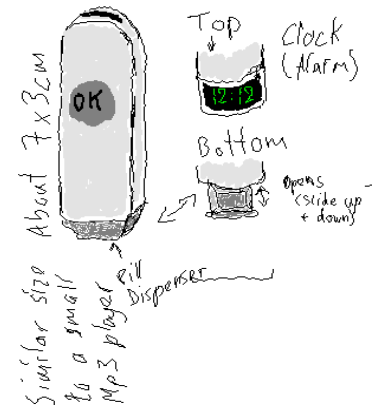
Typical example of box 1 from a paper test
• Many learners did not make significant progress in box 1 on the paper test. It is important to remember that in isolation these early ideas are not a significant indicator of overall capability, this example was from one of the highest scoring paper scripts.

*Typical example of box 1 from PDA test*
• Similar box 1 responses can be seen on the PDA from high ranking portfolios, the only difference being the use of colour.

Different sketching styles

Aware of the standardising effect of some computer imaging programs, where restricted tool sets can make everything look very uniform, but also that complex applications are difficult to master, we were concerned to get the balance right with the e-scape sketching tool. We were very happy with the results from the *Pocket Painter* application that presented a basic set of tools that were easy to understand and quick to learn for novices. For those with more experience, the package also provided features and effects that allowed them to take the representation of their ideas further. We were impressed by the diversity of presentation / visual / communication styles that were evident across the work in box 1 on the PDAs; far more than in the paper test.

combination of notes and sketches

Most of the work in box 1 on the PDA presents
a combination of sketches and notes
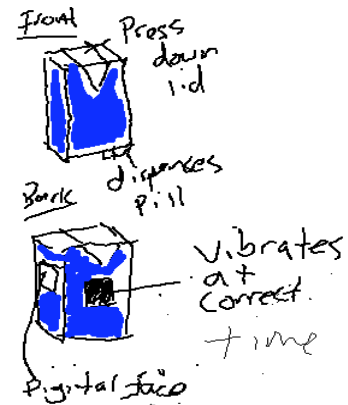
just sketches

It is possible to find box 1 examples in both the
paper and PDA tests where learners have only
used sketches with no notes, but these are not
typical, and it could be that the learner ran out of
time before getting round to annotating their work.

just notes

In the paper version of the test we noted a number of
responses to box1 where learners had just used
notes with no sketches. Interestingly while there are
PDA box 1 examples that are predominantly text,
there do not seem to be any with no sketching at all.
This might be because while writing legibly is
relatively easy on screen, the style of the handwriting
is not as precise as it is on paper, which may have
discouraged some learners from writing.
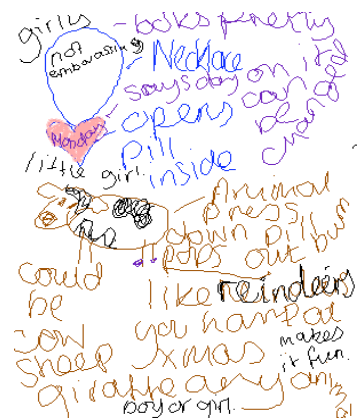
use of colour

With very little training and only a few minutes to jot down their ideas, learners made
surprising, appropriate and dynamic use of the various colour tools to enhance their
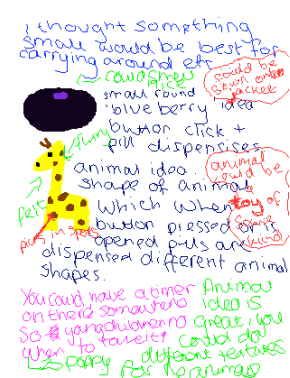sketches and notes in box 1.
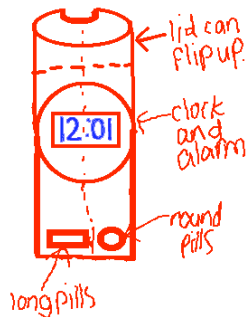
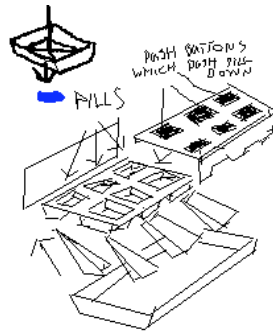*coloured backgrounds*   *coloured fills*   *coloured lines*   *coloured text*

use of tools

The use of the different paint tools allowed learners to create dramatically different responses, from straight-line technical drawings, to more freeform artistic responses.

graphic tools   *perspective drawings*   *paint effects*   *fill tool*

number of ideas

Just as we had seen in the paper tests, some learners started by jotting down lots of different ideas, while others focussed on a single idea and developed this in more detail.

one idea in more detail   *lots of hazy ideas*
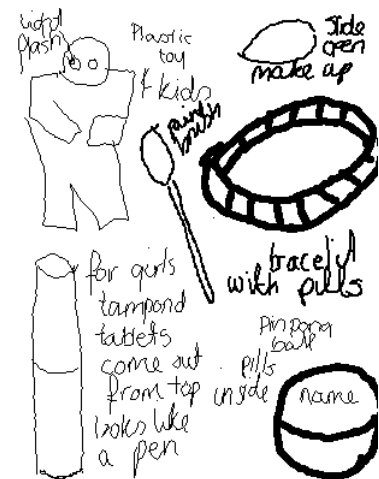
In reviewing learners' use of the PDA to collect and develop their early ideas and the quality of the sketches they produced, our overall impression of the sample is that the digital tools did not appear to hinder the development of ideas in box 1, indeed the range of drawing tools available has facilitated a much richer and more diverse set of responses than we typically saw in the paper tests.

Collaborative sketching and team support

We should remember that the purpose of collaborative work in e-scape is **NOT** so that we can assess group performance. The purpose of the group-work sessions is simply to support the on-going work of individuals. The group operates as a group pf peers supporting each other in their individual pursuit of solutions to the task. The previous paper based projects that we have undertaken showed conclusively that learners valued highly the opportunity to collaborate both in the generation of ideas and in their reflections on them. Early trials with e-scape confirmed and extended the potential of learners working together to support individual performance, establishing the power of zapping files between learners using the various communication technologies (IR, Bluetooth and wireless).

We followed the format of the paper tests as closely as possible and organised learners into teams of three. They worked on their early ideas individually and after 7 minutes the e-scape system sent the work of each learner on to another learner in their team. This new piece of work appeared on their PDA with the instruction to imagine this was now *their own* work and they had to take it forward as if it were their own for the next 5 minutes, adding any constructive criticism and new ideas they could think of. In the paper version of the task there are 3 early ideas boxes (one for each team member) and as they work each subsequent team member can see what the former has added.

As each paper box is the full size of the PDA screen we judged it would be difficult to replicate the paper experience by zooming in and out, making it difficult to keep what the other teammates had done in mind. Instead, we decided that the e-scape system would create 3 separate drawing layers on the PDA, allowing learners to see through to the ideas on other layers as they worked on the task, and (for assessment purposes) allowing assessors to separate each team mate's contribution to the developing idea. Unfortunately this proved impossible within the constraints of the budget, so instead we worked with a single sketch and requested learners to add sketches and notes in different colours.

Lots of text not much drawing

Predominantly learners used sketches with some notes to express their early ideas in box1 when working on paper and the PDA. In boxes 2 and 3 however, when developing someone

70
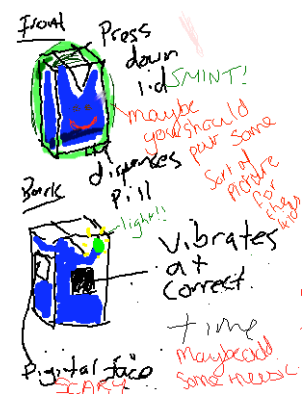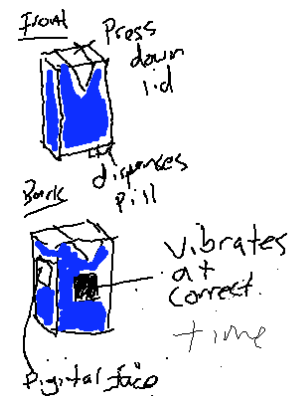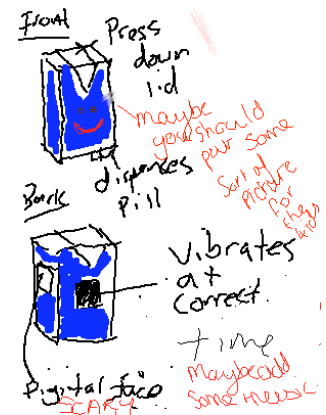
else's ideas, they mostly made notes on the PDA. This was not the case in the paper tests where there appeared to be more of a balance between sketching and notes across all 3 boxes. In fact on paper, learners tended to follow what had come before with a text-rich box 1 prompting a text-rich response in boxes 2 & 3. We think this difference can in part be explained by the lack of space on the PDA as learners filled up much more of box 1 on the PDA than they had on paper, often leaving only small gaps to add further ideas in box 2 & 3.
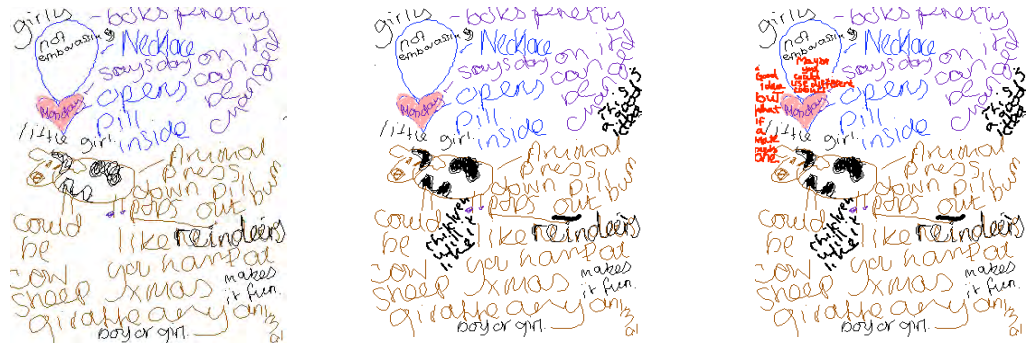
zooming in (tiny text)

To overcome the problem of lack of space, some learners used the zoom tool to write micro-notes for their teammates. It is interesting how some teams developed a common style, the three box 3's below are all from same group.

access to others' ideas

Passing early ideas around in this way, not only supports, validates, builds and sometimes challenges the individual pieces of work, but it also gives each of the team mates access to three potentially very different strands of thought which they could incorporate into their own work. The three examples of early ideas shown below were all from the same group and illustrate the rich diversity that can be achieved using this sharing technique.

high quality Box 1s for whole group (average performance for each overall)

typical quality box 1s for whole group Y10 (some high overall performance)

Y5 sketching performance

Although the majority of the learners who completed the PDA task were from year 9 and 10, we also trialled the system with one group of Y5 learners, and looking through their sketches it is often difficult to tell the difference between primary and secondary early ideas.

comparable quality box 1s for whole group Y5

Learner feedback on sketching with the PDA

In the learner questionnaire, a few learners, surprisingly mostly the boys, indicated that they did not like the PDA for sketching:

- thumbs down to sketching   (25 boys: 6 girls)
- thumbs up to sketching   (5 boys: 11 girls)

Overall twice as many learners agreed or strongly agreed that the PDA was good for sketching their early ideas (65%) as opposed to disagreed or strongly disagreed (35%)

The difference between sketching to think and sketching to communicate
Having spent the first 15 minutes sketching and annotating their early ideas on the PDA learners were asked to consolidate their thoughts on paper. After some particularly rich sketching on the PDA one learner responded to this instruction with "but I can't draw"!

In conversation with learners during and after the activity their response to the PDA drawing tools depended on how you asked the question. It was clear, and not surprising, that in the context of a formal test most learners felt they could create better presentation drawings on paper using familiar tools and techniques. Most had not thought about the difference between this form of formal drawing and quick personal sketches to help get their ideas sorted out, and if they had they did not consider that these would be valued in a test.

Once a distinction was made between drawing to present, and drawing to sort out your ideas, only a few learners still felt they would have preferred to draw on paper throughout. In a future development of the system, where learners have been working with the technology throughout their course of study, it would be possible to offer learners a choice of medium that in itself would provide interesting insights into their attitude to communication.

Many learners were frustrated by the size of the drawing area. We had purposely restricted the size of box 1 on the PDA to the same dimensions as the paper version of box1. This was partly for technical reasons (we were concerned that big files would clog up the wireless system) but also so we could compare digital and paper based responses.

The small screen on the device was also a problem for a number of learners, some reported it as too fiddly and some screens were misaligned and needed resetting. We are confident that with sufficient time and access to familiarise themselves with the tools and format of the devise and to personalise it to meet their particular needs, most learners would be satisfied that the PDA offered an appropriate or better platform for collecting early ideas and taking notes.

In it's present form the PDA is certainly not the best platform for creating presentation drawings, and even after significant time to get used to the tools, it would require a much larger input/drawing area and display screen. We are aware of a number of technologies currently in development, such as mini projectors, virtual keyboards and tabletops, which could be harnessed in the near future to create a more suitable interface for this type of work.

Taking photos using the PDA

The main difference between the photo-storylines in the paper and PDA tests was that learners had control of the PDA camera, whereas in the paper test, the teacher / administrator took all the photos. Handing over responsibility of the camera to the learners had a number of advantages, primarily that the learners could decide how to set up and photograph the progress of their prototypes.

*photo story-line from the paper test*

The PDA system also allowed us to collect more images (3 per hour) so learners could take shots from different angles to show different aspects of their work and provide a far richer pictorial record of the development of their ideas.
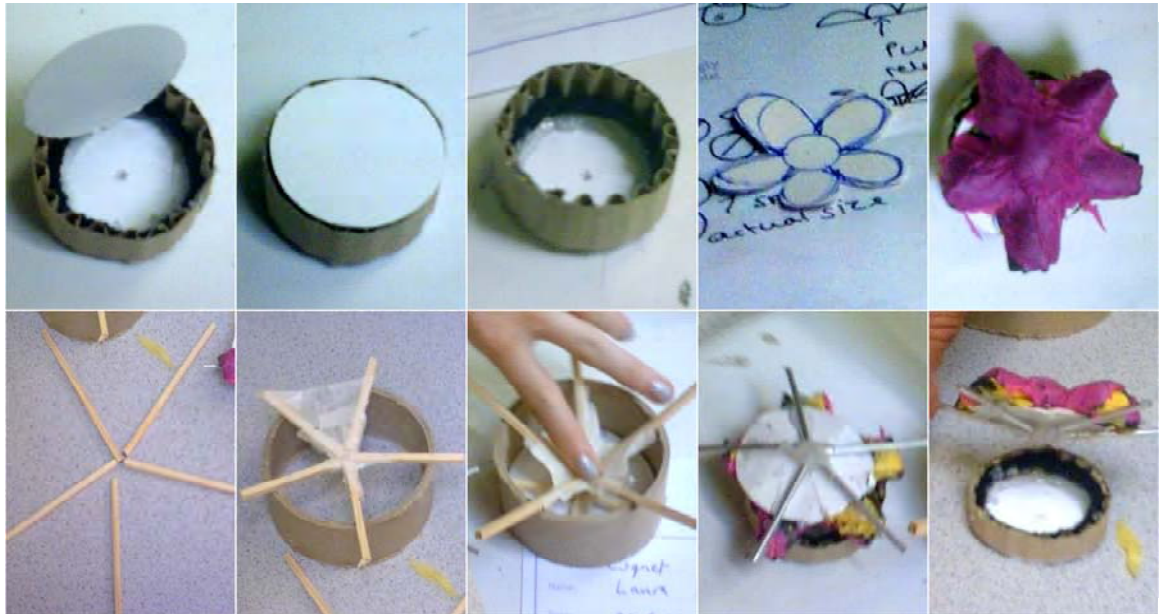

more comprehensive PDA photo story-line

The cameras in the PDA were not designed to take pictures close up in poor lighting conditions. While all the learners picked up the PDA drawing tools quickly, it was difficult to get all learners to take quality photos of the sketches in their workbooks. We provided black felt tip drawing pens to ensure high contrast in the drawings (pencil marks were grey on grey). With the right lighting and positioning it was possible to collect very clear pictures, however learners really needed more time to experiment and better feedback systems in the e-scape system to help them select the best photos to keep and transfer to their portfolio.
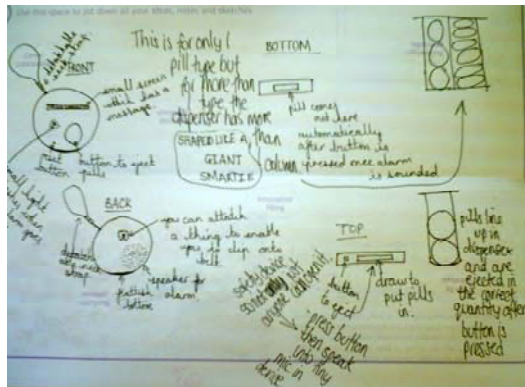
As with digital sketching it is important to reiterate that good photography skills do not necessarily mean you are a good designer, and even if you cannot use a camera well, it doesn't mean you are a bad designer.
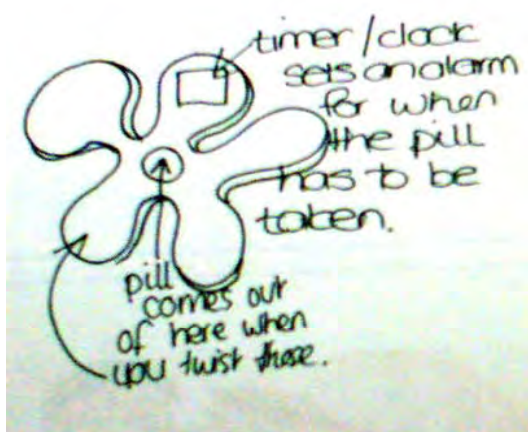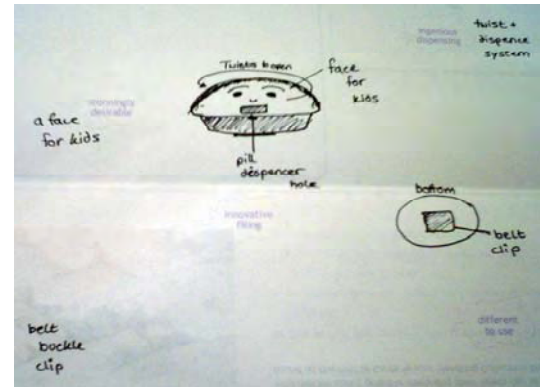

good photo, good sketching, typical D&T (Y10)
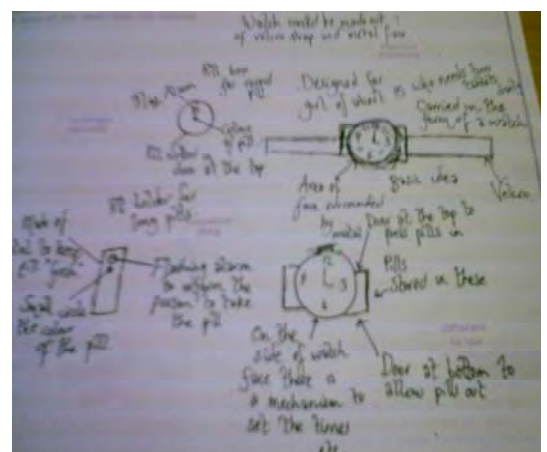
typical photo, typical sketching, good D&T (Y10)

74

good photography, good sketching, good D&T (Y6)

good photography, typical sketching, good D&T (Y10)

Using the PDA

As we have reported elsewhere, learners had no difficulty responding to questions using the various tools for entering text into the PDA. Their experience of handling text on electronic devises, particularly SMS messaging, helped to make this the media format that posed the least problems. The team had assumed all learners would use the screen keyboard to type in their answers, however many learners found the transcriber worked perfectly for them using the default factory settings. This feature automatically converted their handwriting from anywhere on the screen to typed text held in a pre-configured text frame. Other learners were confident enough users of mobile technology to modify the transcriber settings to suit their handwriting, the rest used the screen keyboard with surprising dexterity. Everyone seemed completely familiar with the predictive text system. Almost exclusively, the adults we tried the system, struggled to type in a couple of words into box 5 in the allotted time; hunting and pecking on the screen keyboard. Most learners managed to input at least a sentence, and some managed several.

Box 5 user profiles

At this point in the task it is important to get the learners to think about who they are designing for. This is the first time the learners use the PDA text entry system in the task. As with the paper version of the activity, some notes in the previous sketch boxes may already show an implicit concern for a specific user group, or explicitly target a group before we request the information in box 5.

user group implied independently in box1

general user group suggested independently in box1

specific user group specified independently in box1

We provided a set of user profiles and learners either chose one of these, or extended, or amalgamated, or created their own user profile for their particular design. The following screen shots illustrate the wide range of responses to this sub task:
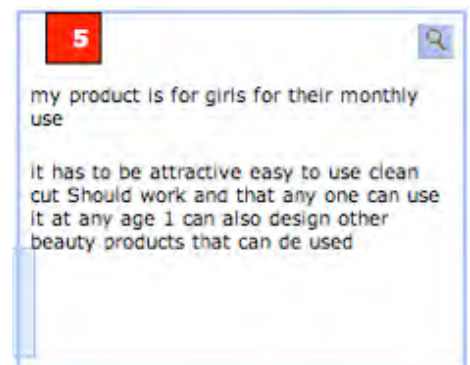
5 children aged 5-8 durable colourful easy to use affective	**5** anyone! my design is small and light! so it can be used outdoors, it has changeable top so kids and teens can have fasionn! 'cool', good for outdoors! and effective! and cost effective (cheep to make)	**5** anyone with a front pocket on there shirt stlyish/mordern that looks gd so it is bought and sold. it must also work properly and be suited for different sized pockets	**5** suart age 10000 likeS to eat fish! the fish is blasted out of a canon ! and stoast eat s
typical box 5	*good box 5*	*qualified box 5*	*whacky box 5*

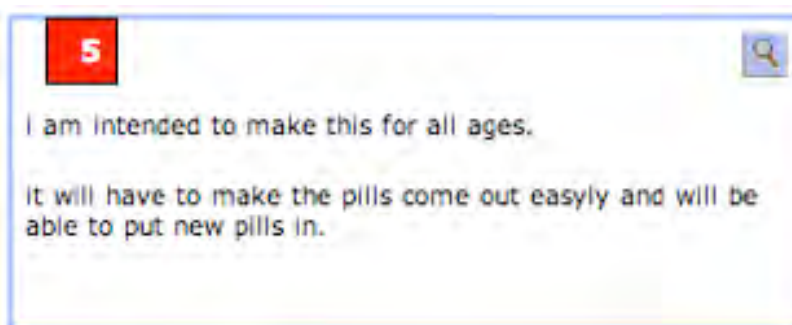|	
5 My intended user would be about 15 years old, a boy or a girl. The colour scheme would alter, depending on the gender of the user. To be successful , my product would have to work properly and provide a secure place for the pill to go. The timers and alarms would have to be reliable and the customer should be able to put their full trust into the watch. It shou	**5** my product is for girls for their monthly use it has to be attractive easy to use clean cut Should work and that any one can use it at any age 1 can also design other beauty products that can de used
comprehensive box 5 *(they only had 5mins for this)*	*empathy for a specific user group*

As with all the other aspects of the PDA task it is important to reinforce that good performance in individual sections is not necessarily related to good D&T performance. It's not the snapshot of where they are that counts but how this changes throughout the activity.

5 i am intended to make this for all ages.

it will have to make the pills come out easyly and will be able to put new pills in.

This user profile is from one of the top ranking portfolios. As this learner works through the task they develop a much clearer picture of their user group which changes from this general statement in box 5, to a much more specific view in box 17 ***"My project is a game for little kids to encourage them to take their pills…"***

This ability to grow and develop throughout the activity is an obvious sign of capability.

A clear vision of a user group for the product being designed is an important reference point for evaluation and without it, it is difficult to steer a project appropriately. Learners' familiarity with digital text entry seems to have helped many to get more information down during this aspect of the activity.

5 my intended clients and users are young children or people w9th small pills that may get lost in their pocket or handbag.

Review Comment:
i think these are good notes as i kept looking back at them when we were making the product and they helped me through the making.

Fast-back review comments

In the paper-based designing in *assessing design innovation*, identifying when learners had made specific additions to their notes was often difficult. It was usually possible to unpick how their sketches and models had progressed, but working out when they had added notes was difficult. By contrast, all the contributions through the PDA are time stamped, making it really easy to tease out a chronological view of each aspect of the work. This ability to place work on a timeline led to the development of a text-based review feature that we called **fast-back**. At the end of the activity we allowed learners to review each stage of the activity and comment on anything they would have done differently if they had known then what they know now.

Some of these comments referred to problems with the equipment and system, mostly early on in the activity:

Review Comment:
i think i designed the product pretty well in the pda but it was quite hard to draw on it as the line or whatever i was trying to draw wouldn't go in the exact place that i wanted it to and the rubber won't rub out tiny little mistakes.

Review Comment:
the design was not good to model, to create a good working system in that small a space was an unrealistic expectation.

Some referred to their ideas and models

Others were more meta-cognitive and reflected on the process often in quite honest and insightful ways.

Comments on Grant Naylor's work

colourful and eaay to access will appear use ful to the elder comunity

it needs to have an easyer way to get the in and out.

Review Comments:
i find that if i have to judge somones work i offen find loads of things rong but find much more that makes it better

Comments from Lucy Shaddock

Review Comment:
my friends helped a lot

Even though this was the last sub task and most learners were exhausted by this stage of the activity, their texting skills ensured that these text frames were a rich source of evidence, not only of their design ideas but also of their attitude and approach.

'Design-talk' using the PDA

When we had completed the paper project we realised that we had missed a significant opportunity by not asking learners to annotate their photos. The photo story-line provided invaluable evidence of **what** learners had done while modelling, but there was still often very little evidence of **why** they had done it. Consequently as we were developing the e-scape system we were determined to correct this weakness and the voice-memo tool was the ideal medium. After each photo session in the **e-scape** system we asked learners to record a 30

77

second audio note (a sound-bite) explaining what was going well with their design and what could still be improved. We were keen to explore the role of talking in the design process, an issue that arose in the previous project (*Assessing Design Innovation*) when a teacher had pointed out a particularly animated conversation between two learners. "If only we could eavesdrop on that conversation we would know so much more about what they are trying to do…" The e-scape system allowed us to explore the potential of design-talk. We outlined above (in part one of this report) our early trials of some possible approaches, and eventually we settled on the sound-bite approach, illuminating the photo-story-line.

Photo thumbnails and sound file buttons from an e-scape web portfolio. Clicking the sound file buttons plays the audio clip using the computer sound system.

We were careful to restrict the length of each audio clip to 30 seconds to avoid data overload for the judges and to force learners to summarise the key points they wanted to make. Typically early sound files are descriptive of the modelling process:

"What is working well is that I have got the shape how I want it, I've glued it and put a pipe cleaner inside the corrugated cardboard to make it secure"

Most learners identified strengths and weaknesses in the progress of their work. Their comments still appeared to be more positively than negatively skewed, but not to the same extent that we had seen in the paper versions.

"What isn't going so well is the top bit of my model because I find it hard to cut out the shape that I want so I need to keep trying with that – but the rest of it's OK"

Some learners demonstrated developmental growth across the series of audio clips, identifying a weakness in an earlier clip and reporting it resolved in the next.

"the thing that will need further development is the basket where you pop up the pills and it has to get caught in the basket because at the moment the basket is in the way of the pills popping up"

" I have found out how to make the pills pop up by making a piece of card with a hole in it and putting some springy stuff over it"

Although some learners, particularly the girls, reported feeling uncomfortable recording the audio notes, it was surprising how, after the first few attempts they relaxed and allowed their authentic voice to come through and some assessors commented that they felt they could identify additional aspects of learners' attitudes as well as intentions, which was not evident

in their drawings or notes. It is difficult to summarise this through transcripts of learners' recordings. We have included a selection of audio file samples in appendix 1.

Lift Pitch

The last audio task we set the learners was to 'pitch' their ideas to the company directors during a short (30 second) lift ride from the first to fifth floor. Given how little training, guidance and previous experience the learners had with this technique the team were amazed how well they coped with it, both in terms of the *quality* of the pitches and the *diversity* of the styles they used.

Some learners described **what** they have done
- how their thing is made (materials, construction techniques etc.)
- what it looks like
- how it works
- how it's used

Some made a **sales pitch** to the user reflecting some understanding of why someone might want their thing
- bright and pink for girly types
- chunky, clear and easy to use for old people

Some made a **design pitch** to the manufacturer reflecting more analytically about why they have done things
- why it's made the way it is
- why it looks like that
- why it's good the way it works
- why it will be good for users
- commercial/economic advantages

There are also different ways in which learners present their pitch including:
- natural/relaxed conversational style
- humorous/whacky/off-the-wall
- dynamic/gripping/intense/dramatic documentary style
- formal/structured and business like
- concise and to the point
We even had one girl who rapped her presentation

And the pitches provided plenty of evidence of:
- the level of honesty/grip
- the speed of delivery
- the level of confidence
- the range of issues vs. the depth they go into
(too much is almost worse than too little - balance in all things)

The following is typical of the type of response to this subtask. As with the strengths and weaknesses, the transcript is not representative of the subtle dimensions the live audio conveys and we have included a selection of lift pitch recordings in appendix 1.

"My pill dispenser is unique and individual and it's got a hand wrist strap that can be worn around the wrist or attached to things like bags. It's got 2 main compartments, which inside have 2 secondary compartments so it can hold up to four different types of pill without confusion. The lids are tight so the pills don't fall out and the boxes can be separated and taken if the person only needs 2 types. It also comes in different colours so that people with different tastes will like them"

Design-talk is a significantly undervalued aspect of design capability, whether it is a focus group interview with clients (to work out what they want), or an intimate team conversations (to work ideas out), or a more formal presentations (to pitch your ideas), communicating what you mean is – for many learners - much more achievable through the spoken word than through drawing or writing.

Tracing the evolution of ideas

A significant advantage of the *e-scape* portfolio system is that it allows copies of learners' responses to the various sub tasks to be saved in each learner's portfolio. If we wanted to trace the provenance, genesis or exchange of ideas across a group in the paper tests it was a significant task to collect all the scripts together and cross-reference the individual items. With the web portfolio, everything a learner has contributed to, and all their teammates comments, are recorded into each portfolio making cross-referencing much more straightforward.

Learner 'A' has jotted down this early idea in box 1. It is clearly derived from the popper-pencil we included in the handling collection, but it is definitely not a snake yet!


This is the box 1 learner 'A' received from one of their teammates. Learner 'B' has drawn a caterpillar and Learner 'A' has suggested using the antenna to dispense pills.

These are the box 4 consolidation sketches from learner 'A' who has modified the early idea for a technical system and synthesised ideas about how it might look and work for the user, inspired by learner 'B's caterpillar sketch.

And here is the box 4 sketch and notes from learner 'B' who has modified the early idea for what the pill dispenser would look like to incorporate ideas about how the system would work, inspired by the antenna dispensing mechanism suggested by learner 'A'.
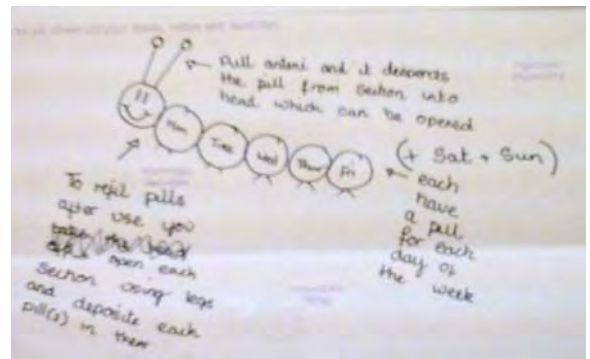

This is not cheating. It is evidence of sophisticated design & technology capability in action as learners merge and synthesise ideas, an important quality which we struggled to evidence in the paper activities. With the digital portfolio, is has proved to be far easier to follow the illusive path of ideas as they passed from one teammate to another.

Growth - the key to capability

In this section we have considered aspects of performance across each of the four media types collected using the e-scape PDA system. Isolating aspects of performance in this way is useful for comparative analysis of assessment systems, but can be confusing in terms of establishing overall capability.

As we have repeated several times, exceptional performance in individual sub-tasks is not necessarily an indication of design ability.
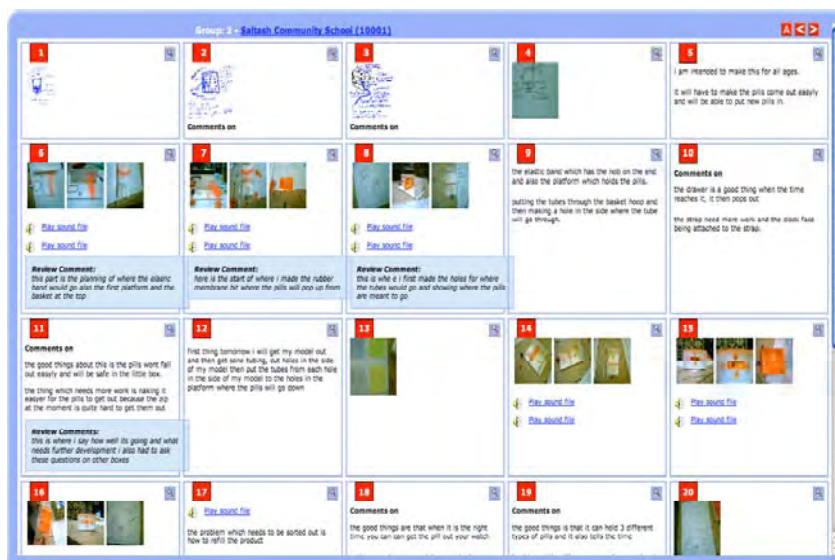
If there were a single indicator of such capability then we believe it would be evident in the *growth* of ideas through the project. Growing ideas is the heartland of design and technology and this is not evident until you see the whole portfolio together and make an overall holistic judgment of how the individual components interact towards a purposeful solution.

**Review Comment:**
I wish the PDA had worked!
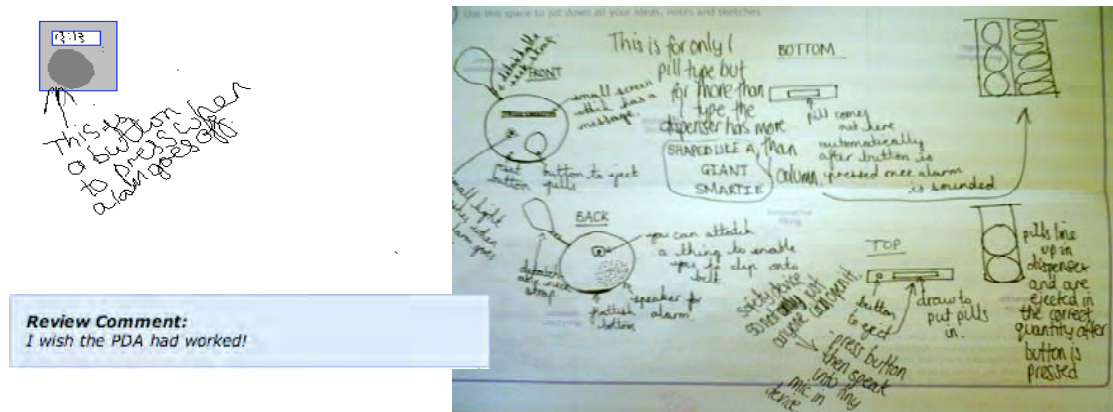
Despite a somewhat shaky start, the detail above from this Y5 portfolio illustrates significant growth in this learners' thinking between box 1 and box 4, a clear indicator of capability.
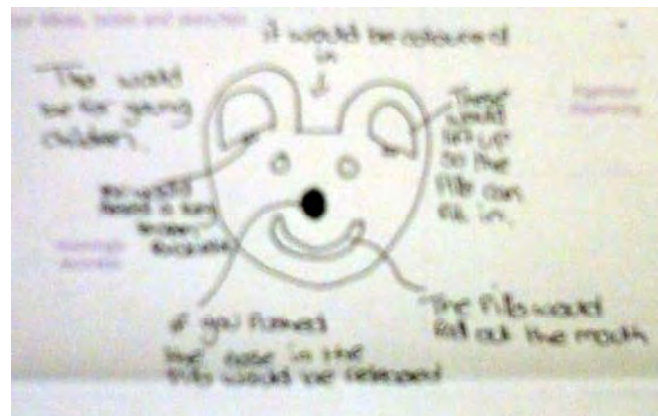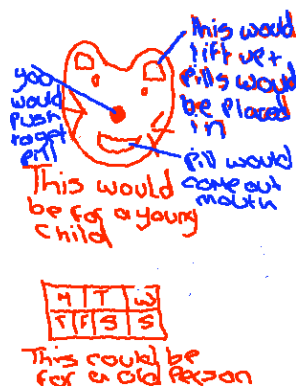
Compare this to the limited growth between boxes 1 and 4 for this Y10 learner. (Both learners' work then develops rapidly and the portfolios rank highly overall)

2.17 findings in the data
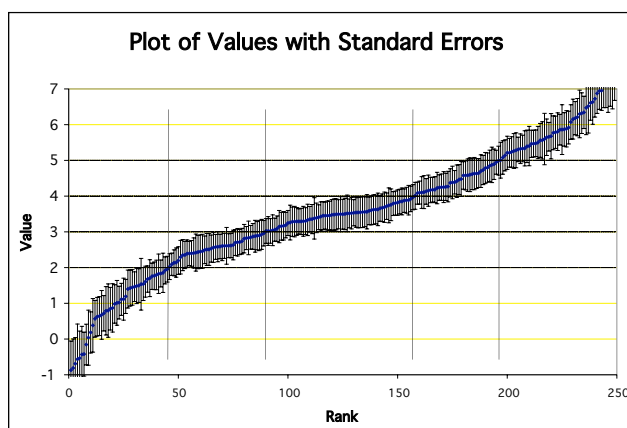
The distribution of performance across the whole of the e-scape sample is shown in this chart. It also shows the Standard Error attached to each portfolio placement.

In his analysis of this distribution Pollitt points out that:

if we exclude the 30 best and 30 worst portfolios the average standard error is 0.41 grades… In formal statistical terms 68% of the portfolios' "true" values will lie within one standard error of the reported value.

Vertical lines are drawn through the grade boundaries to show how many learners would fall into each grade. Note that Grades 2-4 are, by definition, equal size, and that this leads to more learners lying in the central grade than the ones either side of it. (see Pollitt section 2.13)
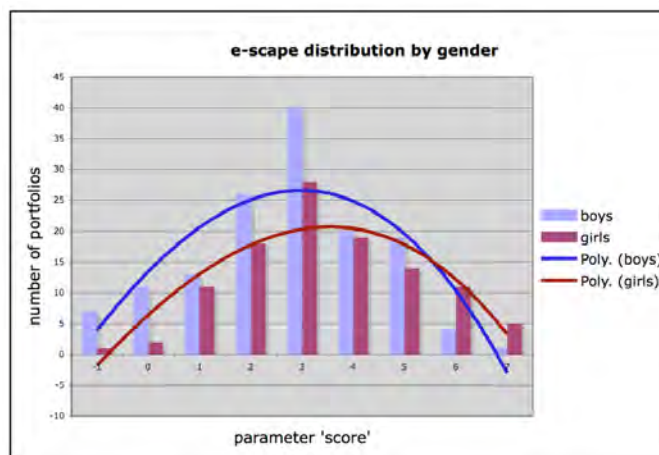
Performance by gender

The e-scape sample was not a gender-balanced sample. Despite our attempts to achieve this balance – by asking teachers to create such groupings – we were often worked with the pre-existing GCSE groups that 'belonged' to the teacher responsible for managing the e-scape pilot in that school. These groups were not always balanced and we ended with a sample of 109 girls and 140 boys.

Their respective performance is shown on the chart here, with the girls mean performance being 3.95 and the boys 3.21.

The performance difference is evident throughout the scale.
At the lower end of the scale there are only 3 girls having a parameter score of less than 1, whilst there are 18 boys in this group. At the top end of the scale there are 16 girls with a score above 6, whilst there are only 5 boys in this group.

The gender differences in performance are clearly evident in this chart. It shows a simple count of the number of portfolios (girls and boys) within each parameter group. The girls are clearly over-represented (in relation to overall gender group sizes) in the groups with parameter scores of 6 & 7, and under-represented in groups with parameter scores of -1 & 0. The polynomial trend-lines illustrate the substantially different peaks that reflect the mean scores reported above; boys just above 3 and girls just below 4.
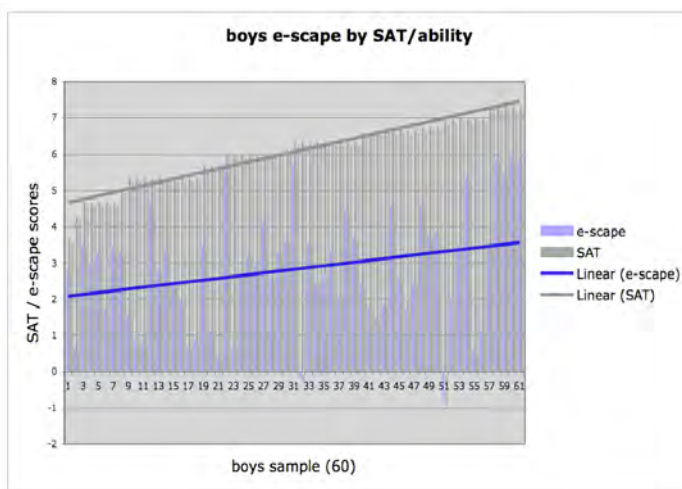
Performance by 'general ability'

Whilst we were not entirely successful in getting balanced gender samples from schools, we were rather more successful in getting balanced ability groupings. We asked schools to provide us with KS3 SAT scores for English, mathematics and science, and six of the twelve schools were able to provide these data for us. From the three SAT scores we created a mean score for each learner and used this as a surrogate general ability measure. The combined mean SAT score across the whole sample (for both girls and boys) was 6.1. In so far as SAT scores can show, the gender groups are of equivalent general ability.

girls e-scape by SAT/ability
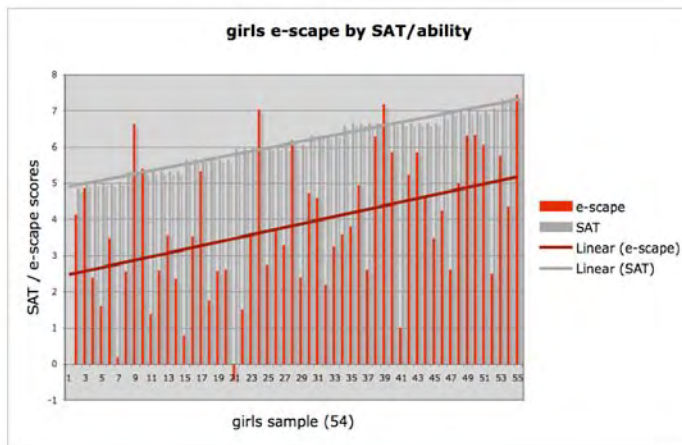
But these scores show interestingly different relationships with their e-scape performance parameters. The two values can be shown together, though it is important to recognise that the SAT scale runs from 4-8 while the e-scape scale runs from -1 to 8

The e-scape performance of the girls group shows some wildly varying scores when related to general ability. Indeed the chart appears to demonstrate a high level of instability. But the trend-line demonstrates that there is a relationship between e-scape performance and the calculated general ability score. The correlation coefficient between the two data sets is not overwhelming, but it is positive (0.42). It does appear that whilst for individuals the data is very unpredictable, for the sample as a whole performance does generally rise in line with the calculated general ability measure.


boys e-scape by SAT/ability

The position with the boys group is interestingly different. Once again in relation to individuals, performance seems very variable in relation to their calculated general ability score. However the trend-line suggests that there is some relationship with general ability and the

correlation in this case is 0.28. It is also evident that the trend-line is at a considerably lower level (2 > 3.5) than that for the girls group (2.5 > 5). It is interesting to speculate on why girls' performance relates more closely to general ability than boys' performance.

Performance by school
It gradually became apparent to us as we worked through the judging process, that the schools were not all equivalent in their learners' performance.

e-scape performance by school & gender
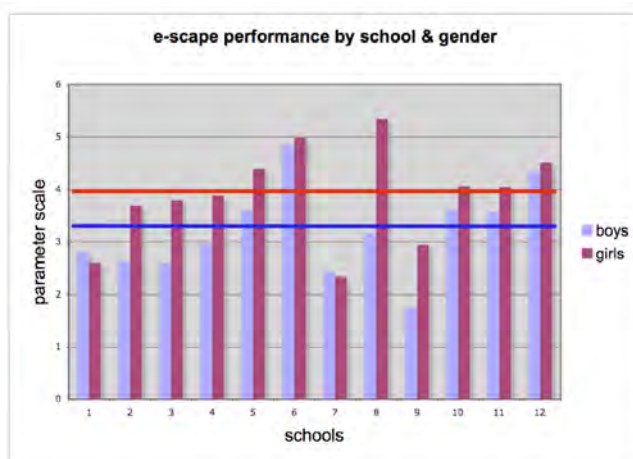
The chart here shows performance by school and by gender; the schools arranged simply in alphabetical order. The mean scores for girls and boys are also shown.

It is clear that there is considerable variability not just in the overall performance levels but also in relation to the gender differences. Schools 1,6,7 and 12 have very small gender differences, whilst schools 2,3,9 and (especially) 8 have very considerable gender differences.

In the APU survey that we conducted for DES in 1991, we demonstrated a far higher 'school effect' than was normal for most subjects and we attributed that to the relative newness of the subject. We speculated that where design and technology was well established in the culture of the school, performance was likely to be better than in those schools where it was new. But this new data is not so easily explained, since in all the schools visited there was evidently well-established practice.

We suspect that the differences evident here are attributable to a combination of factors.
• greater difficulty (in some schools) with the technology
• less familiarity (in some schools) with innovation-centred designing
• less familiarity (in some schools) with hand-held digital tools
• less flexibility (in some schools) with dealing with the e-scape challenge

e-scape performance related to 'light-fantastic' performance
We have explained earlier how the logistic challenges of running the e-scape activity and the light fantastic activity resulted in having to compromise on the idea of the whole sample taking both tests. In the end we achieved two tests for two of the schools in the 1st and 3rd


distribution of 'light fantastic' scores
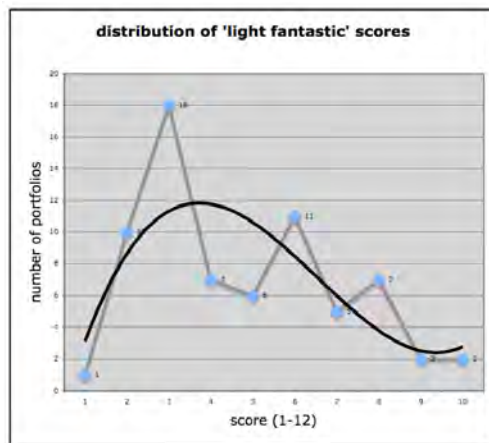
e-scape rounds. These four schools (69 learners) provide only a relatively thin data set compared to the 12 schools (249 learners) in e-scape, but in the circumstances it was all that could be achieved.

The activity originated in the 2004 project *Assessing Design Innovation* and the work this time was marked using the same (1-12) rubric as was used there. The distribution of marks in this 2006 pilot sample is as shown here, and the trendline suggests that the 'centre of gravity' of the work lies around 4 on that scale.

The distribution appears to be very similar to that which we noted in the assessments of the work from 2004. As we reported at that time ...

At the limits, 25% of the sample scored within our lowest band of marks (1,2,3) and 8% of the sample achieved our top band of marks (10, 11,12). This distribution is reflected in the polynomial trend-line. This trend-line shows the general tendency of the performance levels to be 'bottom-heavy' and for progressively smaller percentages of the sample to be able to achieve the upper levels. This we believe is a reflection of the current general performance of learners in design & technology… Design innovation has not received the attention that it deserves and this is one of the reasons why this project was established.
*(Kimbell et al 2004 p 42)*

total distribution

85
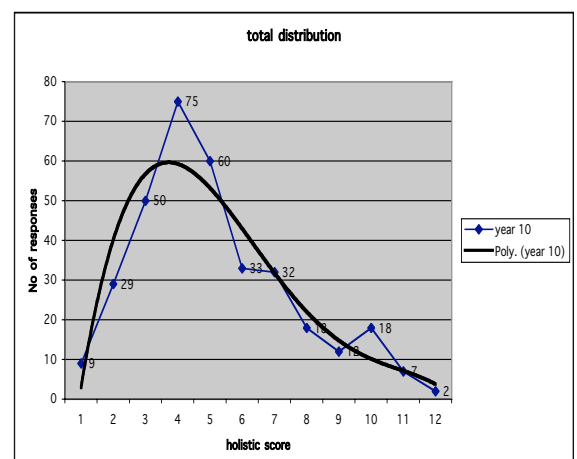
Despite the thin sample therefore, there is reason to believe that the 69 learners performed in line with the expectations for that activity.

distribution of e-scape 'scores'
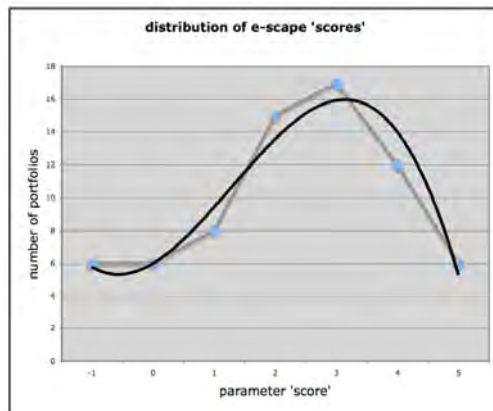
A slightly different picture emerges when we look at the e-scape distribution for the same group of 69 learners. The sample appears to be bunching at the upper end. Whilst 'light fantastic' has a long neck in its distribution curve, e-scape has a long tail.

Nonetheless, the performance of this group is not strong in relation to the e-scape sample as a whole, the mean score being just below 3. When viewed against the whole e-scape sample, it is clear that this 2-test group are underperforming the sample as a whole, and perhaps more markedly so in the case of the girls.

e-scape performance by gender
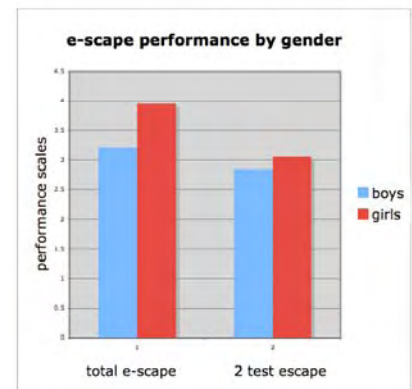
One of the questions that this sample was intended to help us to answer concerns the relationship between performance on screen and on paper. Did good 'on-screen' designers also perform well on paper, and vice versa?

There is only a tiny sample of 22 girls and 47 boys, and there is a very limited correlation between the two sets if figures. For the girls, the correlation is positive but small at 0.2. In the case of boys there is no correlation at all. When we look closer into the data to see why this might be the case, it becomes evident that the inter-quartile statistic suggests that there is indeed a relationship between the datasets, but that any possible correlation is being destroyed by a small number of extremely interesting cases. In these cases, learners have a *very high score* for e-scape and a *very low score* for light fantastic. As examples, the highest scoring girl in escape (5.77), achieved only a 3/12 for Light Fantastic, and the highest scoring boy in e-scape (5.7) achieved only a 2/12 for Light Fantastic.

How are we to explain these very large discrepancies? There are very few girls in this category, and just removing from the list the one extreme-case girl identified above raises the positive correlation to 0.4. There are rather more boys, but by removing just 4 from the sample of 47 the correlation rises to above 0.3. Interestingly, one of these 4 was the profoundly deaf young man mentioned earlier. His work with the e-scape task was the fifth best of the whole boys sample (scoring 4.8), but with Light Fantastic he scored only 2/12 and was in the bottom 15% of the boys' sample

Having identified these 5 individuals, we contacted the schools to ask specifically about the performance of these few. The following comments were reported to us by the teachers we

contacted. At our prompting, they had asked the learners about their reaction to the two tests.

"I didn't take the light fantastic seriously…..I didn't plan my idea carefully enough….thought my design wouldn't work…...I wasted a lot of time. We didn't have enough time to make anything!"

"I enjoyed using the PDA and found it very easy to use…..my idea was better because I had a chance to improve my paper version."

"I loved using the PDA to create my design...really good way to design ideas……I enjoyed seeing my friends look at my design idea."

It is interesting that whilst there are these few cases of really good e-scape performance associated with really poor Light Fantastic performance, there are none in the reverse category. None of the very high scoring Light Fantastic learners performed very poorly in e-scape.

It does seem as though for at least a small group within the sample a *motivational* element goes some way to explaining the misfit between the two data sets. Learners almost always found the digital form of the activity unusual and engaging, and perhaps they performed better than they might have been expected to, and certainly much better than they did in Light Fantastic.

2.18   issues arising

concerning the classroom activity

Many features of the classroom/workshop activity in e-scape were developed initially for the previous project *Assessing Design Innovation*. The structure of the task, the handling collection, the learner booklet and teacher script, the photo-story-line, the review and post-it sessions, the team-based generation of and reflection on ideas, the modelling resources and time allocations … all worked as well as they have done before. They are part of a well-established TERU armoury for controlled coursework projects in school.

There were some new features however.
• The client-cards worked well – offering learners various target groups for their pill organisers. Many learners simply chose one or another of the clients, whilst some used the cards to create an original client for them-selves. The cards did seem to promote the importance of having a clear view of the client.

But the major transformation from previous projects was of course the digital tools and the web-based portfolio, and these innovations have created some dramatic possibilities for the future

• design-talk – using voice files via the PDA – has enabled us (for the very first time in assessment history) to collect the authentic voice of the learner – on task – and present it in

the web-portfolio in a time-stamped slot. Furthermore, we are able to seam together the sequence of voice files into a continuous file of approximately 2 minutes, providing a continuous account (from the learner) of the evolving design product over the 6 hours of the activity. This account highlights the strengths and weaknesses of their emerging solutions and, taken as a whole, provides a really good indication of the capabilities of the learner. In the current version of e-scape, this design-talk is steered by a series of question-prompts and it seems self-evident that the nature of learners' response will be driven by the questions we pose. We need to do further work on this matter to identify an optimum question set.

• the drawing and text tools were digital replacements for existing paper drawing and writing, but – in the hands of learners – became more creative tools than we had dared to hope. Learners' familiarity with texting enabled them to communicate far more than their teachers could with the same tool, and section 2.15 demonstrates the imaginative design response of learners in their use of the drawing tool. In the current version of e-scape, the budget did not allow for the creation of a bespoke drawing package and we rather adopted an existing tool. In the future we would ideally evolve a customised tool-set that would be simpler and easier to access for the very brief time available in the activity (10 minutes here and 5 minutes there) for this early-stage concept drawing.

• the interplay between digital and non-digital (both paper and materials) worked apparently seamlessly and encouraged us to the conclusion that it is not a matter of creating entirely digital or entirely paper-based activities. We can mix and match to maximise the best combination, The advantage of the digital system is that we can seamlessly collect an evidence trace of how this interplay is working, in future systems we can imagine a feedback system that helps learners to monitor this as well as their teachers, and make judgments about how effective their choice of tools/designing processes are as they work through a task.

• we showed in *Assessing Design Innovation* that a single task can be replicated into (in that case) 9 different tasks covering the whole spectrum of design & technology (textiles, systems & control, graphics etc) with an identical activity structure. All that changes is the task. We can reasonably assume therefore that we could do exactly the same with the e-scape version; enabling us to create a matrix of tasks that – taken together – cover the whole spectrum of design & technology.

• moreover, we see no reason why these activities should be restricted to design & technology. Any activity-based task could be structured using the same toolset – be it an English composition, a science investigation, a drama improvisation or a geography problem-solving activity. The key thing is that it is an **_activity_** to be pursued in a way that demands some kind of *performance* within which learners can demonstrate capability. Within those parameters it would be possible to design an assessment activity using the TERU / PDA toolset so that learners create a web-based performance portfolio that can be assessed remotely.

• this issue is particularly significant in the context of the current debates on the status of GCSE coursework. An e-scape model of coursework would be ***structured supervised coursework*** managed with the TERU / e-scape toolset and undertaken in controlled conditions in schools. We believe that there is a great deal of potential in this approach and have opened negotiations with DfES, QCA, Becta and GCSE Awarding Bodies to create a nationally scalable version of e-scape for this purpose. See 2.18 below: 'next steps'.

Within this 'next step' we also propose to explore the possibilities of teachers' monitoring of coursework beyond the classroom. Time-stamped data in the portfolio show ***when*** the work was done; GPS data can show ***where*** the work was done, voice memos show (at least to some degree) ***who*** did it. These technological tools are available and offer a degree of confidence about the authorship of a piece of work, even when conducted remotely. There is much potential to be explored here.

• the e-scape national pilot (June / July 2006) demonstrated that learners adopt the system VERY rapidly. It seems to be a natural extension of their mobile technology/gaming/media-rich youth culture. We should note however that teachers need more support and more time to get to grips not just with the technology, but also with the transformations that this creates for their interactions with learners. There is a tendency for teachers to feel de-skilled by (most) youngsters' digital dexterity and to believe that because of this they have nothing to offer in support of this area. Teachers need help to rebuild confidence in their skills as educational/learning experts, and see that combining their learning expertise with young peoples digital capability offers great potential for developing rich and compelling learning experiences. This too has messages for future work – see again 2.18 below.

• finally in this section it is worth pointing out that we saw very little discernible difference (in terms of attitude and response to the e-scape task) with the year 5, KS2 learners. There are some differences in the quality of the outcomes, and these are attributable principally to less developed modelling skills and – to some extent – a somewhat more naïve approach to design. But when we consider these relatively small performance differences between the year 5 group and the mass of the responses in year 10, we are forced to wonder what has been happening in the 5 years that separate the two groups.

concerning assessment

i) designing the web-portfolio

Since the design of the web-based portfolio had to be undertaken before we had any portfolios to put into it, the decision we made were based on a series of best guesses about how we might tackle the assessment process. Our approach was to order the sub-task boxes sequentially on the screen so that the whole of a learners' work was laid out effectively in a time line giving us an instant overview ('glanceware" in software jargon), and provide a zooming feature to allow assessors to zoom in and out of the detail of each of the sections (See section 2.11)

Several difficulties subsequently arose with this layout that made the use of the 'comparative pairs' assessment more difficult than it might have been. As an example, the layout of box 1-2-3-4 locates the work of the individual learner:

1	their 1st ideas
2	their response to team-mate (a) 1st ideas
3	their response to team-mate (b) 1st ideas
4	a photo of their consolidated drawing in the booklet

We could however have sequenced the web-site a different way:

1	their 1st ideas
2	their 1st team-mates response to those ideas
3	their 2nd team-mates response to those ideas
4	a photo of their consolidated drawing in the booklet

This 2nd arrangement has the weakness that the flow of work from the principal learner is interspersed with the supportive interventions from the learners' teammates, but the benefit is that the consolidated drawing in box 4 makes more sense – since it is a direct result of the work in box 1/2/3.

Several design issues of this kind arose through the judging process and inform how we might re-design the web-based portfolio to make it more effective.

Moreover, the structural hierarchy of the site starts with a list of schools (click on one) that reveals the list of learners (click on one) that reveals his/her portfolio. Having worked for many days through three phases of judging (see section 2.12/13/14) a number of other priorities became apparent in terms of accessing the appropriate portfolios. Three priorities in particular would feature in any subsequent re-design of the web-site.

i) to anonymise the source of the portfolio by using a unique candidate number system that is independent of the school (which could also have a unique centre number).

ii) to display the work not merely sequentially – but in ways that allow judges to customise its layout to suit their preferences (e.g. the relationship between voice files and photos).

iii) to enable judges to 'home-in' on critical areas of evidence more quickly for the purpose of facilitating their judgments

In particular it would be desirable in a future system to provide a set of tools to allow each judge to arrange/view/setup the portfolio in their preferred way, for example by changing the order of components, or the size of each item on screen, or back-grounding some components while fore-grounding others or getting some to play automatically in the background. It was not possible to image and model these possibilities at the outset since we had not at that time been introduced to the judgmental pairs system.

ii) extending the comparative pairs approach

The comparative pairs judging process that we adopted with the support of Alistair Pollitt (see section 2.12) was premised on the holistic approach to assessment that we have always advocated for design & technology (see e.g. Kimbell et al 1991, Kimbell et al 2004). The case for holistic assessment in design & technology lies in a combination of validity, reliability and manageability, and the comparative pairs approach extends the argument significantly, particularly in terms of reliability and manageability.

The reliability of the judging process is significantly enhanced over conventional approaches to assessment, partly because of the multiple comparisons that are made (each portfolio is seen against many other portfolios) and the multiple judges that do the comparisons. Furthermore the system automatically flags up both the portfolios that are causing any difficulty and the consensuality of each of the judges. In either case, anomalies are identified and can be dealt with. We have described this in section 2.12, but it is interesting here to consider some further possibilities of the system.

The potential exists to expand the assessment and learning impact on the people who are involved in the assessing / judging process. Currently (for GCSE) teachers assess their own learners' portfolios – but do so with very limited external reference. They do not see work from other schools and the teachers from other schools do not see their learners' work. This insularity has two downsides. First, teachers' assessments are unable to recognise the wider local, regional or national picture of capability, and second (even more important) the teacher does not get a sense of the variety and strength of other work that is being undertaken. The professional development potential of the current arrangement is therefore limited. It has long been recognised that the real benefit of becoming an external examiner or moderator is that one gets to see so much other work from so many places and at many levels of capability.

Imagine then a situation in which all teachers are e-scape-style judges. They get to see all kinds of work in the website and can compare their own learners' portfolios with those from many other schools. The effect would be to make **all** teachers into external assessors / judges. Quite apart from the assessment benefit of this, the professional development benefit could be substantial.

Going further however, imagine if learners themselves were able to access the website – albeit probably through a different set of gateways. Teachers often try to keep hold of copies of previous work to provide models of performance for their current learners to see and gain inspiration from. Imagine a situation in which learners were not only able to see this other work (in the website) – but were encouraged also to engage in the judging process. Comparing one's own work with that of many others could – if managed properly – provide a very valuable learning tool.

Further refinements are also imaginable.

Current arrangements continually suffer from the challenge of assessing trends over time. Is this year's work better than last year's – or worse? Are standards going up or down? With an e-scape style system it would be possible to integrate previous portfolios (from last year) into the sample for this year – and see where they end up in the rank order. This would provide an immediate measure of the stability (or otherwise) of the scale.

Moreover these insertions from previous years might be benchmarked to indicate grade boundaries and again the system would provide an automatic register of how these boundaries appear in the current scale. Particularly in the context of coursework assessment, one can imagine a new order in which many benefits could flow from the wider application of e-scape style portfolios assessed through a comparative pairs judging process.

Finally however it is important to note that there would also be a knock-on effect in the classroom – concerning in particular the relationship between formative and summative assessments.

Currently teachers often use project-work assessment rubrics as teaching tools, pointing out for learners the kinds of things that they will need to do to get all the marks for this section or that section of the rubric. In an e-scape world however, - whilst there would still be criteria to guide the judging process – they would not work in the same way. Marks would not be allocated against individual criteria and then added up. Rather the judge makes an overall balancing decision about the strength of **this** piece as against the strength of **that** one. The summative assessment process therefore becomes far simpler and quicker than is currently the case. Our judges in this project (particularly the teacher who is currently involved in GCSE assessments) were clear about the relative speed and simplicity of the approach. See section 2.14.

But there is a consequence for formative classroom assessment, for the rubric is no longer something that can be used as a scoring guide to show learners how they might improve their work. In it's place however one can imagine all sorts of alternative and more holistic support systems for learners. Since the judging process is based on a holistic judgment – learners would need to understand what it is about their current work that makes judges say it is better (or worse) than another piece of work. The focus of formative assessment would be on how the **overall quality and impact** of the work might be enhanced.

Furthermore, it also seems perfectly possible to *adapt* the pairs judging process so that it is more than just holistic. We can imagine a rubric in which the holistic judgment is followed by 3 or 4 major category sub-judgments – again using direct comparison with other portfolios.

	Portfolio A	Portfolio B
holistically	wins	loses
category (a)	wins	loses
category (b)	loses	wins
category (c)	wins	loses

The comparative pairs system has never previously been used in this way, but in the context of exploring the relationships between summative and formative assessment this would seem to us to be fertile and important territory.

concerning the technology

i) the PDA

The e-scape activity was dependent upon the PDA. We chose this as the principal classroom tool because of its multi-functionality – to enable drawing, writing, photographing and dictating. There is no doubt that this multi-functionality was important to the success of e-scape, but it is also true that the device was sometimes found to be constricting. The small screen inhibited some in their drawing and writing and it was difficult to use the screen to glimpse a big picture of the whole portfolio.

For these reasons we are interested in the future direction of this technology, and the following examples illustrate the way things are moving.

Virtual – projection – keyboards

The technology already exists to project a virtual keyboard onto a flat tabletop surface. The projection is of a full-size qwerty keyboard and typing onto it is uncannily like typing on a real keyboard – including the 'click' as the key is struck. The projection can be done from a matchbox size device and the signal can be blue-tooth connected to the PDA. So in effect the PDA has acquired a full size keyboard. This already exists.

See

http://www.alpern.org/weblog/stories/2003/01/09/projectionKeyboards.html

We can however speculate on an extension of this idea. If the projector can project a keyboard there is no reason why it could not also project a blank sheet onto the tabletop. In this situation one might draw with a pencil on a 'virtual' plain sheet of paper and have it 'sucked up' by the projection sheet as a digitised drawing. As with them keyboard, the drawing could equally be transferred to the PDA by blue-tooth. The PDA is therefore no longer quite so constricting as a drawing tool, since, just like the keyboard, the digital sketchpad has become external to it.

Speech-to-text systems

During the early stages of the project we experimented with speech to text systems and Chatterbots. The processing power and software developments for/on desktop computers are such that these systems are now viable. Presenting young people with the facility to talk their ideas into a web portfolio and then the option to access this evidence as audio or

automatically generated transcripts would be liberating for some learners. While we can demonstrate how this system might work, we will have to wait a while for it to be available for groups, since identifying individual voice profiles in a multiple-voice free flowing discussion is currently beyond the scope of the technology.

Mini-projection

We are all familiar with data projectors and with the fact that they continue to get smaller and smaller. Five years ago the smallest were like cornflake boxes, but now the smallest are more like individual cornflakes.

This photo is of a prototype from "Light Blue Optics", and the coin is a two pence piece. They speculate on the adoption of this miniature laser projection technology into the PDA or mobile phone and conjure up a vision of the data projector being inside each hand-held device

See http://www.lightblueoptics.com/

In this situation, the whole of a learners' portfolio could be projected onto a screen or wall and viewed as a whole – rather than as a series of PDA screen-sized bites.

So both the input and output systems of the PDA could - in a year or so - be external to the PDA – which becomes merely the processor.

PDAs and mobile phones

When we were planning this project in 2004 the learners we met in schools did not typically have cameras in their mobile phones and indeed many could not see why you would want one. Now it is difficult to buy a phone without a camera and most have good resolution. In the summer pilot, several learners had phones with 'smart' features and many reported the phones in their camera to be better quality than e-scape PDA cameras. Within another 2 years it will be as difficult to buy a non-smart phone as it is to buy a phone without a camera today, and given the ubiquity of I-pods, we can only imagine what the i-phone will do to change expectations of mobile handsets. The point is that whatever we imagine, it will be different when it arrives, and we need to help teachers and schools to get better at embracing and harnessing this liberating but unstable world. The 'personalised learning' agenda that is currently receiving so much attention from DfES and elsewhere leads naturally to a personalised and ubiquitous computing imperative. Mobile, back-pocket, technologies are inevitably going to flourish in this world.

ii) the software (e-scape application)

The e-scape activity in the classroom is managed through the e-scape application, a piece of software that was developed during phase 2 and runs on the PDA. The application is best

described as digital sellotape that binds together some elements of pre-existing software (for drawing / writing / photography and dictation) into a form that could facilitate the e-scape activity.

We have described above how the hardware may well develop in the next few years, and the software is equally capable of development. Perhaps most critical would be the creation of *an authoring tool* that allows teachers to build activities of their own choosing rather than being entirely constrained by the design of the activity we built into the software. We can imagine activities in science, history, music and geography that have an e-scape-like framework of sub-tasks and timings. Such an application could be enormously empowering for teachers, putting them in the driving seat and allowing them to develop and customise activities for their own setting, their own learners, and their own timescales.

2.19 conclusions and next steps

Conclusions

The conclusions to this project are best presented in relation to the four strands of research question that we posed at the outset. These questions have directed our work throughout the project.

Concerning **technological** challenges (e.g. the possibility of connectivity between hand-held devices and websites) we overcame these and succeeded in creating a prototype system. The system was sufficiently robust to be taken into 14 schools during the national pilot in which 300 learners undertook studio / workshop activities and successfully uploaded their portfolios into the website. We have identified areas in which we would choose to develop the hardware and software for future activities, but even as it stands it would have to be regarded as a technological success, and indeed a major technological innovation.

The key point here is that the whole system is driven by a remote server dynamically sending and receiving data to and from the PDA, putting the teacher in control of the sequences of the task and automatically building an evidence trail in the web portfolio.

Concerning **pedagogic** challenges (e.g. the most effective structuring of activity sequences to maximise learners' performance) we resolved these through an extended set of trials of the task, the activity structure, the booklet, the timings and the resources. Each school trial was focused on teasing out a particular set of issues and enabled us to arrive at an appropriately satisfactory end point. The best evidence for its success is that the learners all responded to the activity so well. Whilst the technology was no doubt part of its attraction, the task succeeded in engaging them and the activity enabled them to demonstrate their capability.

The key point here is that everything we did for the purposes of collecting evidence for assessment *also* helps to scaffold the progress of the activity for learners.

Concerning the ***manageability*** challenges (eg is it possible for all learners to have ubiquitous access to digital tools in a normal workshop setting), we again sought to deal with these through the trialling process. We were aware for example that learners would need training with the PDA – but did not know how long or how much. We are aware of the state of most workshop environments and were interested to explore the robustness of the PDA in these potentially harsh environments (e.g. getting dropped on concrete floors). We started with 100 devices and after all the trialling and the national pilot had one screen broken and none lost. The reaction of teachers both to the PDA and the activity was enthusiastic. We have established that the approach adopted for e-scape was indeed manageable for learners, for teachers and for the research team.

The key point here is the infusion of technology into activity. Real-time activity in studios, workshops, playing fields, theatres, science labs and the like, is typically not aligned with digital power. This power typically sits in splendid isolation in the shimmering purity of IT suites. In e-scape we have shown how the technology can get down and dirty and unleash its digital power where it is really needed. And in the national pilot we demonstrated that it was manageable.

Finally, concerning the ***functionality*** of the assessment system, it is perhaps here that the most dramatic conclusions might be drawn. The e-scape approach enables learners to create web-based portfolios directly from their classroom design activity. The web-based nature of the portfolios has in turn enabled us to explore a quite new paradigm for assessment – comparative pairs. The direct connection to real-time activity in the studio/workshop supports the *validity* of the assessment, and the comparative pairs model of assessment enabled us to achieve high levels of *reliability* in the assessment judgments. Moreover the system reports on the effectiveness (consensuality) of each of the judges, and we were all well within acceptable tolerances.

The key point here is that *performance* assessment is notoriously difficult, and at both ends of the system. It is difficult to manage the performance itself in ways that assure equity to all learners and it is difficult to ensure reliability in the assessment. Within e-scape we have created a system that operates right across this terrain. Learners and teachers report that it worked well at the performance end, and the data shows that it produced reliable statistics at the assessment end. The prototype has done enough to demonstrate that it is a functional system for assessment.

Next steps

As we write this report, we are in discussions with DfES, QCA, Becta and the Awarding Bodies concerning the new directions that might be taken with project e-scape. It is clear that the system works and opens up many possibilities for development – both for teaching & learning and for assessment – and the challenge is to decide on the most appropriate next step.

The two major innovations in e-scape have been
• hand-held digital tools linking dynamically to a website to create portfolios
• enabling 'comparative-pairs' judging for reliable assessment

Whilst the 1st one could stand alone – and might be developed by further explorations of hand-held digital tools linking to websites - the 2nd is naturally linked to the 1st, since it is only possible because of the 1st. And both innovations are essentially about **access**.

This 1st innovation enabled us to capture genuine designing/problem-solving classroom activity directly into the website. Every learner had direct access to significant digital processing power in the midst of (and throughout) the workshop activity and without being tied into a computer lab or IT suite. They were liberated to operate as autonomous designers but with 'back-pocket' access to the website. Thereafter, since the judges had immediate access to all the portfolios, all the time, we could – for the 1st time in assessment history - exploit the possibilities of comparative pairs judging for 'front line' assessment.

It is these two innovations that will be at the centre of the proposed next step for e-scape.

With the 1st innovation, we see no reason in theory why any kind of activity-based assessment task could not be substituted for the design & technology task that we developed. English composition, science investigation, or music performance might be the focus of such development and we propose to explore the extent to which the system can be made to operate across disciplines. As part of this, we will create an 'authoring' interface that allows the teacher to build an activity through a series of sub-task steps, and each with a specified timing.

The 2nd part of the 'next step' concerns the integration of the e-scape system into Awarding Body data management systems. For the assessment to work as genuine front-line assessment it has to be seamlessly linked to a national system that allow schools to 'enter' candidates for the assessment. Thereafter it has to enable their performance to judged and the outcome to be managed once again through the awarding process that results in individuals achieving an authenticated grading from an Awarding Body.

We propose to develop a scalable national system of assessment built around – and linking together - these two priorities. The matter has been given added urgency by the current concern with GCSE coursework assessment, which has increasingly been criticised for its lack of trustworthiness in terms of activity administration (whose work is it) and in terms of the reliability of assessment judgments.

The proposed next step would enable us to retain some of the important aspects of coursework, particularly the necessity for learners to tackle *real* tasks – over *time* – and evolve *individual* solutions. But would enable this to happen in a tighter framework of school-administered activity. Thereafter, learners' performance can be judged – using a form of

comparative pairs – to arrive at a highly reliable assessment outcome. The two major drawbacks with coursework assessment are thereby nullified.

Discussions are advanced for this new project – **e-scape phase 3** – to run from Feb 2007 to March 2009, by which time a scalable national system will be operational.

End……….

references

Black P, Harrison C, Lee C, Marshall B, Willam D 2003 *Assessment for Learning: putting it into practice* Buckingham: Open University Press ISBN 0-335-21297-2

Black P & Harrison C 2004 Science inside the black box: assessment for learning in the science classroom London: NFER Nelson ISBN 0 7087 14447

Department for Education and Employment (DfEE) 1999 *Design and technology*: *The National Curriculum for England:* Department for Education & Employment (DfEE) and the Qualifications and Curriculum Authority (QCA). London

Department for Education & Skills (DFES) 2003 *Survey of Information and Communications Technology in Schools* Oct 2003 HMSO - available at http://www.dfes.gov.uk/rsgateway/DB/SBU/b000421/index.shtml

Department for Education & Skills (DFES) 2005 *Harnessing Technology: Transforming Learning and Children's services: DFES e-strategy* - available at www.dfes.gov.uk/publications/e-strategy

Haste H. 2005 *Joined-up texting: The role of mobile phones in young people's lives* Haste H. Director of Research: Nestlé Social Research Programme Nestle - available at http://www.mori.com/polls/2004/nestlesrp3.shtml

IMS Global Learning Consortium Inc. Sept 2004 IMS ePortfolio Best Practice and Implementation Guide available at http://www.imsglobal.org/

Kimbell R Stables K Wheeler T, Wosniak A, Kelly V. 1991 *The Assessment of Performance in Design & Technology: the final report of the APU design & technology project*. SEAC and COI for HMSO (D/010/B/91)

Kimbell R, Miller S, Bain J, Wheeler T, Wright R, Stables K. 2004 *Assessing Design Innovation: a research & development project for the Department of Education & Skills (DfES) and the Qualifications and Curriculum Authority (QCA)* TERU Goldsmiths

Laming, D. (2004) *Human Judgment: the eye of the beholder*. London, Thomson.

Nesta Futurelab (Ridgway J, McCusker S and Pead D) 2005 *Literature review of E-assessment Research Report 10* Nesta Futurelab. - available at
http://www.nestafuturelab.org/research/reviews/10_01.htm

Office for Standards in Education (OFSTED)   2003   *Good assessment practice in design and technology* Ofsted publications – available at
http://www.ofsted.gov.uk/publications/index.cfm?fuseaction=pubs.summary&id=3208

Office for Standards in Education (OFSTED)   2004   *ICT in schools – the impact of government initiatives:*
Secondary design and technology May 2004   OFSTED publications - available at
http://www.ofsted.gov.uk/publications/index.cfm?fuseaction=pubs.summary&id=3649

Office for Standards in Education (OFSTED)   2004 (ii) *ICT in schools: The impact of government initiatives 5 years on*   May 2004   OFSTED publications - available at
http://www.ofsted.gov.uk/publications/index.cfm?fuseaction=pubs.summary&id=3652

Prime Minister's Strategy Unit   2005 *Connecting the UK: the Digital Strategy*. A joint report with Department of Trade and Industry - available at
http://www.strategy.gov.uk/work_areas/digital_strategy/index.asp

Pollitt, A (2004). "Let's stop marking exams". Paper given at the *IAEA Conference*, Philadelphia, September. Available at:
http://www.cambridgeassessment.org.uk/research/confproceedingsetc/IAEA2004AP

Qualifications & Curriculum Authority (QCA) 2004   *E-assessment expert seminar* BECTa 9th Dec 2004

Qualifications & Curriculum Authority (QCA) May 2005 *Assessment for Learning*: website
http://www.qca.org.uk/7659.html

Thurstone, LL. (1927) "A law of Comparative judgment". *Psychological Review, 34*, 273-286

Tomlinson Report 2004 *14-19: Curriculum and Qualifications reform. Final report of the Working Group on 14-19 reform* DfES Publications DFE-0976-2004

Other related and relevant materials and sources

Teens and New Technology – Latest Data on Mobile Phones
by Rosemary Duff, ChildWise
2005
http://www.childwise.co.uk/mobiles.htm

Literature review of Mobile Technologies and Learning

NESTA Futurelab Report 11 2005?
Laura Naismith, Peter Lonsdale, Giasemi Vavoula, Mike Sharples
http://www.nestafuturelab.org/research/reviews/reviews_11_and12/11_01.htm

Beyond the Electronic Portfolio: A Lifetime Personal Web Space
By Ellen R. Cohn and Bernard J. Hibbitts   website
2005
http://www.educause.edu/apps/eq/eqm04/eqm0441.asp?bhcp=1

fd Learning e-portfolios their use and benefits
Bob Banks
June 2004
http://ferl.becta.org.uk/display.cfm?resID=8099

QCA website
http://www.qca.org.uk/7659.html