Vol. 20.3 October 2015

Design and Technology Education: An International Journal

20.3

Design and Technology Education: An International Journal

Design and Technology Education: An International Journal

Design and Technology Education: An International Journal (formerly The Journal of Design and Technology Education) is published three times a year.

Contributions

All contributions and materials for review should be addressed to The Design and Technology Association, 16 Wellesbourne House, Walton Road, Wellesbourne, Warwickshire, England CV35 9JB.
Tel: 01789 470007 Fax: 01789 841955
email: **Neil@data.org.uk**

D&T Association membership

To join the D&T Association, or request further details on membership please contact Jacqui Eborall, Membership Secretary, The Design and Technology Association, 16 Wellesbourne House, Walton Road, Wellesbourne, Warwickshire, England CV35 9JB.
Tel: 01789 473902 Fax: 01789 841955
email: **jacqui.eborall@data.org.uk**

Design and Technology Education: An International Journal is available in the online shop at www.data.org.uk

If you wish to receive **Design and Technology Education: An International Journal** independently, please contact Central Books Ltd, 99 Wallis Road, London, E9 5LN
Website: **www.centralbooks.co.uk/**
email: **mo@centralbooks.com**

Printed by Renaissance, Coventry

Volume 20
Number 3
ISSN 1360-1431 (print)
ISSN 2040-8633 (online)

October 2015

Design and Technology Education: An International Journal Contents

Special TERU Edition (20:3)

Prof Richard Kimbell and Kay Stables, Goldsmiths, University of London

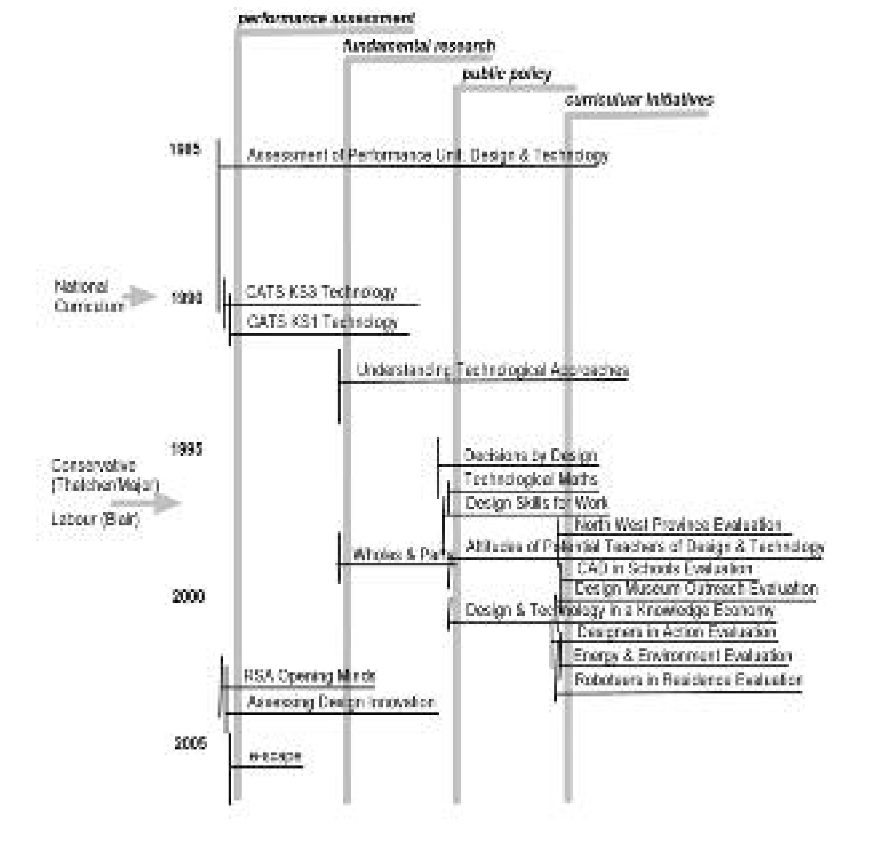
This is an unusual edition of the journal. The Technology Education Research Unit (TERU) at Goldsmiths, University of London has been invited by the D&TA team and the journal's editors to create a Special Edition of the journal. Specifically (and for a reason that will become clear later) we have been invited to compile an edition from papers that we have written based on our research over the last 30 years. One of these papers appeared in an international journal; two in a book that we wrote together that was published in 2007; two were keynote presentations at conferences in South Africa and Australia; and one (oddly) has never before seen the light of day. We have made our selection on the basis of bringing into one edition things that we have written over the years that we consider have particular relevance in today's educational climate.

We are using this editorial to set the scene for the collection by summarizing and categorising the entire body of research that we have undertaken at Goldsmiths, and we need at the outset to clear up the start-date for this body of work. Whilst TERU was formally established in 1990, our research activities really started before that with the APU project (1985-91). So in total our research spans a period of 30 years, 25 of them as TERU. We have structured the story so as to be broadly chronological and interestingly this chronology also reflects a series of shifts in the nature of the work. Originating in research concerning *assessment* we moved progressively through phases of *fundamental* research, *public policy* and *curricular initiatives* before returning once again to *assessment* priorities. We have mapped this chronology in the graphic that appears at the end of this editorial.

TERU AND PERFORMANCE ASSESSMENT

In 1984, the UK Department of Education & Science announced design & technology as a new field of enquiry to be tackled by its research branch, the Assessment of Performance Unit (APU). Established in 1975, the APU's prime task was surveying and monitoring levels of achievement in schools. By the time the design & technology contract was issued, it had conducted extensive surveys in mathematics, English, science and modern languages, typically at ages 8, 11 and 15. Much had been discovered about what learners could be expected to achieve in these subjects at those ages. Progressively, however, a change of focus was detectable in the conduct of those surveys. APU began to focus less on mere monitoring, and more on providing support for curriculum development.

Early APU surveys were seen largely as providing data about what learners could or could not do – and how this changed over time. In curricular terms APU was distinctly non-interventionist. Progressively however, the concern became to understand why learners performed in the ways they did; teasing out learning blocks and helping teachers to enhance learning. APU was increasingly becoming a force for curriculum development. (Kimbell et al., 1991, p. 11) With the 1984 announcement that APU



TERU Projects timeline

wished to survey design & technology, tenders were invited. The contract to undertake the research was won by Goldsmiths.

The proposal enabled a research team to be created in the design & technology department at Goldsmiths. This team was directed to Professor Vic Kelly (a curriculum specialist) and the research was coordinated by Richard Kimbell (a lecturer in design & technology). At the launch of the project, the team additionally comprised Kay Stables (a specialist textiles teacher), John Saxton and Jim Patterson (both craft, design & technology teachers). Other appointments were made during the subsequent 5 years. We found new ways to describe the domains of performance in design & technology and developed new approaches for supporting and enriching learners' performance. We developed this approach into 26 tests that we took into 700 schools across England, Wales and Northern Ireland, and in total we assessed the performance of approximately 10,000 learners. The resulting performance data were analysed from many perspectives, and the final report contained national performance levels analysed in relation to gender, ability, and the curriculum that had been experienced by the learners. We also revealed generalised features of design & technology activities that have serious effects on performance levels, such as the nature of tasks and their contextual setting as well as the structures of activity through which learners tackle those tasks. The full research report was published in 1991 (Kimbell et al., 1991).

But before then, in 1989, other research ventures were appearing on the horizon – not least concerning the planned implementation of design & technology in the National Curriculum. With the imminent prospect of a number of new research and development projects coming into the (re-named) Design Department at Goldsmiths, in 1990 Richard created TERU – the Technology Education Research Unit, as a Unit within which we could draw together all these research and development activities in support of design & technology in schools.

On the strength of *APU Design & Technology*, we acquired three new projects – two of which centred upon approaches to the performance assessment of learners in design & technology classrooms, workshops and studios. Specifically, we were invited to create prototype tests for National Curriculum design & technology – at age 14 (1989–1992) and at age 7 (1990–1992). Both these projects took further the models of research that had been originated within *APU Design & Technology*; the age 14 project being directed by Jim Patterson, and the age 7 project by Kay. Richard directed the third project – developing curriculum support materials for design & technology for the newly created National Curriculum Council – alongside the preparations for publication of the *APU Design & Technology* report.

THE NEED FOR FUNDAMENTAL RESEARCH

APU Design & Technology had been the first large-scale research to be undertaken in design & technology. The subject itself was a new concept – drawn together through a series of curriculum initiatives that gradually coalesced into design & technology in the late 1980s. Plenty of curriculum development projects had taken place in these evolutionary years, but nothing of a fundamental nature to enable the design & technology community to create the conceptual underpinning that is necessary for real understanding of a subject. Design & technology – at this time – was best described as 'what was done' by a group of practitioners who shared a set of ideals about teaching and learning in workshop and studio settings.

In our own national context, these ideals and practices had been rationalised (in 1985) as part of the revision of 16+ examinations. Prior to this point, there had been a twin system of qualifications at 16+; the General Certificate of Education (GCE), for the 'top' 25% of ability of the population, and the Certificate of Secondary Education (CSE) for the rest. In 1985 these two systems were merged into the General Certificate of Secondary Education (GCSE) and the opportunity was also taken to consolidate and update the content of the subjects to be examined.

Two of those GCSE subjects, Craft Design & Technology (work in wood, metals and plastics, graphics and technological systems) and Home Economics (work in food, textiles, child development and home management) were the core of what was subsequently to become design & technology. In both groupings, the role of **designing** was accentuated, and this subsequently became the organising feature that dominated design & technology when it was launched as a 'new' subject as part of the first England and Wales National Curriculum. This new subject drew from all its founding formulations, most notably Craft Design & Technology and Home Economics, but there was at least as much doubt and confusion about its composition and practices as there was clarity and light. The formulation of National Curriculum Programmes of Study and Attainment Targets – built around designing and making – forced the amalgamation of these two groupings into design &

technology as it now (broadly) exists. The disparate traditions and practices created enormous tensions within design & technology. The situation cried out for some fundamental research that could build a conceptual framework to make sense of the beast that had been created.

In 1991, Richard applied to the Economic and Social Research Council (ESRC) for a grant to fund a project to explore – and seek to understand – the practices that proliferated at this point. In 1992 the ESRC approved the award and a new 2-year project was launched within TERU: *Understanding Technological Approaches* to teaching and learning in the curriculum.

In this project we explored in detail real-time projects in design & technology at every school year from year 1 to year 11 in the new National Curriculum (i.e. with learners from age 5 to 16) in every area of design & technology. The approach was broadly to **observe** projects from start to finish – usually 3-4 hrs with year 1 and 2 but as long as 48 hrs with year 11. The observations were built around as common framework – enabling us to make direct connections between the approaches to designing and making across this complete age range.

Analysing these detailed observations (taken over 2 years) enabled us to characterize approaches to design & technology teaching & learning, and describe it in ways that had hitherto not been possible. We published this work in 'Understanding Practice in Design & Technology' (Kimbell et al., 1996).

THE DEMANDS OF PUBLIC POLICY

By the mid-1990s design & technology had become a fixed point on the educational landscape. Having escaped from the obscurity imposed by its fractured history, design & technology – as a single entity – began to assert itself into areas of public life. All kinds of issues began to emerge with interested professional bodies, not least with the UK Design & Engineering Councils, both organisations with certain responsibilities for managing, promoting or regulating their professions who also have a brief to inform and educate the general public about their activities. Particular interest in design & technology is related to:

- Its role as a university entrance qualification
- Its employment value for school leavers
- Its role as an economic driver in a knowledge-economy
- The challenge of recruiting and training teachers

From 1995, we were approached on a range of these issues to run projects that could illuminate areas of public policy. The first of these arose through the Design Council, building case studies of 'good practice' so as to exemplify what was meant by design & technology. However, the bodies for these public policy projects were typically **less** concerned with developing good practice in schools, and more concerned with understanding the distinctive contribution that design & technology could make in areas of public and professional life. Their priority was to seek **conceptual** clarity.

We presented a case to the Design Council, that **designing** is a distinctive way of thinking, and they awarded us a grant for a 2-year project exploring exactly that territory. The project *Decisions by Design* (1995–1997) explored the power of designerly thinking for those who are not (and do not intend to become) designers. How is design thinking similar to and different from 'ordinary' thinking? What is its distinctive character? The successful conclusion of this project led to further projects in the general area of transferable design skills for employment. The first, *Design Skills for Work* (1997–1999), addressed the general question 'what are designers good at, if they are not being designers?' This was followed by a project exploring the attitudes of design students towards a career in teaching – *Attitudes of Potential Teachers of Design & Technology* (1999–2000).

At the same time the Engineering Council – interested in routes from school into engineering – was concerned to explore the role of mathematics in design & technology. The serious drop-off of candidates coming forward with pure and applied mathematics and physics, along with the increasing awareness of the engineering nature of some design & technology, had encouraged some universities to seek students who had successfully completed design & technology Advanced Level examination courses. The project *Technological Maths* – seeking to identify the nature and extent of the mathematics in design & technology – ran in TERU from 1996–1997. A second project for the Engineering Council – *Design & Technology in a Knowledge Economy* (2000–2001) – aimed to locate design & technology within the wider debate about the need for curriculum change to support future knowledge economies.

Towards the end of the 1990s, the National Curriculum formulation of design & technology had worked its way through the entire school population, primary and secondary. It had evolved through two official versions (1990 and 1995, and the 2000 version was looming) as well as a number of unofficial ones, inspired by particular interest groups. A centre of gravity had emerged for the subject, consolidating into forms of classroom and workshop practice that were more commonly understood and accepted. So changes at this point were destined to be less sweeping and more incremental – tweaking the formula rather than slinging it out of the window.

So the need for evidence about the performance of particular approaches to learning and teaching within this curriculum became ever more necessary and in TERU we became involved in all kinds of evaluative projects – seeking to understand and make evident the particular strengths and weaknesses of this or that curriculum initiative or approach.

EVALUATING CURRICULAR INITIATIVES

Ironically, the first of these evaluation exercises was for a foreign government. The presence of design & technology in the UK had for some years been exerting an influence on the international scene, and the consolidated form of National Curriculum design & technology had been influential, especially in the English-speaking world where UK journals and conference speakers were available.

It was the new Mandela administration in South Africa that invited TERU to undertake its first evaluation of a curriculum initiative, funded by the Department for International Development (DFID). In the North West Province – centred on Mafikeng – the provincial curriculum team, in association with a non-governmental organisation (NGO), had undertaken a pilot study to introduce a technology education curriculum for learners in their final 2 years of schooling. The scale of the challenge of undertaking this curriculum in rural schools in South Africa is difficult to imagine in more 'developed' countries:

- Schools with minimal facilities and (sometimes) no electricity
- Involving teachers from subject backgrounds as diverse as geography and Afrikaans
- Traveling huge distances to attend training sessions
- Training for a curriculum that was dramatically different from former (craft) practice
- Resources brought into the schools by van across huge distances
- With the curriculum expert (the van driver) visiting perhaps twice a year

Our evaluation of the curriculum and of the Province's procedures for developing and disseminating it became part of the wider South Africa education debate when technology was absorbed into their national curriculum framework.

Other evaluation projects followed; for London's Design Museum, exploring the effects of their educational outreach programmes; for the Design & Technology' Association (D&TA), evaluating the impact of Pro-DESKTOP computer aided design software; for the National Endowment for Science Technology and the Arts (NESTA), developing a new **systems and control** curriculum with LEGO soft and hardware; for Middlesbrough Local Education Authority (LEA), evaluating literacy developments through design & technology in primary schools; and for the BBC, evaluating their Roboteers in Residence programme that brought expert roboteers into schools to work with learners developing robots for a BBC TV programme.

THE NEW MILLENNIUM

In 2000, a number of related events took place that shaped the activities of TERU over the following 5 years. The latest version of the National Curriculum (NC2000) was launched, with some amendments to the Programmes of Study and the Attainment Target. Most critically, however, it included for the first time a statement about the importance of design & technology in the curriculum. It may seem odd that such a 'vision statement' should not be published until a decade after the original launch of design & technology in the 1990 National Curriculum. The recognition of this need for a clear statement of intent was reflected right across the curriculum – from all subjects – and these statements were drafted with expert subject groups in 1999 as cornerstones for the launch of the fully revised curriculum.

However, the issue ran deeper for those of us concerned with learning through design. The tortuous history of design & technology, and the rapid evolutionary steps that it had progressed through in the decade immediately prior to the establishment of the National Curriculum in 1990, all contributed to the recognition – in the UK Government Department for Education & Employment; in D&TA (the Design and Technology Association), the subject's professional Association; and in Higher Education and teacher education establishments – that the newborn baby would need careful nurturing in the immediate years ahead. Accordingly, the Department for Education & Employment established a Design & Technology Strategy Group to oversee these years and to bring forward recommendations for the immediate future.

One of the earliest tasks undertaken by this group was to analyse the internal coherence of design & technology as presented in its revised version, and specifically in relation to the 'fit' between the newly created vision statement and the Programmes of Study and the Attainment Target, both of which had evolved through three versions of the National Curriculum. Some discrepancies became apparent. Among these was the recognition that whilst the vision accentuated the importance of developing learners' creativity and innovation, and significantly through the vehicle of teamwork, teachers – particularly through the assessment criteria for the GCSE examinations – were not required to acknowledge or reward these qualities.

In the light of these mismatches, TERU was commissioned to undertake a project to reinvigorate the creative heart of designing and develop approaches to the assessment of design & technology that would reward teamwork and innovation.

PERFORMANCE ASSESSMENT AND INNOVATION

In January 2003, we launched the project *Assessing Design Innovation* and in many ways this drew TERU back to its origins in the Assessment of Performance Unit in the mid-1980s. We were back to exploring approaches to performance assessment in design & technology, but with the additional requirement that the approaches we developed should be focused on supporting teamwork and enhancing learner innovation.

But by now we had a great deal more experience of research and development approaches. We were able to draw on the wide range of techniques that we have developed in our earlier work:

- Exploring the nature of design & technology
- Supporting the development of public policy
- Evaluating curriculum initiatives

Over 2 years from January 2003 to December 2004 we worked with a small number of LEAs and schools across the country, and produced models for assessing design innovation that were subsequently not only reported to the (now renamed) Department for Education and Skills and its curriculum and assessment 'watchdog' the Qualifications and Curriculum Authority, but were also shared with the General Certificate of Secondary Education Awarding Bodies. One of the immediate outcomes of this project was the development by one of these awarding bodies of a new form of syllabus and examination based on the approach we had developed in the project. (See OCR Product Design… 'The Innovation Challenge')

In the process of developing our approach to assessment in this project, we explored a range of new technologies to see how they might be helpful. Among these technologies were the use of digital cameras to record learners' emerging work, and of some simple computer aided design interfaces to support their ideation. It became apparent to us that these digital technologies offered the potential radically to transform the assessment process, and we proposed to Qualifications and Curriculum Authority and the Department for Education and Skills that these technologies should be the explicit focus of a research and development project. This proposal came simultaneously with the challenge to the examination Awarding Bodies to accept portfolios on disk. This was – at one level – a natural evolution of good design & technology practice, but – at another level – represented a serious challenge to the established assessment procedures of the Qualifications and Curriculum Authority.

In the light of all these pressures, our proposal was accepted and project e-scape ran through three research and development phases. E-scape Phase 1 (2004-5) was a feasibility phase – looking to see how the digital technologies available at the time might (just possibly) be made to work for learners in classrooms. E-scape Phase 2 (2005-7) was then commissioned – to build a working prototype system that would allow learners in normal studios and workshops to create web-portfolios of their work, and for teachers to be able to undertake web-based assessments of them. Phase 3 (2007-10) was then commissioned as a national pilot-testing programme in association with Awarding Organisations. In total this involved 19 schools and 350 learners (mostly in year 10) and the assessment technology in particular brought a completely new set of tools to teachers making assessments and Awarding Organisations seeking reliable means for awarding grades. We developed Adaptive Comparative Judgement (ACJ), that represents a radically new assessment methodology that we believe is immensely valuable both to teachers and learners.

Stepping outside the boundaries of design & technology was also a feature of a further performance assessment project that we undertook in parallel with *Assessing Design Innovation*. This project, commissioned by the Royal Society for the Arts (RSA), was aimed at exploring approaches to assessing generic competences such as team-working, systematic thinking and managing risk that were being developed through a further RSA project 'Opening Minds: Education for the 21st Century' (Bayliss, 1999). The TERU project, *Researching Assessment Approaches*, was conducted during 2002–2003. Meanwhile, the initial *Assessing Design Innovation* project materials were being utilised in collaborative work with the University of Strathclyde (McLaren et al., 2006) and the Stockholm Institute of Education (Skogh, 2005). The e-scape project was also further developed in other national contexts – in Scotland as *e-scape Scotland* (McLaren, 2012), and in Israel as *Assessment in my Palm*, (Stables and Lawler, 2011; 2012) in both settings with greater emphasis on formative assessment and in Australia where e-scape was used in school examinations for engineering.

Finally, in 2014 we launched a new project that grew from the earlier e-scape projects. Whilst their focus had primarily been on summative assessment – since the brief from QCA related specifically to the award of GCSE – we were aware that much of the power of the approaches that we had developed lay in the *formative* benefit that they held for learners. The '*Formative Assessment*' project is in association with the UK government's innovation agency, Innovate UK, whose brief is to support innovation in the application of new technologies. The project focuses on 'design talk' (discussions between teachers and learners) and the role that technology – specifically artificial intelligence – might play in enriching it.

THE EMERGING STORY OF TERU

The major blocks of research and development outlined here, that we have undertaken within TERU over the last 30 years, were not consciously planned out from the start. But neither were they arbitrarily taken on.

The APU starting point in 1985 was unexpected, and was undertaken with more enthusiasm for design & technology than expertise in assessment research. We have progressively acquired that expertise. But after that first project for APU, the priorities for our subsequent work have reflected the concerns of a new subject emerging into the spotlight of National Curriculum from the relative obscurity of a collection of historical and typically unregarded and undervalued subjects.

One of the biggest difficulties for the new fledgling design & technology was that there was almost nothing in the way of research upon which to base decisions about curriculum, or pedagogy, or assessment. Practice in schools therefore emerged on the basis of hunches and best guesses and things that had worked in the past. There was painfully little foundation on which to build a coherent and progressive vision of design & technology.

> Design & Technology lacks a research base in pupils understanding and learning such as is available in the cases of mathematics and science.
> (DES/WO, 1988, p. 7)

> Craft Design & Technology stands out as the most under researched area of the curriculum. The literature of the subject barely exists.
> (Penfold, 1988, preface p. ix)

TERU was established in response to these challenging observations. Moreover, it was founded on the belief that learning in and through design & technology has some features that make it unusual in the curriculum, and that enable it to contribute positively and uniquely to the education of young people. The research and development that we have undertaken has been informed by this belief and has sought to throw light onto the traditions and practices of teaching and learning in design & technology workshops, studios and classrooms.

In 2008, Richard and Kay compiled a book to tell the full story of our research endeavors in TERU. The book is entitled *Researching Design Learning* (Springer 2008), and this introductory piece – and the chapter on research methodology – are taken from that book. We are grateful to Springer for allowing us to use these pieces.

References

Bayliss, V. (1999) *Opening minds: education for the 21st century*, London: RSA.

DES/WO (1988) *National Curriculum Design and Technology Working Group interim report*, London: Department of Education and Science and the Welsh Office.

Kimbell, R., Stables, K. and Green, R. (1996) *Understanding practice in design and technology*, Buckingham UK: Open University Press.

McLaren, S. (2012) 'Assessment is for learning: supporting feedback', *International Journal of Technology and design Education*, 22(2), 227-245.

McLaren, S. V., Stables, K. and Bain, J. (2006) *Creativity and Progression in transition through Assessment for Learning in Design and Technology Education (CAPITTAL-DT): A pilot project Report to Determined to Succeed Division*, Glasgow: University of Strathclyde.

Penfold, J. (1988) *Craft Design & Technology: Past Present and Future*, Stoke-on-Trent: Trentham Books.

Skogh, I.-B. (2005) 'Innovative performance – how can it be assessed?' in Norman, E. W. L., Spendlove, D. and Grover, P., eds., *Inspire and Educate: DATA International Research Conference 2005*, Wellesbourne: DATA, 161-166.

Stables, K. & Lawler, T. (2011) *Assessment in my Palm: e-scape in Israel evaluation of Phase 1*, London: Goldsmiths, University of London.

Stables, K. & Lawler, T. (2012) *Assessment in my Palm: e-scape in Israel evaluation of Phase 2*, London: Goldsmiths, University of London.

Williams, P. J., (2012) Investigating the feasibility of using digital representations of work forperformanceassessment in engineering. *International Journal of Technology and Design Education*. Vol 22.2, p.187-203

A Tale of Three Pilgrims

Eddie Norman, Emeritus Professor of Design Education, Loughborough University

RESEARCH

I had the title 'A Tale of Three Pilgrims' in mind when I thought about an Inaugural Lecture following my appointment as Professor of Design Education towards the end of my career at Loughborough Design School. Regrettably the appointment had been delayed by several years and by the time the invitation to give the lecture arrived I had retired and a combined Inaugural and Valedictory Lecture did not seem appropriate. Still you never know when things are going to 'come in' and the title works just as well for the Design and Technology Association's invitation to write a final Reflection piece for the journal. The Three Pilgrims are of course guitars! The instruments span my Loughborough career and are symbolic of conceptions of design innovation that have been at the heart of developments during this period.

I joined the Department of Design and Technology at Loughborough University in 1984 and a few years previously, and few miles to the south in Leicester, Paul Tebbutt was perfecting his design for the Pilgrim guitar shown in Fig 1. Having obtained financial backing and brought-in a technical manager, Paul Tebbutt continues the story of the Pilgrim as follows:

'Then there followed a period culminating in exhibiting at the Frankfurt Trade Fair where the World's music dealers and press gave us unprecedented approval; we returned to England knowing that we held the key to success.

Figure 1 Paul Tebbutt's Pilgrim Guitar (1980)

There then followed a difficult two years during which we set about the task of finding craftsmen equal to our desired quality and we have succeeded in gathering together a team of the highest calibre.

The Pilgrim Guitar is craftsman built to a degree of perfection we are sure you will find truly pleasing. Every aspect of our guitars has to meet a standard of quality control seldom found in today's hurried world.'

Pilgrim used the highest quality materials and the most exacting manufacturing methods, and the guitars were correspondingly expensive. I bought mine second-hand, but it still cost rather more than I should have spent. Nevertheless it is still doing sterling service and being played most days by a friend of my daughter. The 1980s were a tumultuous period as the UK struggled to improve its international competitiveness and 'craft' and 'quality materials' as a basis for design innovation were already in retreat. There was a drive towards achieving 'quality through design' and investment in automated manufacture to improve productivity. In general education in the UK, Craft, Design and Technology (CDT) had lost 'Craft' from its title by the end of the decade. Amongst my first responsibilities at Loughborough was being in charge of the Machine Shop and the associated initial training, so I am sure I was at least ambivalent about some of the changes, having been a great admirer of the high level of craft skill that the Metalwork and Woodwork A-level students brought with them to Loughborough. Which brings us to our second Pilgrim.

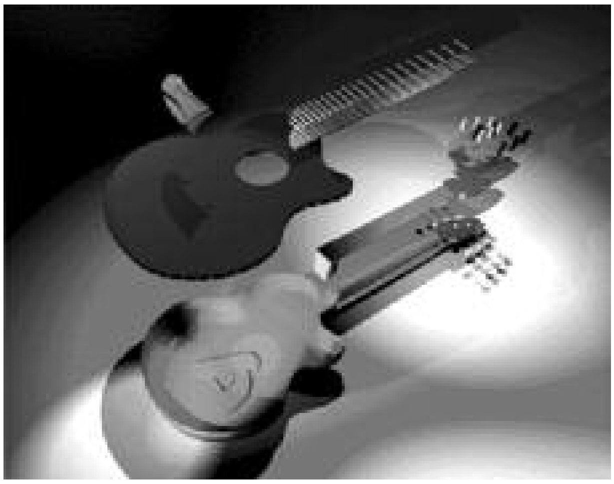


Figure 2 Designing-in quality

I had always agreed with the view that design academics in higher education should endeavour to sustain and develop their practice, which was one of the motivations behind my involvement in the 'polymer guitar' project which I pursued with Dr Owain Pedgley. It was really a research project relating to design innovation and one of the outcomes was a 'Frankfurt Show Prototype' embodying the capability to design in quality as indicated in Fig.2. We were aiming to improve many aspects of guitar performance – including playability, ergonomics, form and cost – and partly influenced by the wasteful production of toy 'plastic' (polymer) guitars that could not be played. These appeared to us to be largely a waste of materials, time, design, manufacture and distribution effort, although some children no doubt got some fun out of them for a short period at least. Anyway, although well-received at Frankfurt, we could not get the business arrangements in place to start manufacture. Pilgrim No.2 was my record of this project. Rob Armstrong had bought up the remaining guitar parts from the Pilgrim factory when it closed in the late 1980s, and he made this instrument for me with a foamed polycarbonate soundboard.

Pilgrim No.2 is also a waterproof guitar that I discovered I needed because after becoming the musician for Charnwood Clog Dancers, I found out that their dancing made it rain! I had previously been using a 30th Anniversary (all wood) Rob Armstrong guitar. It was made from wood that Rob had collected and been seasoning for decades, and I was playing it to looks of horror from informed guitarists as I tried to shield it under an umbrella. I continue to use 'the waterproof Pilgrim' to play for clog dancing and most people do not even notice that the soundboard is polymer. It sounds just like a guitar (as would be expected) and a very good one at that (as Rob Armstrong made it).

I bought Pilgrim No.3 when they became available early in 2010, and largely out of curiosity. John Hornby Skewes Ltd (JHS, a UK-based musical instrument company) had taken the Pilgrim guitar to the Far East where it had been copied. JHS also took an instrument that Rob Armstrong designed in 1980 to be copied as the basis for a Gordon Giltrap Signature guitar, and these formed the start of their award-winning 'Vintage Series'. I have a few issues with Pilgrim No.3, largely surrounding the neck and fingerboard, but otherwise it is a fine instrument. It has a loud, strong sound, which is less subtle than my original Pilgrim, but it is great for playing outdoors and, particularly for Morris tunes. However beyond the technicalities of guitar-making lies a much stronger message.

Figure 3 Pilgrim – Rob Armstrong waterproof hybrid

Figure 4 JHS Vintage 'Far East' Pilgrim

In the mid-1990s when we started the 'Polymer Guitar Project' at Loughborough, the reputation of guitars made in the Far East was for being 'cheap',' OK for beginners' etc. A decade later they were mass producing well-crafted Pilgrim guitars at relatively low cost. Manufacturing in the Far East was now offering quality as well as cost advantages. By the time I retired from Loughborough Design School in 2010, students were building CAD files for their final year projects and then deciding whether to send them to CAM machines on campus or companies in the Far East, and considering which would be fastest! On a more negative note, an American guitar company was reporting difficulties in obtaining supplies because of the emerging world-wide shortage of tonewoods. So the global economic model with the Far East as the manufacturing powerhouse is already beginning to show the 'sustainability cracks' that we had hoped the Polymer Guitar Project might help to alleviate, but, of course, that would only have been a 'drop in the ocean' in combatting the environmental damage that mass production and global trade are creating.

It is certainly a challenging task for future design and technology educators to retain relevance in such a complex, global context, but there is clearly still scope to believe that design innovation has a strong future in the UK. Perhaps the greatest challenge of all is to recognise and understand those aspects of 'craft', 'design' and 'technology' education that enable sustainable design innovation to happen. If the Inaugural lecture had taken place then making some contributions relating to these matters founded on my 28 years' experience at Loughborough Design School would have been its agenda, so it's probably fortunate that I retired when I did. There are some tricky questions here.

Errata

With regards to the research papers 'Technical Objects Between Categorisation and Learning: An exploratory case study in French middle school' page 32 and 'Phenomenology for Introductory Architectural Analysis Courses: The pentagon methodological approach' page 58, published in Design and Technology Education: An International Journal 20.2, the full list of authors/co-authors was not published. We apologise for this omission and the full list of authors is as follows:

Technical Objects Between Categorisation and Learning: An exploratory case study in French middle school

Impedovo, M.A; Andreucci, C.; Delserieys-Pedregosa A.; Coiffard, C.; Ginestié, J.
Aix-Marseille Université, ENS Lyon, ADEF EA4671

Phenomenology for Introductory Architectural Analysis Courses: The pentagon methodological approach

Fátima Pombo, Guest Professor, Department of Architecture, University of Leuven, Belgium
Wouter Bervoets, Postdoctoral researcher, Department of Architecture, University of Leuven, Belgium
Henk De Smet, Professor, Department of Architecture, University of Leuven, Belgium

Tasks in Technology: An analysis of their purposes and effects

Prof Richard Kimbell, Goldsmiths, University of London

Introduction for the 2015 DATA Special Edition

This paper was written in 1993 and was published in *The International Journal of Technology and Design Education.* 1994 Springer 4(3): 241-256. The paper concerns the nature of the tasks that initiate and drive technological activity. It is set in the context of two research projects that we conducted in TERU; the *Assessment of Performance Unit* project in Design & Technology (1985 to 1991) and the Economic and Social Research Council project "*Understanding Technological Approaches*" (UTA) (1992-1994). The former was a large scale national survey of performance in schools - involving tests on 10,000 learners in 700 schools, and the latter is a small scale study (80 learners in 20 schools) examining in detail the processes that learners engage in as they tackle technological tasks. However, the wider context of this paper concerns the English and Welsh National Curriculum (NC) implementation programme that had been launched in 1990. It caused a huge storm both in the curriculum generally and in design & technology (d&t) in particular. In the wider curriculum the assessment arrangements surrounding the Standard Assessment Tasks had been so badly designed that in 1992 teachers and schools had boycotted the whole process. And in d&t, the 'Order' that defined what teachers should do in the classroom/studio/workshop appeared to make very different demands on teachers than had previously been the case. The Order defined d&t in four 'Attainment Targets', the first of which (AT1) was 'Identifying Needs and Opportunities'. This (at least) implied that learners themselves should be doing that 'identifying', and in 1990 that was far from common practice. At exactly this moment we undertook the ESRC: UTA project that enabled us to collect the data that would inform this issue. We followed in detail the tasks that teachers set or negotiated with learners and examined the consequences of these tasks on ' subsequent actions.

I shall focus on two aspects of tasks that are central to understanding how learners respond to them and what they learn in the process. The first of these concerns the end-user; to what extent and in what ways the concept of a client impacts on the tasks that learners pursue in schools? Whilst this client issue had recently been brought to prominence in the NC, the second issue is much more deep rooted in the traditions of design & technology teaching. It concerns the extent to which the teacher controls what goes on - setting the task and controlling events - as against the learner taking responsibility for these matters. How much autonomy do learners have in setting and running projects, and what are the consequences of these levels of autonomy?

Technology and 'clients'

Technology is a task-centred, goal-directed activity. It is a multi-faceted and somewhat amorphous activity rather than a distinct discipline[1]; quite different for example from science or philosophy which have distinct boundaries. Technology makes use of a wide range of bodies of knowledge and skill, but is not defined by them, for the raison d'etre of technology is to create purposeful change in the made world. Something did not exist before, but now - as a result of human design & development - it does exist. We have wheelbarrows, wallpaper, waistcoats and warships because someone (or group) decided (for one reason or another) that they would be good things to have. This is technology. But technology is not just about new things. I constantly try to make my latest model of wheelbarrow (or warship) better than yours. This too is technology. It is a highly focussed activity and it is intensely value laden as should be clear from the use of the word "better". I might mean cheaper, or stronger, or longer lasting, or shorter lasting, or less damaging to the environment, or more damaging. All these are perfectly proper objectives that might make my ...whatever... better than yours for the purposes I have in mind.

So technology is a very human activity and is arguably one of the major distinguishing features of humankind. As Bronowski put it....

> Among the multitude of animals which scamper, fly, burrow, and swim around us, man is the only one who is not locked into his environment. His imagination, his reason, his emotional subtlety and toughness, make it possible for him not to accept the environment but to change it. And that series of inventions, by which man from age to age has remade his environment... I call... *The Ascent of Man*.
>
> (Bronowski, 1973, p.19)

Technological activity is driven by human desires - for comfort, for power, for money, for convenience, for identity. Technology cannot be blamed or praised for anything, for in itself it is entirely neutral. Blame and praise can only be attached to those of us who identify the objectives and who do the designing and developing of new and ever 'better' things. The boundaries of technology are *not* set by

1. Peter Medway (1992) provides an illuminative analysis of the multidimensional nature of the activity.

RESEARCH

our current practices and understandings in electronics or biochemistry or any other existing field. The boundaries are defined by our human desires. This is not to say that developments are always led by such desires, for there are many examples of manufacturers and marketing experts creating and massaging our desires. But the fact remains that any given technological outcome only exists when there is an identifiable client-based need for it. It matters not whether this need/desire is for Sidewinder missiles (very few clients but very wealthy ones - hence sufficient development and production money) or for cups and saucers (very many clients - hence a big market creating sufficient development and production money). In either case the fact remains that technology is client-driven.

What then of technology in schools? There is clearly a bit of a problem here as the people doing the technology are the learners and in the 'real' world they would be servicing the needs of their clients. But being in school means they are part of a teaching and learning programme that is controlled by the teacher. So who is in charge? Surely, either the *learner* is in charge of the activity, responding to the needs of a client, or the *teacher* is in charge, directing the learner into areas that s/he judges will be useful for the learner to experience.

It is clearly a much more complex issue to talk in terms of a client for learners' designing, and the notion was thrown into high relief by the publication in 1989 of the NC documents. Even from the very first of them in The Interim Report of the Design & Technology Working Group (DES/WO, 1988) it became clear that we were being encouraged to locate learners' project work in reality; or rather "in context". These contexts were many and various, the list in the document including the obvious ones of "home", "school", and "business & industry".

This was not in itself particularly far-reaching, for most technology teachers most of the time would expect to locate their learners' activities into some real or contextual framework. There is not only ample evidence that learner performance is far more effective when the tasks on which they are to engage are seen within a wider contextual framework (see e.g. Kimbell et al 1991), but also that learner performance can only really be understood in terms of that context (see e.g. Light & Perret-Clement, 1991). So the implied demand in NC technology for contextualised tasks was neither far-reaching nor particularly threatening for teachers. But far more significant - and infinitely more threatening - was the drafting of the 1st Attainment Target (AT); "Identifying Needs and Opportunities".

> ...learners should be able to identify and state clearly needs and opportunities for design and technological activities.
>
> (DES/WO, 1989)

Shock! Horror! Were learners really being expected to identify *their own* starting points for designing; identify *their own* client with an individual need that might be met? And if so what is the teacher supposed to do other than preside frenetically over the chaos (anarchy?) of a studio/workshop in which every learner is doing something different for their own clients? How, in this situation, would teachers ever manage to construct a teaching programme that showed any kind of progression? Surely structured teaching requires the *teacher* to be able to control the agenda; introducing certain things at certain times. If learners are busily setting *their own* agendas (in response to the imperative in AT1) - to what extent can teachers be said to be teaching?

The issue of whether or not a client is central to the activity has been supplanted by a different and more threatening issue. Who is in charge, the teacher or the learner?

Learner autonomy (learning to be self-directed)
One of the more obvious objects of schooling is to develop the ability of learners to manage themselves; to bring them to the point where they not only understand what it means to take responsibility for their actions, but moreover they have expertise in so doing. Developing learners' personal autonomy would rightly be claimed by any teacher as a central goal for education.

Some school activities lend themselves well to supporting this goal, and other less so. But it is not unusual to find school prospectuses identifying extra-curricular activities as a major area in which this goal of personal responsibility is brought home to learners and is thereby developed. The sporting ethic, the Duke of Edinburgh awards scheme, choirs and plays, neighbourhood support systems and the like all provide opportunities to underline and develop learners personal responsibility within a wider group framework. There is typically rather less emphasis on this in curricular activities - for there is simply less elbow room within which to do it.

But some curricular activities do lend themselves to it - and technology is one of them. In technology we do not need to feel entirely hamstrung (as are our science and maths colleagues) by vast lists of content to be taught, and for many years the basic mode of teaching and learning has been built around "the project". We operate in a

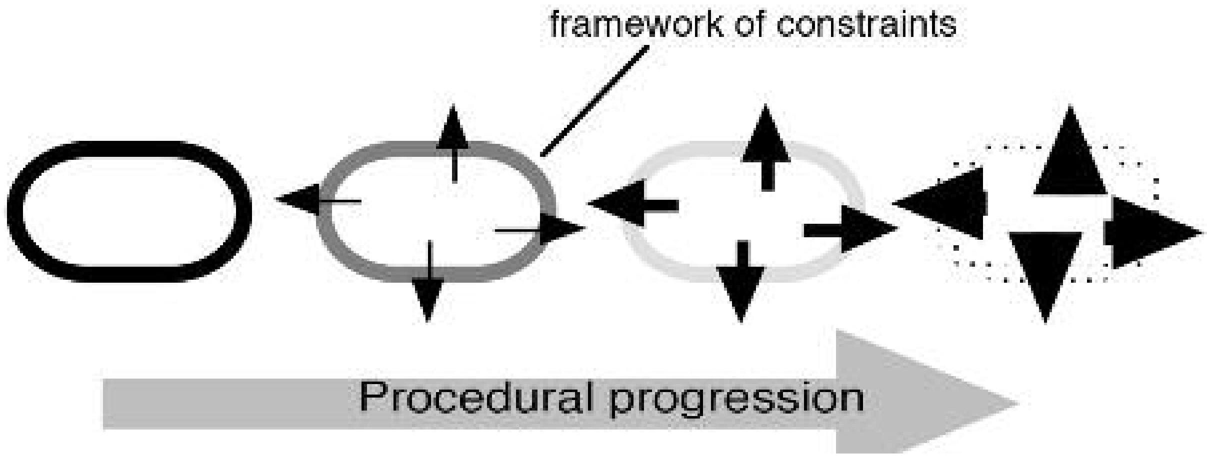


Fig 1. The teacher provides a progressively looser framework of constraints on project-work

studio-workshop environment on projects that typically run over an extended period, and this is an environment and a structure that lends itself nicely to developing autonomous decision-making by learners.

Within this environment, learners need to be introduced to the magnificent breadth of what is possible with materials, tools and a progressively more bewildering array of technologies. But at the same time, we have an ideal setting within which to develop their personal decision-making and responsibility. I have long held the view that technology teachers are almost uniquely fortunate in operating within this rich framework.

> ...the child will move in small steps from almost total dependence on the teacher to almost total independence....The function of the teacher...is to to steer children towards the goal of independent thought and action along the tortuous path of guided or supported freedom.
>
> (Kimbell, 1982, p.16)

From its earliest days in the late 1960s, when Design and/or Technology was first written about as a serious curriculum activity this feature of personal decision-making has been central.

> "Individuals are expected, as they mature, to solve problems on their own and to make decisions wisely on the basis of their own thinking. Further, this independent problem solving is regarded as one indication of the individual's adjustment. It is recognised that unless the individual can do his own problem solving he cannot maintain his integrity as an independent personality."
>
> (Schools' Council, 1975, p30)

"The project" became the standard modus operandi for teachers, and the project would enshrine a subtle balance between the things the teacher wanted to teach and the scope for learners to make decisions for themselves. For example, in a "room label" project, learners might each identify a specific room in the school and design a logo/label to describe what goes on therein. These designs might then get translated into moulds for vacuum forming and the finished plastic mouldings subsequently fixed to the various doors. The teacher would have designed the project specifically to teach the disciplines of vacuum forming, so if a learner produced a design that did not lend itself to this technique, the teacher would negotiate with the learner - manipulating it to the point at which it could be made to work as a vacuum forming.

Through this approach - allowing some freedom within a controlled framework - teachers built their whole teaching course. Introducing metal casting in this project, or electronics components in that one, dyeing fabrics here and automating a pneumatic system there. But technical content was only part of the progression in projects, for there was also an explicit and progressive pathway towards procedural autonomy. Projects would be expected gradually to place ever-greater responsibility on the learner and accordingly the teacher's framework for introducing the content would be ever looser. Early projects would be tightly constrained and would allow little deviation from the parameters set by the teacher. But gradually these constraints would become negotiable and permeable to the point where GCSE projects would be only very loosely controlled by the teacher and A level projects would be almost entirely at the discretion of the learner, involving only tutorial dialogue with the teacher.

Tasks in Technology: An analysis of their purposes and effects

In the hands of a good teacher, "the project" became an infinitely flexible teaching and learning tool. It built technical expertise and procedural autonomy and inevitably therefore produced some outstanding work. But in 1989, when NC AT1 ("identifying needs and opportunities") hit the classroom, this structure for project planning was thrown into confusion. The reason for this confusion lay in a deadly combination of the two issues discussed above, for the two central planks of NC AT1 were that projects would be seen to derive from real "needs and opportunities" of end-users (clients) and that the learners should be the ones to identify these needs and opportunities.

RESEARCH

The two issues merge

The words in the Technology Order appeared to weld these two sets of issues together into a formulation that placed far more responsibility on the learner than would formerly have been expected.

- Ask questions which assist them to identify needs and opportunities for d&t activities in familiar contexts.
- Recognise in their identification of needs and opportunities for d&t activities that the likes and dislikes of users is important
- learners should develop activities which offer opportunities for open-ended research leading to the identification of their own task...

(DES/WO, 1989)

For teachers who were used to the subtle exercise of control through the restrictions they built into design tasks, this was a serious body-blow. What were they being expected to do?

Some very unfortunate activities resulted from the confusions that followed the publication of the first NC Order for d&t. Teachers inevitably drew on their only experience of learner initiated project work - which they had formerly reserved for much older learners at GCSE or even A level. They selected or created contexts in which learners were encouraged to find needs and opportunities for themselves. "The shopping centre"; "the play group"; "the high street" all became targets for hordes of youngsters on the look-out for "needs and opportunities". In some of the more extreme cases these learners ended up designing a road-crossing system, or a youth club or an advertising campaign. No-one can deny that these are genuine design tasks, with identifiable clients and valuable outcomes. But they can so easily be utterly unmanageable and inappropriate as teaching and learning experiences. Inevitably, many young learners found it very difficult to operate in such an unfocussed way and ended up getting lost in the multiple demands of such projects. The teachers felt that they had to allow it to happen - the NC Order appeared to require it - but their instincts told them it was wrong.

It is now a matter of record that things were changed. The Order was re-written (several times) and teachers were exhorted to reassert their control of task setting to focus learner activity more tightly and to worry much less about the wider contextual and client-based setting for it. Four years after the original (radical) publication of the d&t Order, we reverted to a document that would have been readily recognised had it been written six years before.

Research data illuminates the issues

In the early years of the national curriculum there was much debate about how these two issues should be reconciled into teaching and learning programmes for technology, a good deal of heat has been generated - and far too little light. And it was with this in mind that we decided that our new ESRC project "*Understanding Technological Approaches*" should deliberately collect data that would enable us to describe and explain the consequences of the current position on learner performance in the classroom.

The approach taken by the project was broadly to observe learners throughout entire projects - registering data of particular kinds for every five minute interval. Some of these projects were quite short; around 120 minutes, whilst some run for up to 1300 minutes. The projects span all four Key Stages (5-16 years) and in total we observed 80 projects in 20 schools. The data we collected informed a whole range of performance related issues, including *engagement with the task* (speed and intensity of work), *interaction* (with teachers and amongst learners), *direction of work* (what priorities are followed), learner *intentions* (that steer their work), and the *manifestations* of these intentions in terms of studio/workshop behaviour. Given this breadth and detail of data, and given that it was collected every five minutes throughout the projects, this represented an enormous database of 'real-time' learner performance on tasks in schools. And sections of these data illuminate directly the two issues that I have outlined above;

- concerning the 'ownership' of the task in terms of who (teacher or learner) is in control
- concerning the wider world of clients or users and their 'needs'.

Data to inform the locus of control

Among the many observations built into the observer schedule is one that registered the points at which the teacher is *directing* the learner to do something in particular or is *supporting* the learner when they are trying to do something of their own choosing. This provided us with a crude but simple way of representing the axis of control in a project. It is important to remember that we were making no judgments about the value or wisdom of this direction or support - we merely note that it is happening.

Theoretically, the teacher might be directing or supporting in 100% of the 5 minute time slots but in reality this never happens. The following twelve project examples are taken directly from the data and show two things. First they demonstrate how the balance of direction and support indicates who is driving the project (project 2 for example having 4 times as much direction as support). But also the data provides a measure of the 'lightness of touch' of the teacher. Some projects (e.g. no 3) show the learner receiving either direction or support from the teacher in about 50% of the 5 minute slots throughout the project. Others (e.g. in project 7) show the total amounting to only 20%. In this case the teacher is allowing the learner to get along on his/her own for much longer periods.

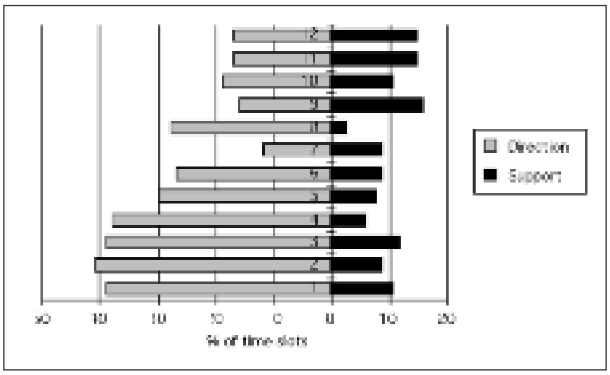


Fig 2. Percentage of project time when teachers are 'directing' or 'supporting' learners

Even this however is a serious oversimplification of the position, for in reality the balance of direction and support is not constant through a project. If the data is plotted against time we can see how this balance varies through a project. For this purpose we have divided the projects into 5 phases, each representing 20% of the time of the project. The following chart shows this balance in a single project but spread over the five phases of activity.

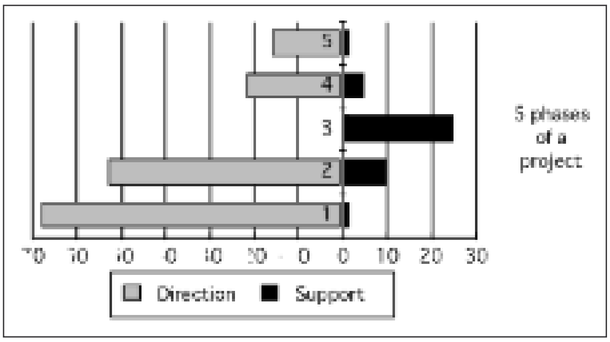


Fig 3. Percentage of 'direction' and 'support' through the phases of a project

In the first phase of the project there was a very high level of direction - with minimal individual support, but gradually as the project gets up-and-running the teacher backs off and in phase 3 spends all her energy supporting individually. The end of the project then reveals further steering by the teacher. The average figures over the life of this project are 30% direction and 8% support (it is project No 5 in the chart in Fig 2).

A quite different pattern emerges from another school (project 9 in the chart in Fig 2). Here the teacher spends as much time supporting the individual as in directing activities - even at the outset of the project. Interestingly, the greatest amount of direction arises in the heart of the project where most of the making will be going on, suggesting a degree of technical instruction of skills/procedures.

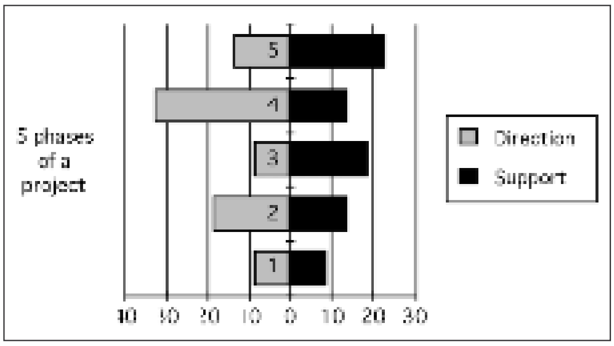


Fig 4. A teacher with 'light touch'

The average figures for this project are very different; 17% direction and 16% support. This not only reflects a more even balance of direction and support, but also indicates a "hands-off" approach by the teacher with the learner working independently (without either direction or support) for significant chunks of time.

It is one thing to describe these effects, and quite another to interpret them and we combined the data to see - for example - whether the differences of approach are associated with differences of outcome in the assessment and evaluation data. It is interesting to note for example that project 5 (30% direction and 8% support) is a secondary school project whereas project 9 (17% direction and 16% support) is a primary school project. Indeed we have been struck by the consistency of this trend in the data. When we plot these data from the whole data-set and organise it according to years (y1- y10) a fascinating picture emerges. We find individual 'support' outweighing 'direction' in primary schools, and the reverse in secondary schools. Moreover the transition between y6 and y7, as learners move from primary to secondary schools, is particularly stark.

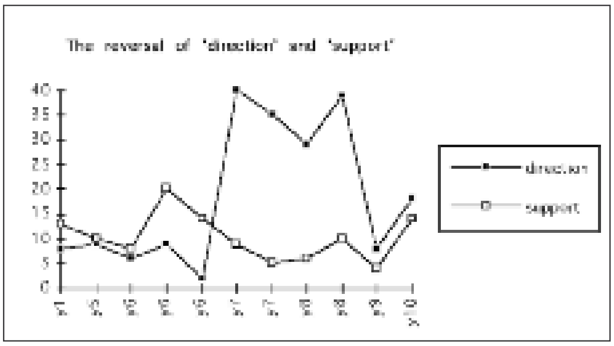


Fig 5. 'Direction' and 'support' data by year group reveals sharp discontinuity between y6 and y7

In y6, it appears to be the norm for teachers to spend much more of their time supporting individually than instructing or directing. In this setting, learners appear accustomed to taking significant responsibility for managing their work, using the teacher to advise and support when problems arise or advice is needed. In y7 the contrast could hardly be greater, with between 30% & 40% of the five-minute slots registering an instruction or direction, either to the individual or to the class as a whole. This is a totally different way of working, and one that must come as something of a shock to learners used to a very different approach to teaching and learning.

In terms of the growth towards personal autonomy, this y6-y7 boundary appears to represent a major step backwards. From a condition of relative independence and responsibility in y6, the learners have reverted to a frightening level of dependency on the teacher. They wait to be told what to do - even when they know perfectly well (and are prepared to tell you) what they might sensibly do next. They seldom do it, preferring to join a queue of other similarly timid souls waiting to ask teacher what they ought to do.

One thing that emerges very clearly from these data, is the extent to which technology projects in y7-y9 in our sample of secondary schools are heavily teacher directed. The HMI report on the first year (1990-91) of implementation of the NC (DES, 1992) criticised the work in y7-y9 in "some schools" where "...learners often spent much unproductive time trying to identify needs". Our project has been observing projects since 1992 and we have seen no evidence of this. Indeed we have observed quite the reverse - and the data outlined above suggests that y7-y9 teachers currently see their role in very different terms to that implied by the HMI criticisms of 1990-91.

Data to inform the role of the user/client

What then of the other major issue outlined above - concerning the role of the outside world and the "client" or "user". As we saw earlier, there is a good prima facie case for suggesting that in order for us even to call the activity "technological", the user's role must be clear. If there is no purpose to a project beyond teaching a skill or internalising a piece of knowledge, then the activity would more appropriately be called craft or science or history (depending upon what kind of knowledge/skill is involved). If learners are genuinely to be designing and making in technological terms, then they are design and making *something for somebody* - even if it is only for themselves or their mum. The user therefore ought presumably to make a significant contribution to the exercise.

In order to explore this dimension through our data, we used a measure that distinguishes between times when the learner is dealing with *user/task* issues, and when they are dealing with *manufacturing* issues. "User issues" would be registered when the learner is considering e.g. *how* big it should be or what shape it might be for people to hold it comfortably (whatever 'it' is). "Manufacturing issues" would be logged when the learner was working out how to manufacture it - or actually doing the manufacturing. We would therefore expect manufacturing issues to outweigh user issues if only because a considerable amount of time on a project is typically spent in 'making'. But in terms of the user/client issue, this approach allows us not only to quantify the extent to which learners are dealing with it - but more interestingly it allowed us to register how this concern changed through the life of a project.

The data shown below are from 47 projects in 11 schools and two matching trends are clear when they are plotted against year groups. Concern with manufacturing issues rises to a peak in y6,y7 and y8, and falls back towards y10. By contrast the 'user' data starts high in y5, drops to a low in y7 and rises back to y10.

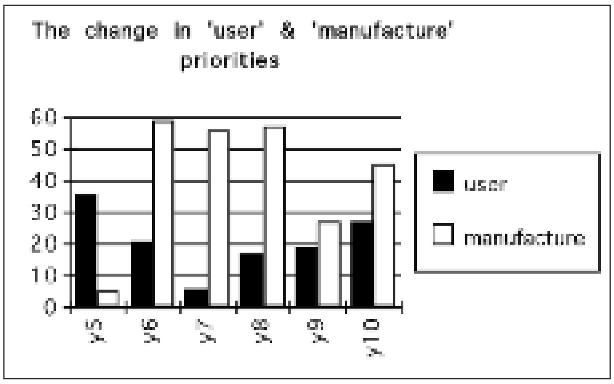


Fig 6. Learners designing for 'users' and for 'manufacture' y5-y10

A somewhat clearer picture emerges if we merge the data within year groups, and the trend in the "user" figures is very clear whilst that for "manufacture" is somewhat less so. In y7 projects in particular there is scant regard to any user and in y6, y7,and y8, there is a dominant concern with manufacturing issues. A more balanced picture then re-emerges in y9 and y10.

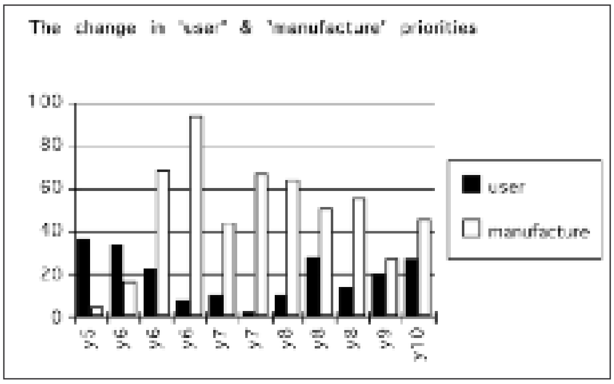


Fig 7. The dominance of 'manufacturing' issues in y 6/7/8, and especially y7

However these are average figures for year groups, and - as before - when these averages are spread across the five phases in the life of single projects we see a very interesting pattern. We can observe for instance the reconciliation of the user/manufacturer balance in the following two patterns taken from a project in y7 and a project y10 respectively.

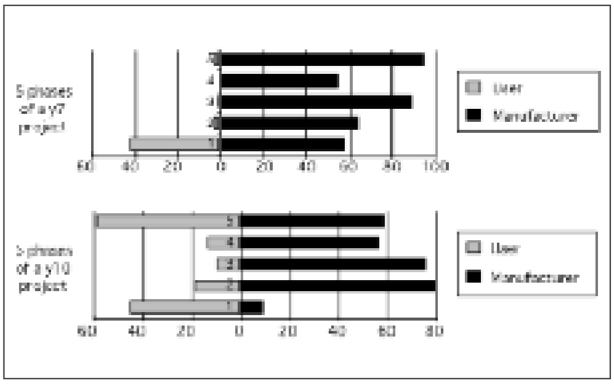


Fig 8. User/Manufacturing data across the 5 phases of the projects

The y10 pattern of 'user' concern is as one might conventionally expect. It is high at the outset of the project (phase 1) when the task is being clarified and detailed, and towards the end of the project (phase 5) when the performance of the product is being evaluated. In between these peaks, manufacturing issues dominate the learner's activity. But the y7 project however reveals a quite different pattern. At the outset (phase 1) there is significant user concern - but this disappears almost totally thereafter, with manufacturing concerns completely swamping all else. These data suggest that whilst the y7 learner did not see the user as significant or relevant to their activity, the y10 learner was significantly influenced by this factor.

Conclusions from the data as a whole, and clues for the future

Taken as a whole, these data suggest that technology projects are seen as very different things in the four key stages from age 5-16. When we combine the observation data outlined above with the more discursive and interpretive data derived from conversations with teachers and learners, the different characters of technology across the Key Stages begins to emerge.

Cultural technology ..."its all around you and always has been" , is characteristic of projects for 5-7 year old learners. Projects tend to be topic-centred across the whole curriculum (e.g. the Saxons) and technological activity derives from within the topic, involving Saxon forts or transport systems.

Problem-solving technology ..."try it for yourself - can you make it work", is more commonly associated with 8-11 year olds. Projects often have a fixed starting point – e.g. a wood strip vehicle chasis - and the challenge is to make it travel as far/fast as possible. It is common here for projects to amalgamate technology with investigations under the auspices of science.

Disciplinary technology ..."you need to know about this", emerges sharply at the start of secondary school (12-14 year olds). Projects are contrived specifically to include a small range of skills/knowledge from the (still largely separate) disciplines on the timetable. Pendants (to teach metal fabrication & enamelling), alarms (to teach simple circuits and sensors), snack-bars (to teach ingredient mixes and processing).

Simulated technology ..."this is how real designers work", progressively emerges with 14-16 year olds. There is a move towards individual projects - identified by the learners themselves and therefore generally having some reality. Within these projects learners are expected to be rigorous in the application of an abstracted designerly process and the development of a portfolio that reflects it.

These contrasted models of how technology should be pursued explain why "users" are largely seen as irrelevant to 12 year old learners at the start of secondary school. It is difficult to take a personalised user too seriously when the whole point and focus of the activity is an instructional one common to all learners in the group. The situation is very different from 5-7 year olds work where the whole experience (e.g. of the Saxons) leads to some awareness of them as living in (and hence users of) castles or wagons. Similarly, with14-16 year olds - where the user re-emerges as significant - it is not infrequently the genuine needs of the user (e.g. best mate/grand-parent) that prompts the project. The four different models of technology also explain the contrasted pedagogies, with top juniors frequently trying to work things out and investigate things for themselves and new entrants to the secondary school learning to do (largely) as they are told.

Given these contrasted models of what technology is about, we should not be surprised that there is no universal interpretation of what a technological task is like. Tasks evolve to fit the picture that teachers have in their heads of what technology is. We can sensibly talk about a 5-7 year old task or a 12-14 year old task - but there is very little common ground between them that allows us to speak about technological tasks *in general*.

That is what we found from the data in the UTA study. The big issue of course is that having *observed and described* this progression of models of technological endeavour, it does not follow that they ought to exist. As the philosopher G E Moore (1903) first observed, you cannot argue from "what is" to "what ought to be". It does not follow that because these trends do exist - it is right that they *should* exist. They may well be completely wrong-headed.

The fact is that in 1993 technology as a curriculum activity from 5-16 was so new and so undeveloped that it would be little short of astonishing if classroom practice was anything approaching coherent across the years of schooling, especially given the very contrasted cultural practices of primary and secondary schools. The national curriculum acted as a provocation to get technology started in many primary schools and even in secondary schools it was a baby in the curriculum. The mid 1960s would be a generous estimate of its date of origin. The levels described in the national curriculum d&t 'Order' sought to lay out a progressive pathway towards capability throughout the compulsory years of schooling, but the pathway was derived not from painstaking observation of what is going on in classrooms so much as from an abstract rationalisation of what *ought* to be going on.

The abstract rationalisation was important, and so too was the detailed observation and analysis of what is currently going on – as exemplified in our UTA project. It is necessary to bring them together and raise the level of debate about what technology *ought* to be like as a whole. Should there be such different models of technology across the age groups? Should we not plan learning activities for one year group with rather more understanding of the qualities that lead up to it in earlier years and flow from it into subsequent ones?

Resolving this matter will of course require teachers to come together and learn to talk - in a common language - about capability in technology. Such a dialogue would allow the profession to develop a securely rooted model of progression towards this capability. Given this wider perspective, we could then profitably debate what tasks should be like and what demands they should make on learners.

References

Bronowski J, (1973) *The Ascent of Man.* British Broadcasting Corporation

Department of Education and Science and Welsh Office (1987) *Task Group on Assessment and Testing - A Report.* DES

Department of Education and Science and Welsh Office (1988) *Design & Technology Working Group - Interim Report.* DES

Department of Education and Science and Welsh Office (1989) *Design & Technology for ages 5-16.* HMSO

RESEARCH

Department of Education (1992) *Technology Key Stages 1, 2 and 3*. HMSO

Department for Education (1992) *Technology for Ages 5-16*. HMSO

Kimbell RA (1982) *Design Education; the Foundation Years*. Routledge Kegan Paul

Kimbell R Stables K Wheeler T Wozniak A Kelly V, (1991) *The Assessment of Performance in Design and Technology. The final report of the APU design & technology project 1985-1991*. School Examinations & Assessment Council/HMSO

Kimbell RA and Wheeler T (1991) *Negotiating Tasks in Design & Technology. School Examinations & Assessment Council?* HMSO

Kimbell RA (1992) Technology - A Problem of Success in *Times Educational Supplement*. Times Newspapers.

Light P & Perret Clement AN (1991) "Social Context effects in Learning and Testing" in *Learning to Think* Ed. Light P, Sheldon S, and Woodhead M.nOpen University Press

Medway P, (1992) "Constructions of Technology: Reflections on a new subject" in *Technological Literacy and the Curriculum* Ed. Benyon & Mackay Falmer Press

Moore G E, (1903) *Principia Ethica*. Cambridge Univ Press

National Curriculum Council (NCC) (1991) "Planning Design & Technology at KS3" in *Implementing Design & Technology at Key Stage 3.* NCC Inset Resources

Nuffield Design & Technology Initiative (1992) *Newsletter 2* Nuffield Chelsea Curriculum Trust and Longman Education

Schools Council (1975) *Education through Design and Craft*. Schools Council Design and Craft Education Project Edward Arnold

RESEARCH

Gender Differences in Technology Illuminated Through Test Performance (outcome) Data and 'Real-time', 'On-task' (process) Data

Prof Richard Kimbell, Goldsmiths, University of London

RESEARCH

Introduction for the 2015 DATA Special Edition

This paper was written in 1994 as an internal TERU paper – it has not previously been published. It draws from two research projects that gathered data on gender differences in performance in technology. As with the Tasks in Technology paper (also included in this Special Edition), the wider context was the early years of the National Curriculum and specifically concerning the Standard Assessment Tasks (SATs). We were aware of the sensitivity of the gender data, essentially that girls seriously outperformed boys and the concomitant concern that the tests themselves might contain implicit bias, so we undertook a systematic review of the data from our two TERU projects that could inform the matter. The first provided 'outcome' data from APU tests (15 year olds in 1988 – Kimbell et. al., 1991). The latter, derived from the Understanding Technological Approaches (UTA) project (Kimbell et. al., 1994) allowed us to crosscheck these data with 'process' data derived from classroom observations (across all school years from 1-11 in 1992/3 -). I focus on two specific aspects of gender performance that were highlighted in test findings:

- concerning 'active' and 'reflective' response modes to tasks;
- concerning design proposals in relation to 'users' and for 'manufacture'.

The test data suggested that whilst girls were better at the reflective aspects of performance, the active aspects were split between girls and boys - the boys outperforming where proposals for manufacture are involved. Our observation data has modified this outcome somewhat, confirming the girls' out-performance in the *reflective* domain as well as in their ability to make proposals for the user. More surprisingly however, they have also showed themselves - in our 'real-time' observation of classrooms - to be more prepared to get involved in making proposals for *manufacture* as well. Taken together, this represents a comprehensive out-performance of the boys by the girls. It was this finding - along with other related ones - that interested the BBC "Panorama" team in our research and which we subsequently contributed to their programme "The future is female" (BBC, 1994). We do however have to exercise some caution with these findings, since the observation data - whilst being extremely deep, rich data - only relates to a small sample of 80 pupils when compared to the immensely broad (10,000 pupils in 700 schools) test data. The two data sets are therefore best explored and interpreted together - and the alignment of the assessment/observation frameworks makes this possible.

The data being examined

This paper represents an attempt to reconcile two quite different sets of data. The first kind (from APU) is *assessment* data derived from looking at the results of pupils' performance on a series of technology tests administered to approx 10,000 pupils in 700 schools through England, Wales and Northern Ireland. It is "outcome" data. We designed tests; we sent them to schools; the pupils worked them and returned their work to us; and only then could we examine what they had done and how well they had done it. By contrast, the second data set (from UTA) is based on 'real-time' observation of pupils working in studio/workshop/ classrooms on projects that were either set by the teacher or devised by pupils. This is 'case study' data of live school-based technology projects. There are 80 case study projects set in 20 schools, which were observed over a period of 18 months between Sept 1992 and March 1994. These cases studies therefore provide "*process*" data through the life of projects.

The importance of reconciling these data

There is a very good reason why it is important to attempt to reconcile these two data sets. Technology is generally recognised as being a procedural activity. There is no such thing as a right and wrong answer to a design task - rather there are better or worse solutions and the principal aim of technology in the curriculum is to enhance pupils' procedural capability in taking a task through to an appropriate resolution. When we were beginning the process of developing tests for the APU project, we were therefore very conscious of the potential contradictions in what we were doing. We were seeking to assess a *procedural* capability through largely *outcome-based* testing. This difficulty led us to take a particular stance in the development of the tests - specifically seeking to generate tests that reflected the processes of design and development.

> We took the view that we should not focus on conceptual understanding for itself, or on the decontextualised display of any particular communication skill, but rather in the extent to which pupils *can* use their understandings and skills when they are tackling a real task. Capability in design and technology involves the active, *purposeful deployment* of understandings and skills - not just their passive demonstration. Isolated tests of knowledge and skills were therefore quite inappropriate and we had to look toward the development of test tasks that could give us a measure of active capability.

> Given this starting point, we developed the idea that tests might be constructed that provide a 'window' *through which we could observe the process in action* - with the size of the window being defined by the time available...we hoped that it would be possible to see (and assess) the central procedures of the activity as well as the extent to which they were resourced by conceptual understanding on one hand and expressive facility on the other. We thus derived the three principal dimensions of an assessment framework.
>
> (Kimbell, et. al., 1991, p22-23)

This approach led us to the development of four strands of short tests:

- Starting points
- Early ideas
- Developing solutions
- Evaluating outcomes[1]

Despite the fact therefore, that our APU tests might be described as outcome-based, nevertheless the outcomes provided evidence of the *processes* and *sequences* used by pupils in tackling their tasks. Accordingly, we were subsequently able to develop and substantiate hypotheses about the processes of design and development used by pupils in technology. [See Kimbell, et. al., 1991 sections 11-16, and Kimbell & Wheeler, 1991 (a) and (b)]

We were well aware at the time of writing these documents that we would ideally need to substantiate the matters raised in those reports through some genuine process-focussed observation of technology in action in classrooms. Accordingly, with the support of the Economic and Social Research Council (ESRC), we launched the UTA project to do just that. We adapted the APU assessment framework into an *observation* framework that allowed us to follow the work of individual pupils through the totality of their projects (however long they were) whilst registering many of the same elements of capability as had been in the APU framework. Moreover we decided that - in order to shed some light on the *development* of capability and hence to illuminate issues of progression - we should choose our case study sample from right across the age range (year 1 – year 11).[2]

In the following pages, I shall explore two principal sets of issues that emerged through the APU data and that can be examined more fully through the UTA data.

Issue (i) Two sides of capability (active and reflective)
The analysis of pupil responses in the APU survey led us to postulate the existence of two sides to capability; active and reflective. The latter is essentially the ability to see all the *issues* that need to be tackled in a task, whilst the former is the ability to respond to those issues in actively making, and developing, *proposals*.

In identifying these two sides of capability we were emphatically not recommending their separation. We argued that in design and technology such a separation would be damaging, and we developed a model of technology as *thought in action* rather than thought separate from action. Having said that however, it is important to recognise the relative strengths and weaknesses of pupils in relation to these two sides of capability, and the teachers with whom we worked found it a useful diagnostic device to begin to consider remedial (balancing) strategies in cases where pupils were demonstrating clear imbalance of the two in their response to tasks.

We identified several recurrent trends of such imbalance particularly in relation to pupil gender. (See Kimbell, et. al., 1991, section 15)

> Generally, girls do far better on the more reflective tests than boys, and boys do somewhat better than girls in the more active tests. In other words, girls appear to be better at identifying tasks, investigating and appraising ideas, whilst boys seem to be better at generating and developing ideas.
>
> (Kimbell, et. al., ibid)

In order to get a purchase on this issue in terms of the UTA *process* data, we needed to reinterpret these findings into a project-work mode. Such an interpretation might lead to the following assertions;

- Boys are more active and girls are more reflective in their response to tasks.

---

[1] These four test structures - which we originated in 1986, piloted on a large scale in 1987, and which formed the backbone of the national survey in 1988 - are uncannily close to the four attainment targets that subsequently appeared in the 1988 Interim Report of the national curriculum working group for design & technology.

[2] APU data is from 15 year old learners only.

Gender Differences in Technology Illuminated Through Test Performance (outcome) Data and 'Real-time', 'On-task' (process) Data

- Given that the start (sorting out the task) and the end (evaluating the outcome) of a project are typically more reflective, girls will be more comfortable at handling these starting and finishing phases of the task.
- Given that the middle of the project (making) is typically very active, boys will be more comfortable at handling this central phase.

If we examine the UTA data for 15-16 year olds (in England, Key Stage 4 - KS4), i.e. the data that is closest (in age terms) to the APU data, these assertions appear to match very closely with the evidence. KS4 is the point in English schools when pupils are preparing for external assessment through GCSE examinations and the projects observed were part of this preparation. The data showed that girls engage with the early (typically reflective) part of the task far more readily than do boys. Our index of engagement is on a three-point scale "motoring" (fully engaged) "poddling" (in tick-over mode) and "stationary" (effectively off task). In the KS4 projects we observed, the girls are motoring in the early reflective stages of the project, and the boys only begin to get on terms with them in the middle of the project in the more active making stages. This is shown in Chart 1, which indicates the level of engagement by boys and girls as they move through their projects. Each project was subdivided into five equal phases, indicating the engagement in the first 20% of the project, second 20%, and so on.

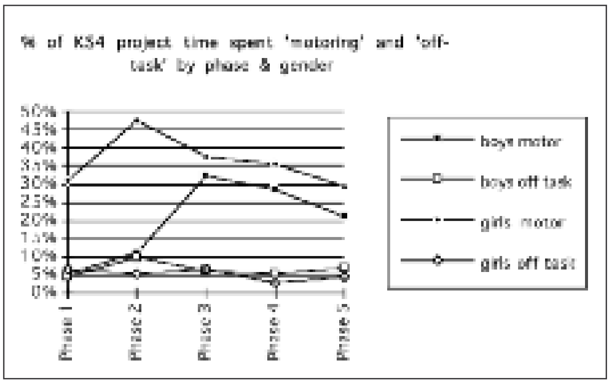


Chart 1 Differences in engagement across projects between girls and boys in KS4

It is a somewhat startling fact that, on average, the boys in our KS4 sample spent the first two-fifths of their project at a very low level of engagement (a mere 5-10% motoring) when compared to the girls (around 40% motoring). Given that GCSE projects range up to 50 hours of timetable time, this represents a prodigious waste of valuable time.

To lend further weight to this analysis of the critical early stages of the project, it is interesting to observe what *kinds of things* the boys and girls are doing - quite apart from the intensity with which they are doing them. In these early stages, the boys are much more likely to be doing (in "poddling" mode) a range of 'active' things (e.g. modelling) whereas girls are more likely to be doing "reflective" things (e.g. investigating/evaluating).

Taken together, the data on *engagement* with the task and the data on the substance of the activities being pursued suggests that our three assertions (above) are broadly true. But the analysis provides a fascinating illustration of the importance of 'real-time' observation data when trying to interpret performance in technology, for the most interesting feature of these data concerns the changes in performance *across the phases of the project.* Chart 2 shows the ways in which boys and girls engage differently in 'active' and 'reflective' modes.

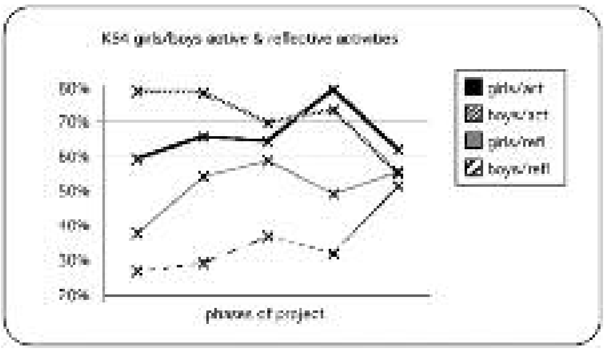


Chart 2. Differences in active and reflective modes of engagement between girls and boys in KS4

The boys start the project by being much more active than the girls and end up being less so. The boys start by being far less reflective than the girls and end up much closer. The boys' performance *starts off* with an enormous disparity between the active (79%) and the reflective (26%) modes of response and ends up much more balanced (61% and 51%). Girls' approach is more balanced throughout. By using these 'real-time' data, we can comprehensively confirm a significant pedagogic finding from the APU data - but which (at that time) we could only infer from performance on different tests.

> ...boys are more able to get to grips with reflective aspects of capability when they are practically engaged in developing a solution, and especially so when they are able to do this through more practical modelling activities. Girls on the other hand would appear to be more able...(to do this)... without the benefit of such practical engagement. (Kimbell, et. al., 1991, p.215)

Gender Differences in Technology Illuminated Through Test Performance (outcome) Data and 'Real-time', 'On-task' (process) Data

The boys' engagement in practical activity enables them progressively to gain access to reflective issues. The girls appear more likely to be able to hold a balance throughout the activity. One important question that flows from this is the extent to which this significant difference in the performance styles of the gender groups is reflected in earlier data, i.e. from KS3 (11-14 year olds) and KS2 (8-11 year olds). Our UTA data allows us to examine these same issues across this wider spectrum of schooling, and three initial differences about active/reflective responses at KS3 are obvious.

First there is far less difference between boys and girls than there is at KS4. Broadly, the curves follow each other closely, with reflective activities growing through the project (from 20% to 40%) whilst active activities decline (from 90% to 70%). Second, the actual levels of active/reflective activity are more extreme than they were at KS4. At KS4 the averages were 68% active, 37% reflective. At KS3 the averages are 78% active, 32% reflective. Third, the profile of performance (boys and girls) is far closer to the boys profile at KS4 than it is to the girls' profile. Boys and girls at KS3 respond very like the boys at KS4.

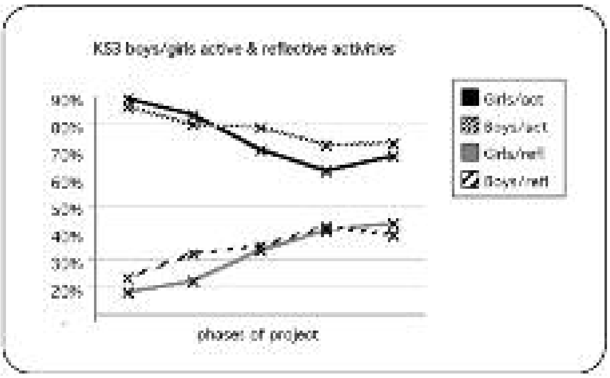


Chart 3 Differences in active and reflective modes of engagement between girls and boys in KS3

These trends illuminate further our earlier analysis of KS3 technology (see Kimbell, 1994; Stables, 1995). We had already characterised KS3 technology as being "disciplinary" technology in the senses that (a) it is more instructional than any other key stage and (b) that it is instructional of the *skills and knowledge of the material workshops* at the expense of design skills and experience. In the far more tightly teacher controlled environment of KS3 technology, it is not surprising that individual pupil differences are squeezed out and produce far more homogeneous data. Moreover the focus on skill-acquisition - at the expense of designing - creates the more extreme active/reflective imbalance of responses.

What then of the position at KS2? Might one expect performance to be more like that at KS3 or KS4?

The data indicates three important features about the performance of boys and girls at KS2. Firstly it is very similar; the boys and girls profiles are almost exactly matching. Secondly profiles are significantly different to those at KS3; there is a better active/reflective balance throughout the project. Thirdly the KS2 profiles (girls & boys) match more closely to the girls KS4 profile than to the boys KS4 profile.

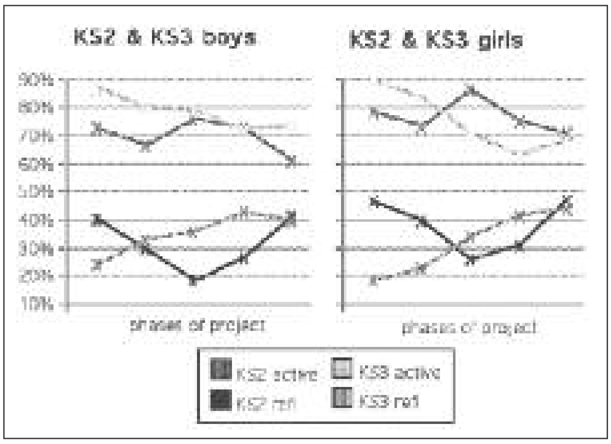


Chart 4 Differences in active and reflective modes of engagement between girls and boys in KS2

The conclusions that one might draw from this analysis of active/reflective response styles of girls and boys across Key Stages 2, 3 and 4, are as follows. It would appear that girls and boys performance at KS2 is very similar in style and relatively balanced in terms of active and reflective modes of response through the life of the project. At KS3, boys and girls profiles are still very similar, but are quite different to those at KS2. The profiles indicate that in the projects there is an early preoccupation with active modes of response at the expense of the reflective (more doing than thinking) and that this is gradually brought more into balance as the project proceeds. At KS4, the boys' performance looks very similar to the girls & boys KS3 pattern (starting from great imbalance and moving towards balance) while the girls performance is closer to that which girls & boys exhibited at KS2 (greater balance through the project). The boys appear to be more influenced by their KS3 experiences than the girls.

If it is true - as we suggested earlier - that at KS4 *"the boys engagement in practical activity enables them progressively to gain access to reflective issues"* then it is as much a comment on KS3 learning and teaching in technology as it is on the boys themselves. For at KS2, they were - equally with the girls - quite able to grapple with the reflective as well as the active throughout the task.

Issue (ii) Developing design proposals in relation to the User and to Manufacture

When pupils are making design proposals in response to a task there are two broadly distinguishable facets to be dealt with;

- developing proposals in terms to the users of the products/systems; (e.g. so it is comfortable to use and the right size)
- developing proposals in terms of the manufacturing constraints; (e.g. ensuring that it can be assembled easily and won't fall apart)

In the APU data, these two facets of the task threw up some interesting differences in the balance of concern of the gender groups.

> It would appear to be the case that girls are generally significantly more able at developing products in terms of the user, whilst boys are more able at actively considering the manufacturing dimension. Both the general trend and this gender difference are demonstrably present in test 3iA where girls - of all ability levels - outperform all boys in 'user' developments, whilst boys - of all ability levels - outperform all girls in the 'manufacturing' developments.
>
> (Kimbell, et. al., 1991, p.217)

The dangers of the short term testing of essentially long term procedural qualities appear to be highlighted by this finding which - at first sight - is *not* confirmed by our UTA data. These 'real-time' data suggest that girls are prepared to deal with user issues *and* manufacturing issues at equivalent levels to the boys, indeed often to higher levels. As evidence of this, the following chart highlights pupil performance in this area at KS3. It shows a clear advantage to the girls as the project takes its course.

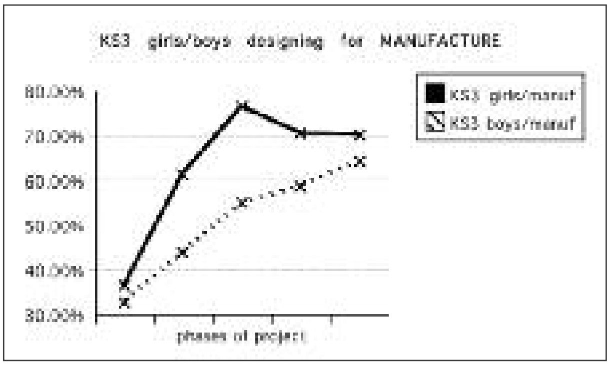


Chart 5 Differences between girls and boys focus on manufacture in KS3

At the outset of the project, neither the boys nor the girls take manufacturing issues too seriously, but these form a major concern from the mid point of the project onwards. Parallel (though not quite such extreme) results emerge at KS2 and KS4. How then are we to interpret this in the context of the APU data? Chart 6 illustrates the differences between girls at different ages. The first point to observe is the extent to which these data relate to the phases of the project, and moreover the phase pattern at each key stage creates another pattern. At KS2, girls concern with manufacturing issues varies only slightly across the project (42%-57%). But at KS3 the max-min span is significantly bigger (36%-77%) and at KS4 it is bigger still (20% - 77%).

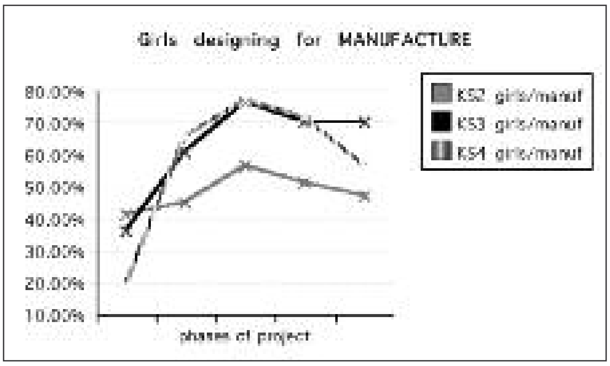


Chart 6 Differences between girls' focus on manufacture at KS2, KS3 & KS4

The girls in our UTA sample appear to be learning to concentrate their energies on particular things at particular times - and manufacturing concerns are increasingly seen as appropriate in the middle of the project and less appropriate at the start and towards the end. Progression across the key stages would appear to be characterised by increasing specialisation and focus and it is very difficult to accommodate this in short-term testing. Incidentally, an exactly reciprocal curve exists in their designing for the user, which starts at a high level - dips through the mid-point of the project - and rises again towards the end, as is shown in Chart 7.

There are two stages in the reconciliation of these long-term process-based findings with those from the short-term APU tests. First we need to recognise that, at least in part, we have exposed two kinds of limitation in the APU test results.

- the limitations of paper-based testing for measuring concrete (manufacturing) concerns.
- the limitations of using short-term measures of long-term capabilities.

Gender Differences in Technology Illuminated Through Test Performance (outcome) Data and 'Real-time', 'On-task' (process) Data

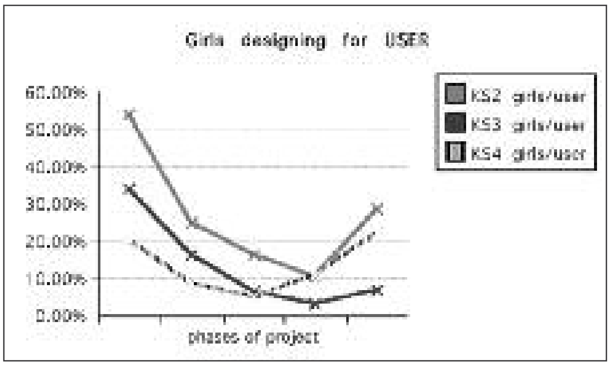


Chart 7 Differences between girls' focus on the user at KS2, KS3 & KS4

> We recognised these problems at the time, qualifying our findings in the following manner;
>
> …it may well be that...the manufacturing demands are very remote from the task...and do not typically arise until much later in the activity. It may be therefore that they (the pupils) are less able - or less prepared - to get involved in this manufacturing dimension in the early stages. (Kimbell, et. al. 1991, p.218)

However, this is not the whole story, and the second stage of the reconciliation lies in recognising that APU surveys are composed of largely random samples of pupils - whereas (at KS4 at least) our UTA sample was focused on a self-selecting group of pupils who have chosen to do technology as an examination subject. Again this is an issue that we noted at the time.

> Because of the emergent condition of design and technology among the schools, it would have been rash to rely solely on this randomly selected sample of pupils for testing.... Accordingly we decided on a policy of enriching the random sample with further 'target' samples drawn from courses of particular interest...our pupil samples were therefore composed - both for the pilot and the main surveys - of a blend of randomly selected pupils and pupils that we knew were pursuing certain courses.
>
> (Kimbell, et. al., 1991, p.41)

The obvious next step therefore was to see what performance levels were like in those 'target' samples that would be more akin to our UTA pupil sample. We found the *performance differences* between girls and boys in the target samples to be *significantly reduced*.

> …the general rule governing the performance of design and technology curriculum groups (as opposed to the control group) is to even out some of the gender imbalance. Whilst girls are generally stronger on user developments and boys on manufacturing developments, if we look at the girls in the design and technology curriculum courses there is a clear picture showing girls on these courses to be scoring more highly than the control group. Of particular interest is the inclusion within this of higher scoring for developing proposals for *manufacture* - in 70% of cases where this quality is assessed.
>
> (Kimbell, et. al., 1991, p.218)

These data are clearly far more compatible with the findings from our UTA sample, which show the girls matching, and even outperforming, the boys.

Conclusions

This paper grew from the realisation that technology - being a procedural activity - presents very real difficulties to anyone seeking to measure performance in short tests. Our APU experience persuaded us that it was possible to derive valid data on performance in this way - but we were always aware of the limitations of that data. The UTA project, whilst broadly confirming our findings, illuminated the limitations of short tests and the extent to which real-time observation of pupils on task can flesh out and enrich our APU performance measures.

For the purposes of this paper we chose to focus on gender issues in performance, and specifically on two APU findings;

- concerning 'active' and 'reflective' response modes to tasks
- concerning design proposals in relation to 'users' and for 'manufacture'

These two sets of issues have a structural relationship that spans the whole of capability in technology and that might be represented as shown in Figure 1

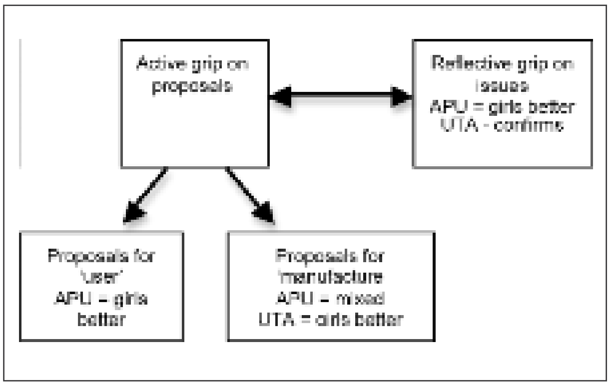


Figure 1 Gendered relationship between action and reflection, as shown by APU & UTA data

Gender Differences in Technology Illuminated Through Test Performance (outcome) Data and 'Real-time', 'On-task' (process) Data

Our APU data suggested that whilst girls were better at the reflective aspects of performance, the active aspects were split between girls and boys, the boys outperforming where proposals for manufacture are involved. Our UTA data has modified this outcome somewhat, confirming the girls' out-performance of boys in the *reflective* domain as well as in their ability to make proposals for the user. More surprisingly however, they have also shown themselves, in our 'real-time' observation of classrooms, to be more prepared to get involved in making proposals for *manufacture* as well. Taken together, this represents a comprehensive out-performance of the boys by the girls. It was this finding - along with other related ones - that we contributed to the BBC Panorama programme "The future is female" (BBC, 1994).

We do however have to exercise some caution in making this assertion, since our UTA data - whilst being extremely deep and rich data - does only relate to 80 pupils in 20 schools. Furthermore since we focused a majority of our sample into KS2 and KS3, we have only 3 schools and 12 pupils in our KS4 sample. Moreover those schools and pupils were not chosen to be (indeed they could never be) a representative sample, and we must therefore be careful not to assume that these findings are generalisable to all pupils in all schools.

This illustrates the value of our two contrasted data sets. Our APU data is immensely broad (10,000 pupils in 700 schools) but the performance data is restricted to test responses. Our UTA data is immensely deep 'real-time' data but it is insufficiently broad for generalisable conclusions to be drawn. This paper is the result of our first foray into the combined data where we consider that the value is in the way that the two sets of data have added insight to each other.

References

Assessment of Performance Unit (1981) *Understanding Design & Technology*, HMSO

British Broadcasting Corporation (1994) "The future is female", *Panorama*, 24th Oct 1994, BBC

Kimbell, R., Stables, K., Wheeler, T., Wozniak, A. and Kelly, V. (1991) *The Assessment of Performance in Design & Technology; The final report of the APU Design & Technology project*, SEAC/HMSO

Kimbell, R., & Wheeler, T., (1991a) *Negotiating Tasks in Design & Technology*, SEAC/HMSO

Kimbell, R., & Wheeler, T., (1991b) *Structuring Activities in Design & Technolog*y, SEAC/HMSO

Kimbell, R. (1994) "Tasks in Technology; an analysis of their nature and effects" *International Journal of Technology and Design Education*. Vol. 4 pp 241-256, Kluwer Academic Publishers

Kimbell, R., Stables, K. & Green, R., (1994) Understanding Technological Approaches: Final Project Report to the Economic and Social Research Council (ESRC)

Stables, K. (1995) "Discontinuity in Transition; pupils experience of technology in year 6 and year 7" *International Journal of Technology and Design Education* Vol. 5 pp. 157-169, Kluwer Academic Publishers

Researching Design Learning: Research methodology

Prof Richard Kimbell and Kay Stables, Goldsmiths, University of London

Introduction for the 2015 DATA Special Edition
This paper was originally published as a chapter that formed a methodological end-piece for 'Researching Design Learning' - the book that we published in 2007 with Springer publishers. In the chapter we draw together many of the priorities, approaches and 'rules of thumb' that we have developed over the years as our research experience gained momentum and that are exemplified through the research projects described in the book. Through this piece we explore the 'fit' between the values of researchers and their clients and how this has impacted in what we have researched and the ways we have gone about the activity. We explore the lessons learned, the critical role of research design and about the design of research instruments, interventions, data collection and data analysis approaches. We explore some of the approaches we have used for bringing research to life for ourselves and others, making it understandable, meaningful and immediate. We conclude by confirming our view that researching is a very design-like activity.

Springer publishers have kindly agreed to the re-publication of this slightly modified chapter as part of this Special Edition.

This chapter reflects on the ways in which we were emboldened to get into research in the first place and the ways in which we engaged in research as a designerly kind of activity in which we have felt free to exercise our creative talents. This does not mean that we have assumed a license to be dilettante, but rather that, given a specific research challenge, we developed all kinds of tools (sometimes very unusual ones) to give us some purchase on the issue in hand. Sometimes these tools have empowered us to *gather* data more effectively, sometimes to *organise* those data in new ways, sometimes to *analyse* data and sometimes in the *presentation* of data. We conclude with the point that researching is a very design-like activity.

Starting points and the challenge of values
Any research methods guide will underline the importance of getting a clear starting point, and we would absolutely agree with that. Teasing out the questions that one is trying to answer through the research is a necessary and sometimes complex process. The more precise the questions are, the easier it is to decide what will count as data to enable us to answer them. Part of the complexity in this process of elucidating research questions derives from the common occurrence that the clients/sponsors of research are unclear themselves about exactly what they want. It frequently takes a good deal of negotiating to dig out what they *really* want to know. The process is just the same as when a lay-person commissions a designer or architect or gardener to generate a new product/living space/garden. The lay-person will typically have some vague notions of what they want. They might have cut out pictures from magazines or (in rare cases) sketched for themselves what is in their head.

But it then remains the job of the designer/architect/gardener to bring their expertise to the task. This is 'what-if' time. What if it was like this? What if it did that? Would it be good if? Would you like it to do that? In doing this, the creator is not throwing *solutions* at the client, but is rather trying to tease out their response to see what excites or interests them. The process is all about digging out the *values* that the client is trying to embody in the work. Are we after a peaceful/tranquil garden space; or a formal architectural space; or a space of light and movement; or; or; or.

It is precisely the same with research clients. We offer up tentative solution-types to gauge reaction and thereby get a better grip on what is really wanted. Are they looking for a *statistic* that will convince a policy body or a collection of case study examples to *illuminate practice*? Or do they seek to shape that practice in particular ways? Not infrequently the client will say 'yes' 'yes' and 'yes'… we will have all of that. At which point it is our turn to point out that *everything* is not an option unless there is lots of time and money. So we help them to prioritise what they *really* want, and what might be a nice added extra. These underlying value debates then directly shape what we might do in the research both to aid our understanding of the issues in the data and to help our clients and stakeholders to get messages across.

But teasing out the clients' priorities is only part of the complexity of finding a starting point. For overlaying them are the priorities that we ourselves bring to the task. We are not just jobbing researchers looking to earn a crust by doing anyone's bidding. We have our own set of priorities

RESEARCH

– typically concerning designing and learning – that we are always interested to understand better. Since we are reasonably well known in research circles, most of the clients that approach us do so knowing that these are our concerns. It is therefore not difficult to find research questions that are appropriate for the client and of interest to us. But there have been some cataclysmic fallings-out over this matter, and the project in 1991/2 in which we developed the first round of KS3 Technology SATs (formal, externally set, Standard Assessment Tasks – SATs - for 14 year olds) provides an interesting case.

This was a hugely valuable project that we obviously wished to be a success. But this eventually proved impossible because of the conflict in values between what the client, the School Examinations and Assessment Council (SEAC), wanted and what we were prepared to do. We worked quite comfortably on the first round of development and produced a set of prototype SATs that seemed good to us (they produced the important data to inform learners' performance against the requirements for the National Curriculum for design and technology). But as we moved towards the second round of development, the terms of the brief were drawn far more starkly. SEAC really did want tests – with right answers – that could be marked with certainty against a checklist. We wanted test *activities* (like the ones we have previously created for the Assessment of Performance Unit (APU) ones (Kimbell, et. al., 1991), which could be assessed using teacher judgements.

We were not prepared to develop tests of the kind that SEAC demanded, since we judged the position of SEAC to be completely wrong at every level; for schools, for teachers, for learners and for design & technology more widely. So we did not do it and we were removed from the development process. And we lost a huge amount of money. This one case brought home to us very clearly that client values and researcher values have to be (at least somewhat) aligned before any research venture can succeed.

Research Design

Assuming that a clear set of questions has emerged from the negotiations establishing a project, the research design becomes a critical aspect, and moreover a part that offers great opportunities for creative thinking. From the priorities identified at the outset we have to create a *design* for the research that stands some chance of achieving the desired outcome. What are we going to do? How are we going to do it? Central to the answer to both these questions is another one: what will count as data?

Think yourself into our shoes at the outset of the *Decisions by Design* project for the Design Council (Kimbell et. al., 1997; Kimbell and Stables, 2007). We had an absolute alignment of their values and priorities with our own. They were interested (and so were we) to see how the lay-person's everyday decision-making process might be the same as, or different from designerly decision making. How might we do that?

The context of the project lay in schools (the Design Council's 'Total Schools Design' initiative) so it made sense to us to think about lay people in schools. Since we would need cooperation at a reasonably high level, it also made sense to target the school management team. We also wanted to have both primary and secondary schools involved. But how many? And from which schools? We already recognised that if we were to get at their decision-making processes, we would need some significant blocks of time working with them to allow them to develop sufficient trust in us. We also wanted to be able to sit around a table with them all at one time.

Using these thought processes we settled on the idea of six teacher fellows (three primary and three secondary) each selected from the school's management team, and committed to giving 12 days of their time to the project over a year. But that was only half the problem, for where would we get the contrasted designerly decision makers – and how would we get them together?

Goldsmiths has a flourishing PGCE programme of teacher education, and each year we take in a group of fresh young design graduates who have an interest in becoming teachers. So we had a captive audience of trained designers. Could use them?

In the end, we operated a double procedure. First – mostly in their own schools and in their own time – the teacher fellows were asked to draft a 'fly on the wall' description of what had happened in their school when an important decision gets made; e.g. about school development planning, budget making, timing for a new school day, or disciplinary procedures. We wanted a full account of how the decision came to be made, recording all the things that might have contributed to that specific decision-making process. The aim was to gain a comprehensive account of why and how the decision got made in the way that it did.

Then, through the subsequent term, the teacher fellows observed our PGCE designers at work on a group-based design project. Four sessions were dedicated to working with students who were asked to work as they would normally do in design activities. Each group had a teacher

fellow assigned to follow their development. Throughout the term the teacher fellows took on the roles of participant-observers in these design activities, and moreover they were required to reflect upon their experiences:
• Analysing the design techniques used
• Debating their strengths and limitations
• Reflecting on the transferability to other problems and settings.

In the end, the teacher fellows were astonishingly lucid about the differences between their own and the design students' decision-making processes. This research design – as with all research designs – was in part based on debates of principle. But at the same time it was also based in part on the pragmatics of what can be done in the time available and with the resource at our disposal.

We could have done something very different. We might have sent questionnaires to thousands of people (some designers and some not) inviting them to tell us about their decision-making processes. We might then have analysed the differences and (possibly) derived some statistically solid data. We judged however – as a point of principle – that we had to put our 'subjects' into decision-making mode and ask them to observe and reflect upon what happened. This is far more demanding and time consuming, but (in our judgement) far more likely to reveal the realities of decision-making. Having made that research design decision of principle, we then had to manage the pragmatic consequences of who, when, where and how. Perhaps we should note here that we have never – ever – used blanket questionnaire techniques. We believe that questionnaires can be useful when administered in person to get particular bits of information from people we have worked with and who understand what we are doing and why we are doing it. But our own response to 'blind' questionnaires through the post or on the High Street makes us vary wary indeed of attaching any significance at all to any resulting 'findings' from such blunderbuss techniques. We also recognise, however, that this instinct is informed by our basic philosophy of research, which is to lean more towards interpretive than positivist models.

The challenge of research design frequently rests on the trade-off we have illustrated here from *Decisions by Design*. What we would like to do in principle – set against what we have the resources (time/money/expertise) to bring to the task. The end result has to be convincing and worthwhile, but equally it has to be do-able.

Instrument design
It has frequently been the case that our projects have involved the development of new instruments for promoting learner performance or for collecting data of one kind or another. Once again it is our designer instincts that pop to the surface when faced with these challenges.

For our first project (APU) we developed some very different response booklets for learners to work through over a 90-minute design task. This and subsequent versions (e.g. for Assessing Design Innovation and for e-scape; Kimbell & Stables 2007) are cases of instrument design where the priority is to find ways of promoting design performance in a short time but without losing the integrity of real designing behaviour. In fact, in these cases, the booklets have to be seen alongside an administrator script and an assessment rubric. Together they comprise the 'instrument' and a huge amount of time, experimentation, trialling and modifying was involved in the original and subsequent versions.

But a very different challenge arose in the *Understanding Technological Approaches (UTA)* project (Kimbell et. al., 1994) that we undertook immediately after the *APU* experience. We were very aware of the limitations of the APU determined to investigate 'real' project work, over 'real' time, and with all year groups from Year 1 to Year 11. This was 1992 and the National Curriculum had made design & technology compulsory for all learners throughout these compulsory years of schooling. So what went on in these projects? Did teachers do the same kinds of things in all these years? The research design issues were interesting and essentially we settled on an approach that required us to be *observers* of activity in the classroom. But as any research manual will testify, being an 'observer' is far from straightforward. Do we intervene and ask questions of the learners or not (participant or non-participant observers). Do we record the process with audio or video? Do we explain who we are and what we are doing – or do we pretend to be wallpaper? The question that dominated our thinking was 'what is it we are going to observe'? A class full of learners working on a task will generate a prodigious amount of 'stuff' to be observed. Are they smiling or frowning? Talking or silent? Working in groups or alone? With numbers or drawings or words? Engaged or off-task? Undertaking interesting or banal work? Mechanical or visual or digital? And so on ad infinitum.

Moreover, the research design was based on using a small team of researcher-observers, each taking a set of schools and somehow observing the *same things*. We had to decide what *was* to be observed and what was *not*. We had to develop an instrument that would allow simple (but

RESEARCH

specific) observations to be transformed into recorded data. It was, by some distance, the most comprehensive observation-based project we have undertaken, and the instrument we developed for it had a number of interesting features.

Designing effective observation

The first challenge arose from the fact that we were, in each school, attempting to observe a **process** in action; a process of design & development by learners, managed by teachers. But whilst processes are continuous, observations represent a moment in time. So how many moments need to be observed in order to gain a 'true' record of the evolving process? This is a bit like plotting points on a graph. How many data points are needed to render a valid representation of the curve?

The question is informed by how long it takes to make the observation. Is it an instant thing or does it take 30 seconds, or 1 minute, or 2 minutes? As an example, if we were really trying to observe how much of a lesson the learners were smiling and how much they were frowning, then that is pretty well an instant decision and the observer can just hit a tick/cross list. But there are 25 learners in the class, and 25 ticks/crosses will take (maybe) 1 minute in total. So for any one learner we end up with *episodic* data – every minute. But we still get (say) 90 bits of data per learner per double lesson, and in reality of course we need FAR more other data to make sense of their activity.

What and who to observe?

All kinds of data might inform our understanding of what is going on. We would like to know about the *task* they are undertaking; about the specific *subtask* that they are doing at this moment; about whether they are doing it alone or in a *group;* about whether the teacher is *interacting* with them or not; about what *kind* of interaction it is; about their *engagement* with the task (motivated or disenchanted) and so on. Once again, all this takes a significant amount of time (say 1 minutes). But there are 25 of them in the class – so now we have episodic data every 25 minutes on an individual. This is clearly not adequate to reflect the evolving activity. We were forced by this process to focus our observation not just on specific things but on specific learners; and we chose four learners in each group to follow in detail. The choice of these four was done very carefully in discussion with the teacher. We asked to follow:

- The very best designer
- Two middle of the road designers (ideally one male/one female)
- A low ability learner who was nonetheless making progress with design & technology.

All four needed to be good attenders as there was little point collecting a huge quantity of data on learners who were frequently absent. The decision to follow four learners was made in association with other related decisions and involved a difficult optimising process:

- How many observations do we want to make?
- How long does it take to make them?
- How many learners can we follow?
- How episodic does the data therefore become?

In the end we evolved a system with an episodic cycle time of 5 minutes. In that time we could observe the detailed behaviour of four learners across a rich variety of data. But our decision might have been different. It might have been more data on fewer individuals; or more data on more individuals with a longer episodic cycle. This is the hard stuff of design decision making in research.

Transforming text notes into tick-lists

To an extent we were able to speed up the process of data capture. Initially, we just had an A4 pages with lines ruled across it leaving us with 50 mm of space for each 5 minutes in which we scribbled as furiously as we could to capture what was going on. We had a time box in each slot and could fill that in before the lesson started (e.g. 9.05, 9.10, 9.15, etc.), and we then used the empty space to make notes on what was happening. We had four sheets – one for each learner.

Through a series of school trials we gradually derived a list of things that we believed were more important than other things and that were happening all the time – like communicating. So we evolved a tick box to identify whether there was a teacher/learner interaction at the moment of observation. More than that we were able to identify what *kind of interaction* it was, at least in terms of who initiated the interaction. Was it initiated by the teacher (providing guidance/instruction to the whole class or to a subgroup containing our observed learner) or to the individual learner? Or was the interaction initiated by the learner (seeking specific support from the teacher)? Two ticks in related boxes could now represent a complex interaction, the noting of which had previously taken a lot of free text.

Pace

Having observed only a few lessons it became obvious to us that we needed some measure of the learners' engagement with the task. We wanted to distinguish between learners who were disenchanted or disengaged or just off-task, from those that were fully engaged, crashing ahead purposefully and at pace. In trying to record these data we identified a middle category who were in what might be termed 'tick-over' mode; doing enough to be seen (by the teacher) to be working, but more going through the motions than making real progress.

We evolved a categorisation of these behaviours into a three-point tick box:
• **Stationary** – going nowhere/off-task
• **Poddling** – in tick-over mode
• **Motoring** – fully engaged, making real dynamic progress

With intervening categories (e.g. between poddling and motoring) we had a 5-point scale to capture this level of engagement. It proved very easy and reliable to note and the resulting data rendered really valuable insights into learning and teaching practices.

S	P	M	time
dir	sup		
T	C	M	

Definitions

Level of engagement	*Teacher intervention*	*Issues child is dealing with*
S = Stationary	dir – direction	T = Task issues
P = Poddling	sup = support	C = Communication
M = Motoring		M = Making issues

Figure 1. Engagement and interaction elements of the observation framework

Behaviour or intention?
One of the problems of observation data is that some of the important things that are happening in a classroom are not observable. This is not because the learners are hidden behind a cupboard or facing the wrong way – but because the important thing is literally not externalised as behaviour. Rather it is going on at an inner level of cognitive processing. One of these inner levels that interested us a great deal was learners' design intentions. You cannot *observe* intention. It is not a *what* thing or a *how* thing, but a *why* thing. As a result of repeated trials we had created a list of observable behaviours enabling us to capture (with a simple tick) all kinds of workshop-related activity – are learners measuring, cutting, filing, shaping, drawing, etc. The lists initially got longer and then shortened as we categorised and streamlined them. But designing is *purposive* behaviour and the more we collected the behavioural data the less important it seemed to be. Does it matter if a learner is filing a shape out of a piece of acrylic sheet or whether they are hammering a piece of metal? What matters is why they are doing it.

• Is the acrylic filing in order to produce a finished *object* or *component*?
• Or is it to produce a *template* that can be marked around to produce standard components?
• Or is it to produce a *transparent template* that can be marked around at the same time as seeing something important through it?

These different ways of thinking that might inform the edge-filing of a piece of acrylic might reflect significantly different levels of designerly behaviour. Even though the behaviour is the same.

The only person who knows what the intention is of a piece of behaviour, is the person exhibiting that behaviour. So we were committed to talking to learners about what they were doing in order that we could understand why they were doing it. It was for this reason – amongst others – that we chose to be *quasi-participants* in the observed lessons; rather than pretending to be wallpaper. However, our questioning always remained 'neutral' seeking out why they were doing something rather than commenting on whether we thought it a good thing to do, or suggesting other things that they might be doing. The learners became accustomed to us constantly moving around the room noting things on pads, and that we might occasionally wander over to see (and ask about) how they were getting on.

Intentions	**Manifestations**	
generating	discussing	making
mod exploring	thinking aloud	-cut
developing	looking	- join
modify	drawing	- fit
detailing	reading	- mould
constructing	writing	- mix
planning	listening	- finish
organising	waiting	- base
investigating	arranging	- add
receiving	selecting	preparing
evaluating	measuring	testing
reviewing	marking out	cleaning up
recording		off task
explaining		
presenting		
seeking help		
intentionless		

Figure 2 The list of design intentions and their manifestation noted on the observation sheet

RESEARCH

We acknowledge that this observation process will also have changed to some extent the behaviour being observed. This is the perennial dilemma of the observer. The more you get involved, the more you find out. But the more you get involved the more you influence what happens, when what you really want to know is what would be happening if you were not there. There are difficult trade-offs to be made here, but the importance of *intention* in design behaviour is so overwhelming that we were obliged to gather it. Often we felt confident in *inferring* an intention from the combination of behaviours we had noted. But occasionally we had no alternative but to ask. Noting the intentions behind the behaviours and the way the behaviours were manifest, gave a rich picture of the different ways learners approached their designing.

Bringing data alive: the art of data compression
The *UTA* project generated oodles of data collected from countless hours of observation. *APU Design & Technology* (Kimbell et. al, 1991) that preceded it, had generated even more – though this was test performance data from a huge sample. But the question inevitably arises with so much data. How should we set about making sense of it all?

Our general approach to data analysis has typically involved a search for patterns in the data and (being designerly folk) we work better with visual patterns than with any other kind. So wherever possible we find ways to represent the data graphically so that trends and anomalies stand out as visual signposts to something interesting that might be happening.

This approach was one we developed during *APU Design & Technology*, very much supported by the team's decision to buy its first Apple Macintosh computer. Up till this time we relied on posing a research 'hunch' to the team's statistician who went away to run a very time consuming data analysis process on the College mainframe computer, producing for us (often 24 hours later) the answer to a question we were no longer interested in. With the introduction of our first 'Mac', we could suddenly explore the data for ourselves, ably supported by the statistician, and utilise the Mac's simple graphics software to visualise our findings. The following examples illustrate this approach.

First, when exploring data on the comparative analysis of girls of different abilities, we noted that sometimes the lower ability girls did considerably worse than the mid-ability girls, whilst sometimes they were almost on a par with them. Our hunch was that this had something to do with the way the tests were structured, and so we presented the data in such a way that the most loosely structured tests were at one end of a continuum, the most tightly structured at the other. As can be seen from Figure 3 below, the more tightly structured the test, the better the performance of the lower ability girls – and equally interesting – the apparent lack of importance this has for high ability girls.

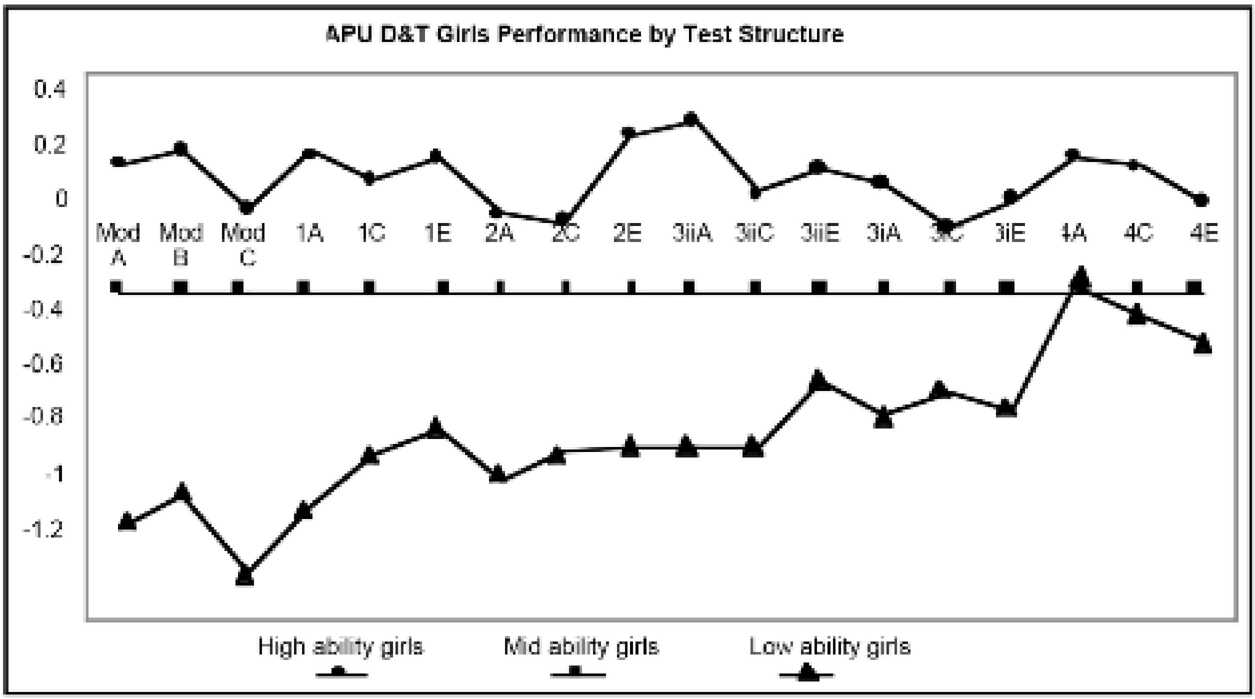


Figure 3. Presenting performance data based on loose/ tight task structures in the APU project

APU holistic performance by context (A, C & E) and gender

* indicates 5% sig.
** indicates 1% sig.

	People	Environment	Industry
T.1 Boys	2.19	2.03	1.71
T.1 Girls	2.42 *	2	1.96*
T.2 Boys	1.89	1.77	1.99
T.2 Girls	2.02	1.85	2.21**
T.3i Boys	1.97	2.23	1.87**
T.3i Girls	2.06	2.19	1.58
T.3ii Boys	1.96	1.8	2.41
T.3ii Girls	2.08	1.77	2.29
T.4 Boys	2.15	2.09	1.96
T.4 Girls	2.5**	2.40**	2.29**
Mod Boys	2.29	2.55	2.65*
Mod Girls	2.38	2.37	2.38

Figure 4. Average holistic performance data across all tests in the APU project

Using graphics helped us make sense of the data for ourselves – and also when communicating this with others. By contrast, we show below the raw composite data from the APU project for holistic performance of girls and boys across the three design contexts that the tests were set in. To the naked eye, it seems an unintelligible set of figures.

As we considered this data we were conscious of a gender effect related the context of the test – girls tending to outperform boys when the context focused heavily on people, more mixed effects when the focus was on industry and virtually no effect when the emphasis was on the environment. Presenting the data in a graphic form makes this effect far more visible, as is shown in Figure 5.

Giving further consideration to the 'mixed messages' of the industry- focused test, we became aware that there were two effects in the data – context *and* test structure. At times these effects were working in the same way for a gender group, at times they were working in opposition. So once again, using the same raw data, we could show these different effects graphically.

The *UTA* project also provided several classic examples of how this pattern-seeking approach yielded interesting interpretations of the work that learners were undertaking. The starting point in seeking patterns involved developing approaches that make it possible to compress huge quantities of data into relatively simple data sets. First, we entered all the observations as raw data in a spreadsheet. So, taking the example of *interaction* between teacher and learner, we had a column in the data record showing (for

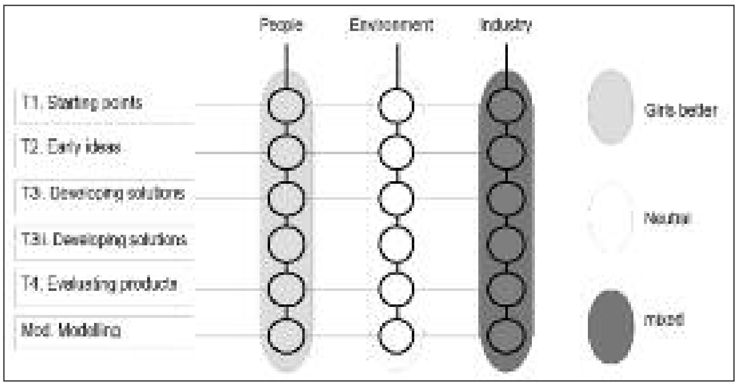


Figure 5. Graphical presentation of the numerical performance data

every 5 minutes period) whether the learner was interacting with the teacher and – if so what kind of interaction it was (e.g. *directive* from the teacher or *supportive* sought by the learner). These data were represented in a single code within the column, and the column ran for the entire duration of the project. A typical case was a project with 14 year olds that ran for 485 min, with 97 units of coded data.

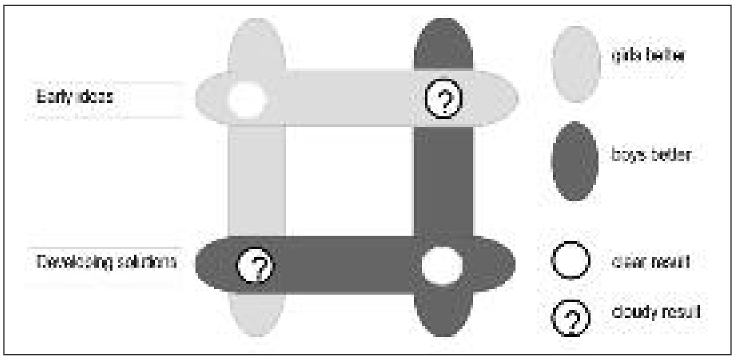


Figure 6. Graphical presentation of numerical data to indicate the effects of context and test structure

From this data we could see, over the life of a project, what percentage of time the learner was seeking support from the teacher, and conversely what percentage of time teachers were being directive. Since we had identical data across all 11 years, it was then a simple matter to represent it graphically (Figure 7). With startling consequences, for immediately it became obvious that something odd happens in the transition from Year 6 to Year 7.

However, whilst charts of this kind are highly informative of generic data, they also tend to hide *trends* in data because of the averaging effect across the life of the project. Since we were concerned with designing as a real-time rolling

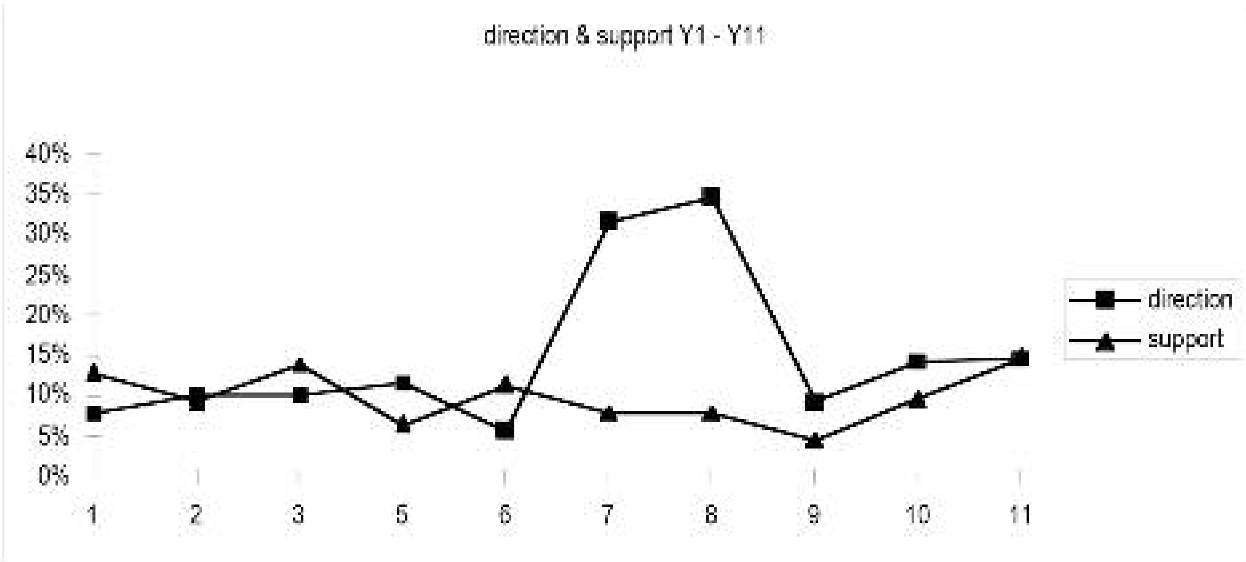
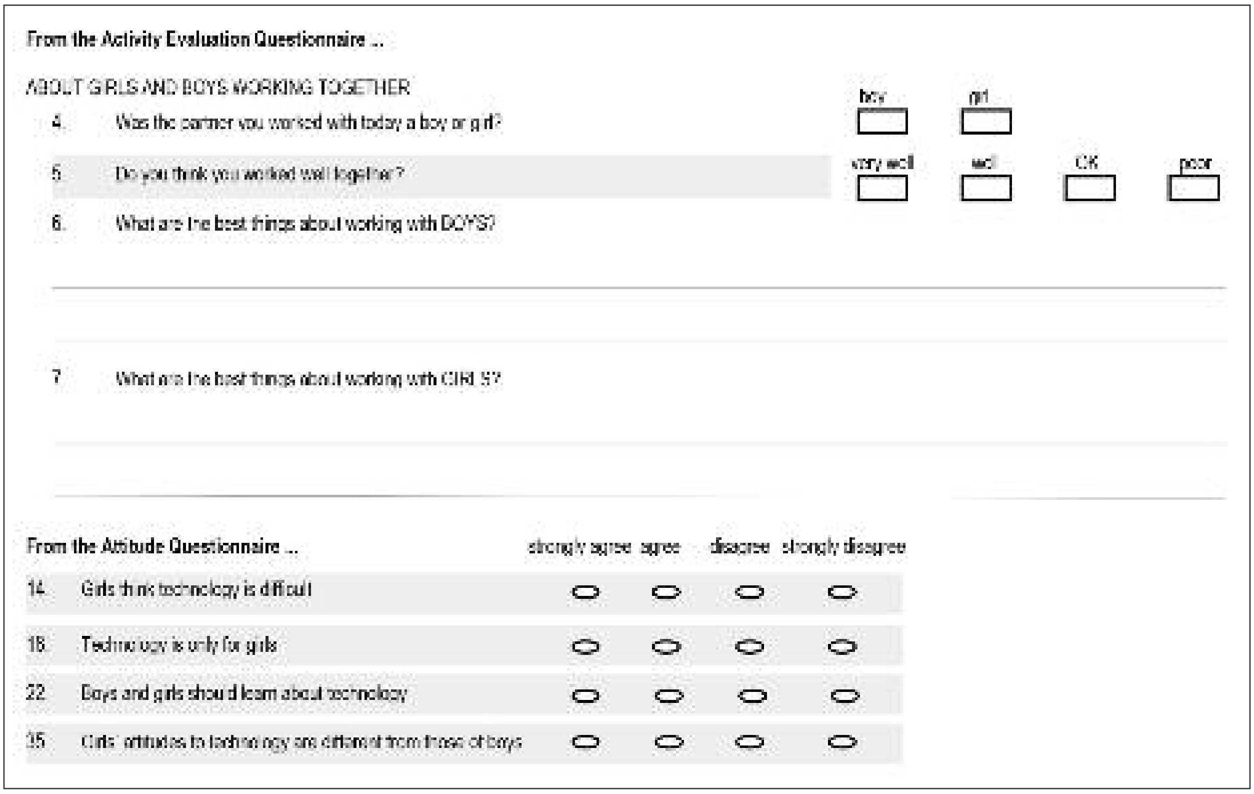


Figure 7. The levels of teacher direction and support across the UTA projects



From the Activity Evaluation Questionnaire ...

ABOUT GIRLS AND BOYS WORKING TOGETHER

4. Was the partner you worked with today a boy or girl? — boy ☐ girl ☐
5. Do you think you worked well together? — very well ☐ well ☐ OK ☐ poor ☐
6. What are the best things about working with BOYS?
7. What are the best things about working with GIRLS?

From the Attitude Questionnaire ...

		strongly agree	agree	disagree	strongly disagree
14.	Girls think technology is difficult	⬭	⬭	⬭	⬭
18.	Technology is only for girls	⬭	⬭	⬭	⬭
22.	Boys and girls should learn about technology	⬭	⬭	⬭	⬭
35.	Girls' attitudes to technology are different from those of boys	⬭	⬭	⬭	⬭

Figure 8. Collecting qualitative and quantitative data in the North West Province Technology Education Project Evaluation

process we clustered the data into five project phases, created simply by taking the first 20% of project time as phase 1, then the second 20% and so on. The resulting data-maps were *condensations* of the data. We referred to these condensations as *data-maps* since they enabled us to take huge amounts of data and reduce it down to a form in which we could create simple graphic representations of the trends that lay within it. (This is illustrated more fully in the Gender differences in technology paper, also included in this Special Edition).

Gathering data from different perspectives

As has been clear throughout the projects, we have been equally comfortable with combining research approaches – qualitative and quantitative, and different tools – if we judged they would provide us with rich data to inform our research questions. This has involved us, for example, developing parallel interview structures, where effectively the same question is being asked to different stakeholders. A clear example of this was in the *North West Province Technology Education Project* (Stables et. al., 1999; Kimbell & Stables 2007) where we used the same question structure to interview teachers involved in the project, their school principals, provincial and NGO officers, and through group interviews, the learners themselves. We have used this approach in a number of projects as the approach helps us to gain insight into an issue or situation from a range of perspectives.

A further approach to gaining a rounded, fuller picture of an issue has been to gather linked qualitative and quantitative data, the former allowing us to explore patterns and trends in the data, the latter to *illuminate* those trends and speculate on their meaning. The *North West Province Technology Education Project* also provides an example of how we collected a range of data that helped us to explore gender differences in and between the learners from schools involved in the initiative and those from the control schools. Figure 8 illustrates different types of data we collected: demographic data (the gender of the respondent and who they worked with); quantitative data about whether they worked well together and what their attitude to gender-related aspects of technology were; and qualitative data through a 'free response' question on the 'best things' about working with boys and girls. The composite insights provided allowed us, for example, to examine in detail the collaborative dimension developed through the initiative (Stables, 2000) and the capability and attitudes it enhanced (Stables & Kimbell, 2001).

Visualisation to support data capture

The process of rendering abstract ideas into visual form is something that we have consistently sought to do, and not just for analysis and interpretation purposes. Two other instances are worth a brief reference. In the *North West Province Technology Education Project*, learners were assessed on design tasks derived from a similar approach to the APU project and the resulting work was to be assessed by associate researchers from the project development team who had to be trained in making holistic assessments. We decided to operate this through a two-stage process. First, using an assessment rubric, we worked through the learners' responses looking for evidence of the qualities identified in the rubric. Second, having identified the evidence, we sought to attach values to it, enabling us to assess all the work consistently.

It was in the first of these processes that we used a very simple, but effective, visualisation tool. We provided 'high-lighter' pens for the assessor team, using different colours for different qualities in the rubric. This highlighting process – done in pairs – then led to a group debate about the qualities concerned. Is this an example of quality X … and if so does it reflect high level performance or poor performance? Do you agree that that is an example of quality Y … and so on.

This sharing process – based on highlighted evidence – proved very helpful to assessors who were then moving on to value the work.

A different kind of visual approach was used for data capture in *Attitudes of Potential Teachers* (Kimbell and Miller, 2000). We were interviewing graduates from design, engineering and related degree programmes to tell us about their experience of design & technology in schools. Rather than merely present them with a bald list of bullet points to complete, we sought to appeal to more graphic/designerly instincts, and created the thumbs up/thumbs down images (Figure 9). They wrote their keywords inside these two images. We cannot say that it worked better than bald listing, but it did create an impression and it did work.

In both these cases the techniques might be thought to be barely noteworthy. But data capture is often a delicate and difficult exercise. In the first case (assessors colour coding) the learner responses are highly complex with many kinds of qualities interlinked and overlapping. The colour coding was a do-able task that simplified the process of assessment, perhaps not by a lot, but maybe by just enough to make a difference. In the second case, we can

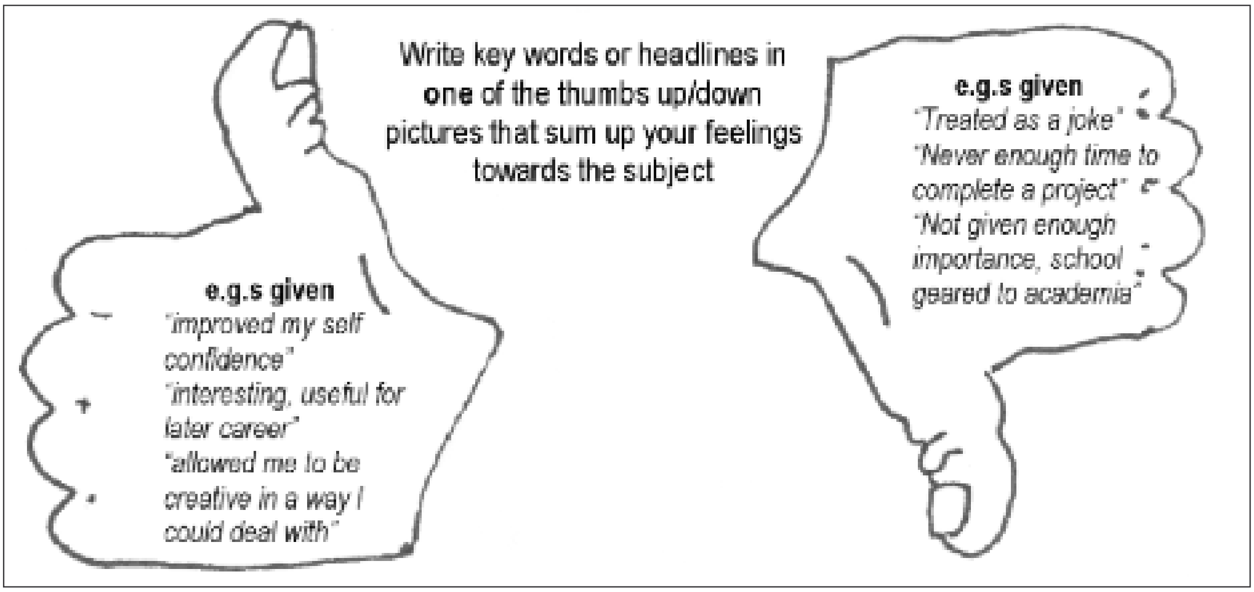


Figure 9. Using 'Thumbs up, thumbs down' to collect key words for positive and negative experiences

sympathise with those who find filling in forms a tedious process. So anything that we can do to lighten the task – and maybe raise a smile – is worth doing. It might just make the difference between engagement and disengagement. It is also an approach we have increasingly used with learners – as young as 8 years old – where the symbol of the thumbs give more instantaneous meaning than words could.

This latter technique also exemplifies how we have typically used everyday, and often vernacular, language to create metaphors for concepts we wish to share with research participants – as we did with the use of *stationary*, *poddling* and *motoring* as metaphors for learner engagement and pace. We have also used a *wow* <> *yawn* continuum for assessing creativity.

Research as Design, Design as Research

We recognise that the form of this chapter has implied a degree of linearity to the process of research. *First*, sort out your research questions; *then* resolve the research design; *then* design the data capture system…and so on. We tried several ways to organise the story of our research approach, and in the end it seemed best to do it this way. But we would like to enter a caveat here that cautions against a too sequential view of research processes.

Whilst it is broadly true that sorting out research questions is a primary task, and that it leads into questions about research design, as soon as we get inside a task we have found it helpful – and even necessary – to model what the data might look like and how we might capture it. This modelling process typically involves not only mock-ups of instruments of one kind or another – but also trials to see what happens when they get used. Sometimes this process reveals other features of the research task that we (perhaps belatedly) come to see as important and decide to find out about – so we modify the research design, redesign the instruments and trial it all again.

As with designing, the process is iterative; starting with a view of how we think the research task will shape up and what it involves, and then moving forward through a series of iterative steps (innovation–modelling–trialling, reviewing: innovation–modelling–trialling–reviewing) until we get to the point at which we think enough of the confusion is ironed out and the instruments work sufficiently well and reveal enough of the things we are interested in. Because at some point we have to draw a line under these iterations, cross our fingers, and just press the 'go' button.

The whole researching process is, as we keep saying, just like designing.

References

Kimbell, R., Mahony, P., Miller, S. and Saxton, J. (1997) *Decisions by Design: A research project commissioned by the Design Council* (March 1995 - April 1997) Final Report, London: Goldsmiths College/Roehampton Institute.

Kimbell, R. and Miller, S. (2000) *Attitudes of Potential Teachers of Design & Technology*, London: Department for Education and Employment (DfEE).

Kimbell, R. and Stables, K. (1999) *South Africa: North West Province Technology Education Project - an evaluation*, London: TERU, Goldsmiths College University of London.

Kimbell, R. and Stables, K. (2007) *Researching design learning : issues and findings from two decades of research and development*, Hardback ed., Berlin: Springer.

Kimbell, R., Stables, K. and Green, R. (1994) *Understanding Technological Approaches: Final Project Report to ESRC*, London: Goldsmiths University of London.

Kimbell, R., Stables, K., Wheeler, T., Wozniak, A. and Kelly, A. V. (1991) *The assessment of performance in design and technology*, London: SEAC / HMSO.

Stables, K. (2000) 'Learning technology through collaboration: examining the impact on learners in South Africa of introducing technology education through using collaborative design processes', in Theuerkauf, W. and Graube, G., eds., *International Conference of Scholars on Technology Education*, Braunschweig, pp 9.

Stables, K. and Kimbell, R. (2001) 'Technology Education in South Africa: Evaluating an innovative pilot project', *Research in Science Education*, 31(1), 71-90.

Stables, K., Kimbell, R. and Molwane, O. (1999) 'Technology Education in S. Africa: An evaluation of the impact of an experimental high school curriculum with particular reference to teacher pedagogy and student group work', in Roberts, P. and Norman, E., eds., *IDATER 99: International Conference on Design and Technology Educational Research and Curriculum Development*, Loughborough University, Loughborough University, 177-183.

Researching performance based assessment: authenticity in assessment activities and processes to support the development of learner capabilities

Kay Stables, Goldsmiths, University of London

RESEARCH

Introduction for the 2015 DATA Special Edition
This paper was originally presented as a Keynote address at the Southern African Association for Research in Mathematics, Science and Technology Education 21st International Conference, held at the University of the Western Cape, Cape Town, in January 2013. I was asked to present a Keynote that focused on assessment and, mindful of the conference theme of making Mathematics, Science and Technology Education socially and culturally relevant in Africa, the paper took the concept of authenticity as a major thread. The presentation drew on a number of TERU research projects, including one that Richard and I had conducted in South Africa in 1999. Using the projects as case studies, issues of authenticity were explored in relation to summative and formative assessment practices and related pedagogic approaches. Through an exploration using validity, reliability and manageability as lenses, the presentation offered some concluding comments on possible challenges and the potential of drawing on the research presented in a Southern African context.

The paper here, that documents the keynote, has not previously been published.

Introduction
Over the last 30 years, researchers in the Technology Education Research Unit at Goldsmiths have been investigating ways of assessing learner capability, initially Design and Technological capability and latterly across a broader curriculum base. The focus of the research has been to find ways of understanding learners' abilities in procedural settings and so has focused on creating authentic assessment activities that generate authentic evidence of capability. A considerable amount of this research has been in the context of high stakes summative assessment, developing valid, reliable and manageable assessment activities that can be used in the context of national assessments. An underlying model has been established - the 'unpickled portfolio' (Stables & Kimbell, 2000; Kimbell & Stables 2007) that structures short assessment tasks, documented through portfolio-based responses, set in problem/challenge-based scenarios. Initially paper-based and more recently digitally captured, these assessment tasks have been used with primary and secondary aged learners, across a range of curriculum areas and in contrasting national settings, including, in 1999, in South Africa (Stables et al., 1999, Stables & Kimbell, 2001).

This paper takes a journey through a series of the research projects, from the first in the late 1980s where the initial approach was developed, creating dynamic, iterative assessment portfolios on paper, to recent projects that use mobile technologies to capture evidence of capability directly from learners as they convey their ideas and thinking through audio, video, text and image based tools. The journey provides insights into fundamental concepts behind the structure of the assessment tasks and portfolios – holistic performance, procedural capability, the iteration of active and reflective sub-tasks and authenticity in tasks and evidence. A framework illustrating how learning intentions can be mirrored with assessment intentions shows how constructive alignment (Biggs, 2003) can be achieved. Case studies from research projects illustrate how the model has developed to be effective in formative, diagnostic, summative and evaluative settings. The case studies also show how the model supports Problem Based Learning, enables collaboration and team work within an assessment setting, facilitates peer and self assessment by learners and enables a range of learning styles to be taken into account in collecting assessment evidence. It also reveals how teachers and learners can become involved in a radical approach to making assessment decisions – using Adaptive Comparative Judgement (Pollitt, 2012; Seery et al, 2012).

The provenance of the assessment activities
Rejecting linear models of process as being more about management than designing, our research afforded us the opportunity to explore alternative, more authentic perspectives. We developed a holistic and iterative view of designing that focused on active and reflective processes and the progressive relationship between these as a designer (or learner) progressed an unformed 'hazy' idea through to a well-developed prototype. (Figure 1)

This model of process doesn't deny features of more linear models (identifying problems, conducting research, generating and developing ideas, finalizing solutions, and evaluating). Indeed it recognizes that these 'sub' processes are present in large measure. But what it does reject is that they occur in a given, prescribed order. Rather

Researching performance based assessment: authenticity in assessment activities and processes to support the development of learner capabilities

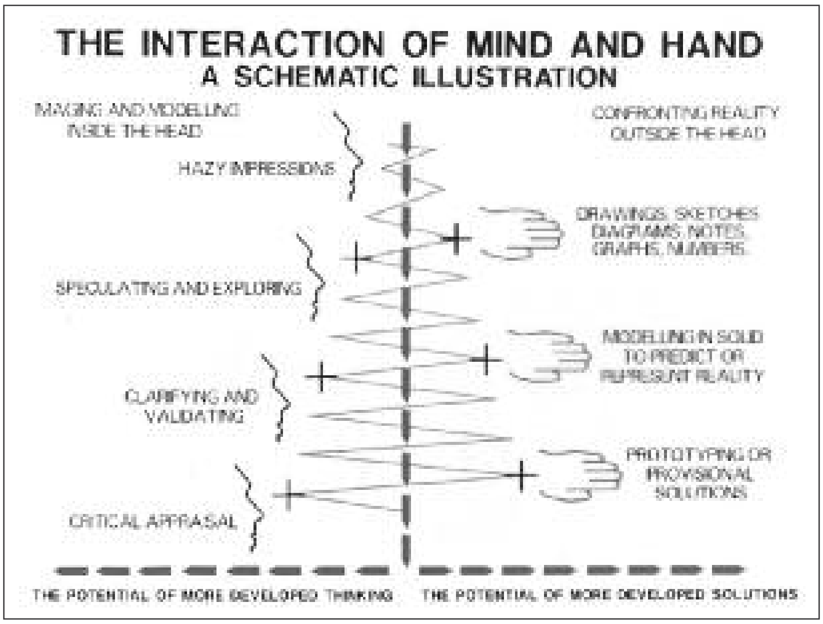


Figure 1. The APU Design and Technology Model

it accepts that ideas are initiated throughout the process and that in moving towards 'developed solutions' problems are solved, research is conducted and judgments are made, driven by the desire to reach a prototype. This idea has resonance with the concept of design tasks as 'wicked' tasks, an idea introduced by Horst Rittell in the 1960s (Buchanan, 1995), that speaks to the nature of design tasks as indeterminate; with no clear, correct answer; in which the designer is operating on shifting sand, without all the knowledge required up-front; and managing ill-formed client expectations. In characterizing this complex process, Lawson (2004) described the process of designing as being like playing chess with minimal rules, but a clear intent.

> Designing then, in terms of chess, is rather like playing with a board that has no divisions into cells, has pieces that can be invented and redefined as the game proceeds and rules that can change their effects as moves are made. Even the object of the game is not defined at the outset and may change as the game wears on. Put like this it seems a ridiculous enterprise to contemplate the design process at all. (Lawson, 2004, p. 20)

This description presents a design process as highly complex and at the same time captures the essence of its reality. It matches well with the view of process that we set out to assess. The challenge was to work out how we could assess capability evidenced through such a process, and to do so validly, reliably and in a managed way. An underlying approach was established through the initial APU research project – what we came to call the 'unpickled portfolio' (Stables & Kimbell, 2000; Kimbell & Stables 2007). We chose this label because of the way evidence of capability was generated and captured in a short time frame – as opposed to more typical long projects where learners are steeped – or pickled – in the good ingredients of designing, learning and teaching experiences. The activities are structured through a series of sub-tasks that are choreographed to enable a dynamic relationship between active and reflective modes of designing. In the 'high stakes' assessment mode, standardisation is increased through the use of an administrator's script that prompts each aspect of the activity and controls the time spent. All evidence of the work produced is documented in a portfolio.

The importance of authenticity

From the outset of the original research project, we were concerned with authenticity. As has been described above, an immediate concern was for authenticity of process. This was based on the premise that if you want to know if someone is capable, then you need to be able to see them operating in practice. Put simply, if you want to know whether a learner can design, then you need to create a situation (or activity) in which they have the opportunity to design and, through this, to make explicit the evidence of their designing. So tasks, activities and challenges also needed to be created and, in parallel with our concern for authenticity in the process, we were equally focused on authenticity in the assessment tasks learners were presented with.

Broadly speaking, we have focused on two aspects of task creation, the context in which the activity is set and the way in which the activity is structured. With the former our belief is that the task should be embedded in a context that is relevant to the learner and is presented in a way that allows them to engage and take ownership of their task. In order to achieve this the task should be 'issues rich' such that there is complexity - the learner has plenty to get their teeth into and be challenged by. To address this we have introduced a number of ways of introducing and 'fast forwarding' learners to the starting point of a task using devices such as stories, short videos, scenarios etc.

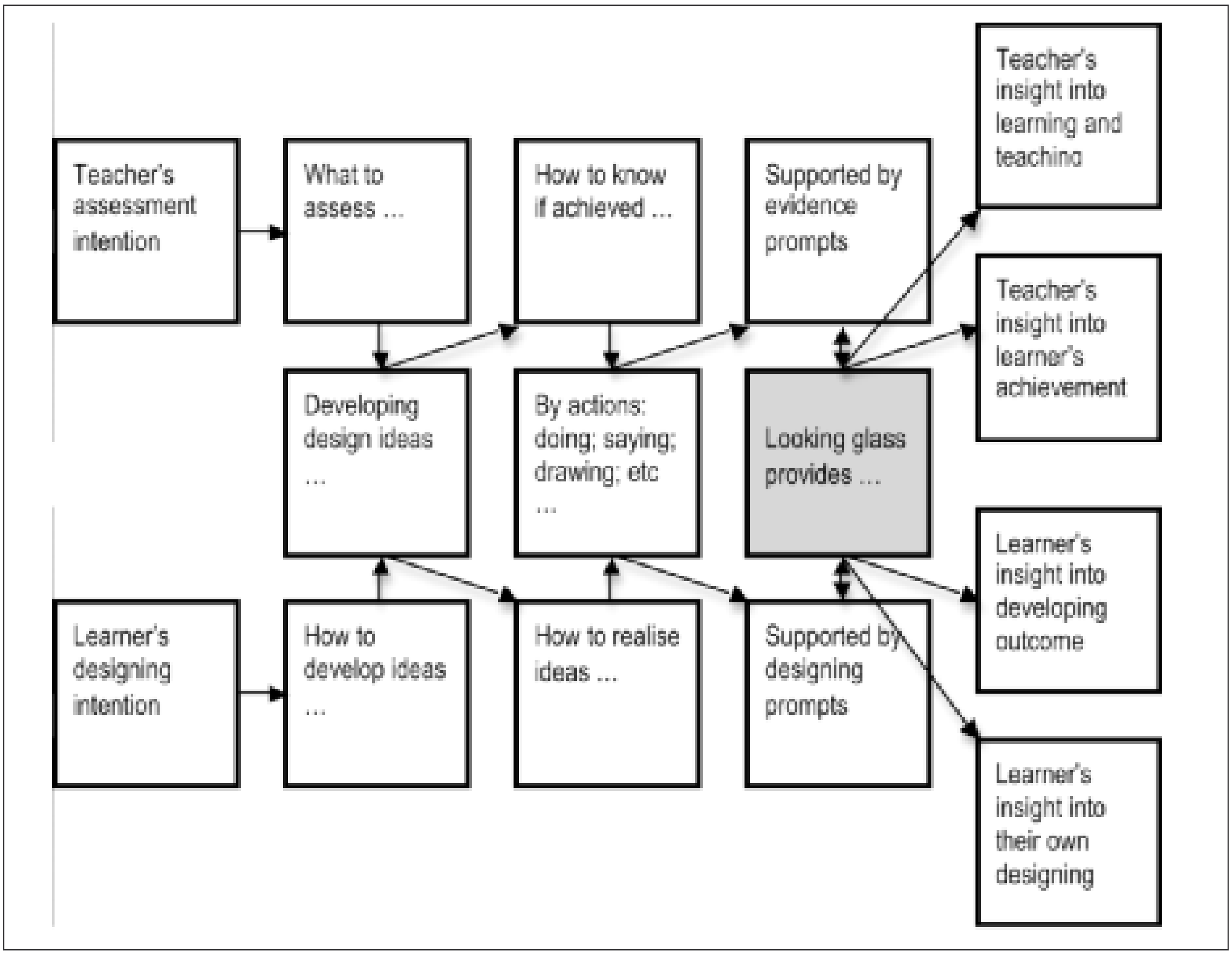


Figure 2. The mirror effect of effective evidence prompts

Having created a starting point, the detailed structuring of the rest of the task is equally important if it is to provide authentic evidence of capability. It is also important that this evidence relates to the assessment criteria that have been set. In all of the research and development work on assessment that we have undertaken, there has been an undeniable link between the nature of the activity that the learners have been asked to engage in and the nature of the criteria that have been used to analyse their work. In essence, there is a reciprocal relationship. If the activity is authentic from a design viewpoint and at the same time provides explicit activity prompts that draw out evidence of the qualities under scrutiny, then both the activity and its assessment are likely to be valid. This relationship was first made explicit when we were devising activities to assess young children's (5-7 year olds) technological capability (Stables, 1992) and we have found it to hold true in subsequent research.

Assessment activities, whether teacher-led or imposed by an external body, attempt to generate evidence of what a learner can do by prompting some kind of response. At the simplest level this might be asking the learner a question. This exposes the learning to the scrutiny of the assessor and is the first and most obvious purpose of evidence in an assessment setting. At a deeper level however if (in the eyes of the learner) the activity is sufficiently authentic, then the prompted display of evidence enables the learner also too to 'see' (probably for the first time) the evidence that they have just created. Reflecting on this evidence enables the learner to improve whatever they are doing. So not only do assessors gain insight into the learner, but so too do learners themselves. When managed effectively, their thinking is laid bare for them to see and to benefit from. It is as if the performance is being observed in a mirror – and a mirror where both the teacher and the learner can see double-sided reflections that support both summative and formative assessment and also learning and teaching.

Researching performance based assessment: authenticity in assessment activities and processes to support the development of learner capabilities

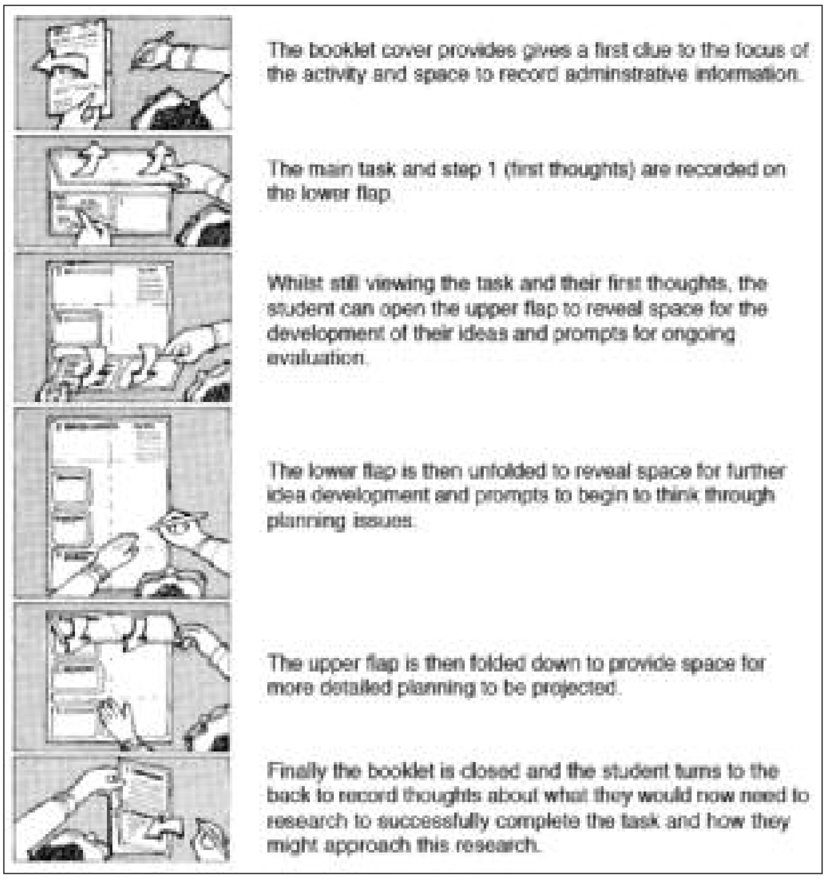


Figure 3. The unfolding booklet of the "unpickled portfolio"

In this way we are creating constructive alignment (Biggs, 2003) between the desired learning and the evidence created for assessment. In formative assessment the model also provides for alignment between current and future teaching and learning and supports the learner in self-assessment. We have found that the more we embed and iterate active and reflective evidence prompts into an assessment task, the more we build meta-cognitive potential for the learner – helping them to make their own learning visible. We have explored ways of enhancing self-assessment towards sustainable assessment (Boud, 2000), for example by adding prompts for learners focusing on the following:

- I was best at…
- The easiest thing was…
- The most difficult thing was
- Today I learned…
- I want to get better at are…(McLaren et al., 2006)

Authenticity and criteria for assessment

If constructive alignment between learning and assessment is to be achieved, then the assessment criteria themselves need to be 'authentic' in terms of what they are attempting to reveal. In the original APU project the model of process we developed became important in identifying a framework for assessment. Our hunch was that we would need to consider three aspects as important

- The inclusion of key elements of the process: identifying and addressing issues in the task; having a grip on generating ideas and developing solutions; appraising their thinking with a sound, critical eye
- Interconnectedness of the iteration between thought and action
- Viewing all evidence holistically.

The latter of these is particularly important as it allows us to see elements of the process at whatever stage they appear, rather than anticipating that they appear in a neat, linear fashion. It allows us to take an overall position on the learner's achievement in the assessment task and then to look inside to identify the strengths and weaknesses within their work. In the APU project we explored this hunch empirically – we had 20,000 portfolios to analyse. 100 teachers worked with us as assessors and their first instruction was to look critically at all the work in a portfolio to see what the learner had done, what they had tried to achieve and how they had approached this, and then to make a holistic, professional judgement about the overall quality of the work. The teachers were supported in making this judgement through a training process and through having 'exemplar' scripts (portfolios) that had already been assessed by the research team. Following this initial judgment, they were provided with a rubric that asked them to make progressively smaller, more focused judgements. Counter-intuitively, the smaller and more focused the judgments became, the less statistically reliable they were found to be. (Kimbell et al., 1991)

This holistic approach to assessment, initially developed as a research tool, has been explored, augmented and developed through a range of our projects and has been shown not only to support the authenticity in the process, but also act as an important professional development tool for those engaged in the assessment process – as will be illustrated by the case studies that follow.

Researching performance based assessment: authenticity in assessment activities and processes to support the development of learner capabilities

Capturing capability – learning styles and designing styles, going digital
Having set the learners an assessment challenge, learner responses are collected through a portfolio, structured through a series of prompts. In the initial project, the portfolio was designed as a large (A2) sheet, folded into a booklet that progressively unfolded as the learner moved through the assessment activity, as shown in Figure 3. This design allowed learners to have sight of their work as it progressed, rather than hiding it from themselves by turning a page.

The booklet encourages learners to draw and/or write, as they see fit, and has proved to be an effective way not just of capturing evidence of capability, but also revealing different styles of designing that have also related to learning styles (Lawler, 1999; 2006). In recent research we have moved the portfolio into digital mode through the use of mobile devices such as PDAs, mobile phones and netbooks. This move has afforded even greater opportunity to take account of learning and designing styles and preferences, as learners have been able to document their process through the task using a combination of text, drawing, audio, video and photo tools. This has provided both flexibility and speed in the ways in which ideas and thinking can be documented. For example, to convey an idea and the thinking behind it, a learner can:

- draw on the PDA screen and then annotate by writing freehand on the screen;
- photograph a sketch model and annotate by adding a voice memo;
- use the video facility to 'fly through' a sketch model, or show it in action, using voice-over;
- draw an idea on paper, photograph it and add a voice memo;
- draw and annotate an idea on paper, then photograph it;
- photograph a sketch model, and then sketch or annotate freehand on top of the photo image on screen;
- any combination of the various techniques above.

The increased range of ways in which learners can capture the evidence of their ideas and their thinking has further enhanced the authenticity of both the activity itself and the evidence of capability that is generated through it.

Case studies
The case studies that follow have been chosen to illustrate how the model has developed to be effective in formative, diagnostic, summative and evaluative settings. They show how peer and self assessment can be facilitated and how learning styles can be taken into account. They also illustrate how collaboration and teamwork within an assessment setting can be enabled. In addition to the APU project, the following three projects will be drawn on to exemplify these various aspects.

The North West Province Technology Education Project Evaluation (NWPTEPE) (1999).
As part of the South African Curriculum 2005, this project, funded by UK Department for International Development (DFID) and South African NGO PROTEC, was a three-year pilot of a Technology Education curriculum in a number of schools in the North West Province of South Africa. The project was conducted with Years 10, 11 and 12 learners and ran from 1997 to 1999. The DFID commissioned TERU to evaluate the impact of the pilot. We were required to assess the capability of learners that had engaged with the curriculum in pilot schools in comparison with learners in schools that had not. The assessment activities, based on the unpickled portfolio model, were designed to take account of the features of the pilot. This meant that the assessment activities explored learners understanding of materials and processes, energy and power, and communications technologies through problem-based approaches and teamwork. The evaluation compared 10 pilot with 10 non-pilot schools. In each school 18 learners were involved in the assessment activities. In addition teachers, headteachers and learners were interviewed about their experiences. To support capacity building, six South African fieldworkers contributed to the evaluation. (Stables et al., 1999)

Assessing Design Innovation (2002-2004).
This project, funded by the UK Department for Education and Skills, was prompted by a concern for the way in which assessment was driving creativity and innovation out of the D&T curriculum. The focus of the work was on developing high-stakes assessment activities for the GCSE (16+) assessment in D&T. The activities were created in conjunction with practising teachers and examination Awarding Organisations. The unpickled portfolio approach was developed into a six-hour activity focusing on creativity and innovation. The project included two important innovations. The first was the use of "critical friends" (Costa and Kallick, 1993) within the assessment process and the second was the introduction of 3-D modelling, evidence of which was captured through photographs taken throughout the activity and pasted as a digital storyline into learners' portfolios. The activities were adopted as a model of "constrained assessment" and now feature in the menu of assessment activities available within GCSE exams in D&T. (Kimbell et al., 2004)

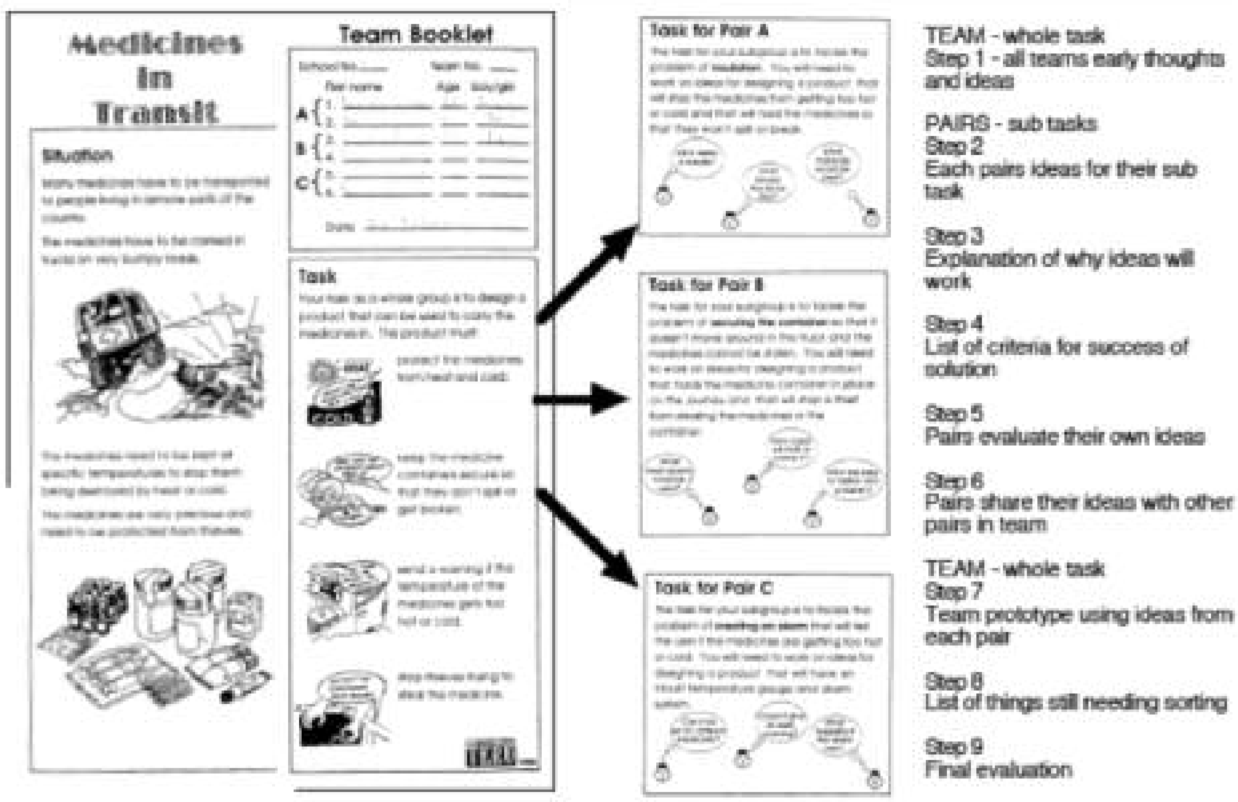


Figure 4. The challenge for the NWPTEP assessment task.

e-scape (e-solutions for creative assessment in portfolio environments) (2004 -2009)
The e-scape Project built directly from Assessing Design Innovation and explored further the possibilities of digital capture in performance-based assessment activities. A system was created that enabled assessment activities to be designed and presented to the learners through mobile devices such as PDAs, mobile phones and netbooks. The work undertaken by the learners, is documented through text, voice, video, photo and drawing and is synchronised dynamically with a web space while the learners are working. Assessment of the work takes the idea of holistic assessment one step further to allow for Adaptive Comparative (pairs) Judging, explained below under Making Assessment Judgements (Pollitt, 2012). The e-scape project has explored the development of dynamic digital portfolios and pairs judging in different disciplines, across age groups, and in a number of different countries including Scotland, Australia, Ireland and Israel. The Israel e-scape project, entitled Assessment in my Palm, expored the use of e-scape for formative and summative assessment in ongoing class projects. (Stables & Lawler, 2011)

Creating authentic contexts
When assessing performance capability, engaging learners at the outset of an assessment activity is important. Two examples are given illustrating quite different approaches to doing this.

In the APU project this was challenging for two reasons: first the whole activity had to be conducted in 90 minutes and second the learners were being assessed on 'design and technological capability' before a subject of that name had been created within the UK curriculum. This meant that many students were being assessed on something they didn't study in school. In order to 'fast forward' the learners into the assessment activity, we created a series of short (6 minute) videos that presented issues-rich snapshots into a particular scenario. For example, one focused on the challenges of the elderly, carrying heavy food shopping, reaching to store it in low and high cupboards, opening packaging and preparing and cooking foods. In addition, the learners were put into 'role' – they became part of a design team with individual responsibility for certain phases of development.

The NWPTEPE presented a different challenge. First we had to create a task that had relevance for learners living

'LIGHT FANTASTIC' TASK

A light-bulb company wants to minimise packaging waste and extend the product range they offer. They want a new range of light-bulb packaging that people won't throw away.

Your task is to come up with exciting ideas for light-bulb packaging that people won't throw away because it transforms into interesting lighting features & structures.

By the end of the activity you must have produced

- a working light-bulb package containing everything for the lighting feature;
- an assembled lighting feature:
- a persuasive argument for your product to attract purchasers.

Outline structure

1. read task to the group and establish what is involved
2. explore a series of 'idea-objects' on an 'inspiration table' and in a handling collection designed to promote ideas for transformation
3. put down first ideas in a designated box in the booklet
4. swop work within team - for further development by team mates
5. work returned to 'owner' to consider which ideas to pursue
6. teacher introduces the modelling/resource kit
7. learners develop their ideas through drawing – and/or through 3D modelling.
8. learners reflect on the user of the end product and the context of use, before continuing with development
9. at set intervals, learners pause and throw a 'questions' dice, e.g. 'how would your ideas change if you had to make 100?'. Answers recorded in their booklet
10. approximately every hour photos of modelling taken to develop visual story line of evolution of design ideas
11. end of 1st morning, learners reflect on own and team members work
12. 2nd morning starts with celebration of work from day 1 using 'post-it' notes to highlight 'best' idea, 'wackiest idea' biggest problem' and 'next steps'.
13. prototype development continues
14. hourly photos and pauses for reflective thought continue
15. final team reflections on each others' ideas and progress
16. learners 'fast-forward' their idea - what it will look like when finished

Figure 5. The structure of the Assessing Design Innovation tasks

in South African townships. Second we had to create a level playing field for the learners who had not experienced the radical curriculum of the NWPTEP. Next in addition to assessing procedural capability, the task had to provide opportunities for learners to show understanding of materials and processes, energy and power and communication technologies. Finally the starting point had to set the learners up to work in teams. Our approach to this was to create a scenario around safe transportation of medicines to rural communities in hot climates where road conditions are poor. We presented the task as a challenge for a team of six, made up of three pairs of learners, each pair taking on an element of the challenge. The structure of the task is presented in Figure 4.

These two approaches, the videos and the team challenge scenarios, allowed us to quickly transport the learners into settings where design challenges and opportunities were opened up and quickly got the learners up to speed. They also resourced the learners with understandings of issues to be addressed whilst leaving space for the learners' own ideas and experience.

Structuring the activities

The APU project set the blue print for structuring an activity through iterating active and reflective prompts to learners. This approach is illustrated here by the Assessing Design Innovation project. Figure 5 provides an overview of the six-hour activity. The 'light fantastic' brief created a model for further challenges, created by experienced teacher-examiners. As with all of our tasks, once the challenge has been introduced we encourage learners to articulate, through drawing and/or writing, whatever vague and early ideas they have. This has typically been a solo

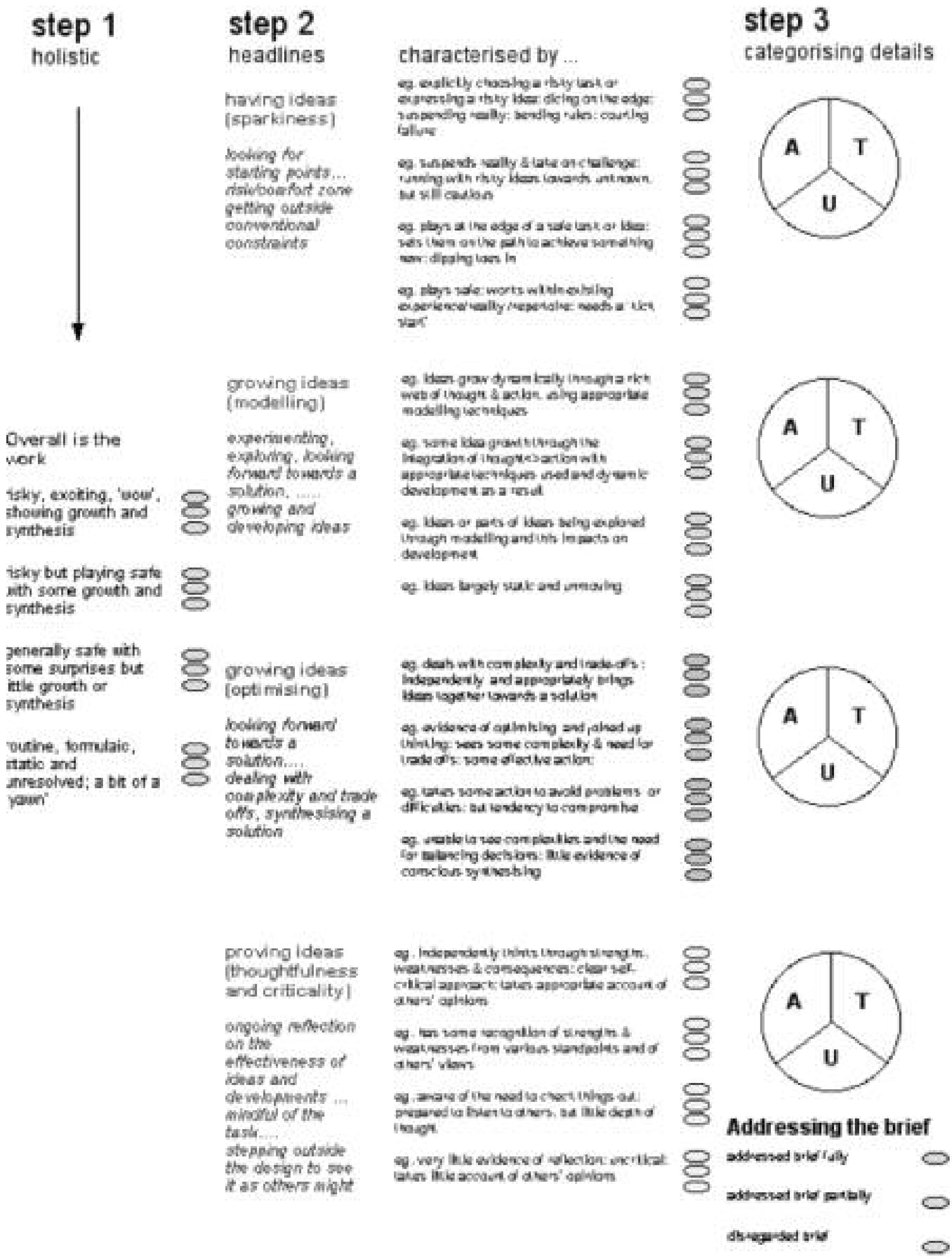


Figure 6. The assessment rubric from Assessing Design Innovation

activity, but in Assessing Design Innovation while each learner was working on their own task, they sat within a supportive "critical friend" group of three. The first interaction between the three is shown in Figure 5 as steps 3, 4 & 5 – swapping initial ideas for development & critique. Once ideas have been returned and reviewed by their original owner, individual development gets underway. Here, for the first time, we introduced 3D 'sketch' modelling. We had become acutely aware through our work with teachers that, if they were freed from the burden of assessment and asked to focus entirely on supporting creative responses, teachers provided the resources and encouragement to enable learners to engage in 3D modelling at a very early stage of designing. The response from the learners was impressive in terms of developing ideas. But as models were developed, re-worked, destroyed and re-built, the evidence of designing was being lost. Our solution was to take digital cameras and printers into the assessment space and photograph the designing six times during the activity. The images were immediately printed and returned to the learners so that a photographic storyline of their development was set down almost in real time. The total activity was broken into two sessions of three hours. Towards the end of each we prompted peer and self-assessment within each group of three. The final stage involved each learner 'fast forwarding' their ideas to show what the finished solution would be like if it were taken to production.

It was the introduction of digital cameras into the assessment activity that caused the initial shift towards developing the digital 'unpickled portfolio'. As suggested above, our primary reason for the photographic record of modelling developments was to capture evidence. What we hadn't bargained for was the significant impact the action had on the learners' design thinking. Once the first photo had been taken, they started to anticipate subsequent photos – using them as staging posts and as an impetus to push their progress. This dual value – of capturing evidence for assessment and supporting the development of the ideas – encouraged us to explore the further use of digital tools for both the assessment and development of capability.

Teamwork and collaboration

Teamwork and collaboration have featured regularly in our assessment activities – for example the use of "critical friends" illustrated through the Assessing Design Innovation project. At times collaboration has taken a supportive role only, at times it has been considered in the assessments being made. This was the case in the NWPTEPE where we wished to see the impact team working had on performance. Consequently, in addition to assessing evidence of technological procedural skills and the application of knowledge, we also sought evidence of 'team working', as characterised through "group decision making, addressing the whole task, amalgamation of ideas, supportive interaction". (Kimbell & Stables 1999, p. 7) Through the evaluation of the pilot we also found considerable further benefits of team working, particularly the positive attitudes engendered between girls and boys. (Stables & Kimbell, 2001)

Making assessment judgments

So far, the case studies have shown how evidence for assessment is generated and collected. As has been explained earlier, our approach to making assessment judgements has broadly adopted a holistic model. This has involved us first creating a rubric that provides a set of characteristics of holistic performance and then identifies the key elements within this, each of which also has a set of characteristics. The rubric used for the Assessing Design Innovation project is shown in Figure 6 as an example rubric where the emphasis is on assessing creativity and innovation.

Assessors are then asked to review the whole of a learner's portfolio and make an overarching holistic judgement, guided by the rubric and often also by exemplary material. The assessors then review the work again, looking for evidence of the first element – in this rubric, having ideas. Assessors are encouraged to look for evidence at every stage of the work, not just at the outset. Having identified the evidence, they make a judgement about the quality of the work, using the rubric descriptors as a guide. This process is repeated for each element. What is important is that, whilst closer scrutiny of the work may result in the holistic judgement being changed, it is because the assessor's understanding of the work has improved, not because a numeric relationship is made between the holistic judgement and the judgements against each element. We believe this approach provides an authentic judgement – a view that is echoed in the comments from the following teacher-assessor.

> One of the major strengths of holistic judgements I see is its flexibility…in which you can give credit to students for what they have actually done rather than whether they are able to "tick the boxes" to match a set of assessment criteria. (Kimbell et al., 2009)

This process has been taken one step further in e-scape through the use of Adaptive Comparative Judgement. This is a system that operates online, within the digital portfolio database, that identifies pairs of portfolios for assessors to judge holistically, and through multiple judgements being made on the portfolios, creates a rank order of performance. Statistically, this system is extremely reliable and is also seen as being fair.

> The judging system feels to be fair; it doesn't rely on only one person assessing a single piece of work. It removes virtually all risk of bias.... It feels safe knowing that even if you make a mistake in one judgement it won't significantly make a difference to the outcome or grade awarded to the student as other judges will also assess the same project. (Kimbell et al., 2009)

We have also found considerable 'added value' in the impact that engaging in the judging process can have on both teachers and learners. Teachers have found it valuable to look at work of multitudes of learners that they don't know, as a way of understanding how different learners have responded to the assessment challenge. For learners the process appears even more powerful. Two different settings exemplify this. The first was with undergraduate students where it provided a valid and reliable approach to peer assessment. (Seery et al, 2012) The second was a pilot within the main e-scape project, where a group of Year 10 learners who had undertaken the assessment activity, were trained to act as assessors and experience the judging process for themselves. Not only were their judgements consistent with the adults, they found the exercise highly illuminating because of the insights gained from assessing each other's work. They commented that they felt better prepared for future work. (Kimbell, 2012) This pilot opened up the potential for a more democratic approach to assessment where learners could join in the process alongside their teachers, even in the context of high stakes assessment.

The value of the approach for Southern Africa

This paper has presented an account of a particular approach to assessment - authentic assessment through performance based portfolios - that has attempted to address issues of reliability, validity and manageability within a system that is fair and equitable in the assessment of procedural capability. The approach also aims to supports the development of the learner's capability. But is it an approach that has value in the context of Southern African schools education? More so, considering the theme of this conference, does it have cultural and social relevance?

This question is best answered by those working in Maths, Science and Technology education within Southern African schools. However, I will make some comments towards exploring this area, and do so by considering the question through the lenses of validity, reliability and manageability and within my limited understanding of these challenges within Southern Africa. I am aware from reading that there are real tensions at play, for example, highly aspirational curriculum documents that are being implemented without adequate resources (World Bank, 2008).

Conceptually, the approach to assessment that I have presented provides a structure towards constructive alignment between teaching, learning and assessment – something that has equal priority in Southern African curriculum documents that promote Outcomes Based Education (OBE). However, there is evidence that achieving this is problematic.

> If the implementation of new curricula demands new forms of assessment but the implementation of assessment practices and instruments lags, the curricular changes have little or no chance to make it into the classroom. It is a common observation and result of numerous researches across SSA that the lack of alignment between curriculum intentions and assessment, and the quality of assessment and examinations remains a major obstacle for curriculum implementation at large. (World Bank, 2008, p.62)

Even when used for high stakes assessment, through the emphasis on authenticity our model takes a learner-centred approach seeking to capture evidence of genuine, procedural capability. This too has resonance with Southern African curriculum aspirations, but also appears to be in tension with actual practices, where it isn't clear that assessment of anything other than factual content is valued. Stears and Gopal (2010), highlight this issue whilst exploring alternative assessment practices in science with Year 6 learners, making the case that assessing learners through reference to the understanding that is shown through their everyday life experiences may be an important for-runner to developing and assessing knowledge of science concepts. Referring to Donald, Lazarus & Lolwana, (2002) they comment that

> Unfortunately, the value departments of education, learners and the general public attach to marks do not bode well for an approach where learners are assessed by interpreting their actions, attitudes and emotions (Stears and Gopal, 2010, p.595).

The World Bank report makes a somewhat starker statement claiming that

> Modern curricula in Sub-Saharan Africa formally aim at learning outcomes like comprehension, application of knowledge, methodological and social competencies,

> and problem solving. Current assessment and examination practices are limited to the recapitulation of memorized facts. (World Bank, 2008, p 57)

In discussing the "Teach for examination success" issue, the World Bank report (2008) highlights a further tension in the value that is placed on different assessment practices, stating "assessment and qualifications that only test for methodological and social competencies lack the achievement of clear exit skills, and have proven to lead to an "anything goes" attitude in the classroom." (World Bank, 2008, p. 58). This prejudicial attitude towards qualitative aspects of learning and assessment is not unique to Sub Saharan Africa. But it does pose a problem when considering validity in assessment practices. It also links to issues of reliability. Where reliability is linked to an expectation of right or wrong, yes or no, answers and 'clear exit skills' then qualitative judgments are viewed with suspicion. The focus on this perception of reliability in assessment appears common in the Southern African context, but again, this is in conflict with curriculum aspirations for learner centered learning and OBE. Perhaps the statistical reliability that has been shown through our approach to holistic judgement can be used here to support more qualitative practices, which are surely more socially and culturally appropriate.

A further issue that I would consider to be important and challenging is the extent to which assessment is teacher dominated. This can be seen within curriculum and assessment documents where assessment is seen as something that is 'done to' not 'with' learners. This issue has been highlighted by Beets and van Louw (2005, 2011).

Through a focus on holistic assessment and comparative judgment our research is supporting an approach which is not only learner centred but actively seeking ways of further democratizing assessment, and there are indications that this would be welcomed by educators in Southern Africa, but could be a challenging concept for policy makers.

An inescapable issue raised by our approach is the very real challenge of manageability, and particularly the importance of managing resources, including teachers' time to understand, adopt and implement new initiatives. The specific issue raised by the value we have seen of making digital resources available cannot be ignored. Again I am aware of the contrasting perspectives presented, for example by the e-Learning Africa 2012 report (Isaacs & Hollow, 2012) that provides a view of African youth as 'digital natives' and highlights the positive impact of ICT on learning, while, in contrast, the challenges highlighted in the World Bank report even of insufficient textbooks, amongst other scarce resources. One point that we have made consistently about our own approach is that pedagogy comes first; technology can then act as an enhancer. The fundamental principles and approaches we have taken are not reliant on new technologies. The world does not stand still and the challenge is to make sure that, as technologies are more available, they support rather than replace good pedagogic approaches to teaching, learning and assessment.

My comments may seem simplistic, but I feel the approach I have outlined has strong potential to support social and cultural relevance in assessment practices, even though the challenges in doing so are many. In writing this last section I can't help but reflect back on the brave, radical curriculum development that took place through the NWPTEP in the late 1990s and how this was welcomed by teachers and learners alike. The issues raised here have parallels to those raised through the NWPTEP, not least by the learners themselves who felt hugely empowered to learn through problem solving, in groups, supported rather than dictated to by teachers. These learners moved from technology classrooms into other disciplines where they demanded that teachers in their school adopted the same pedagogies that their technology teachers used, seeing these as more relevant and supportive to learning. Equally, they engaged wholeheartedly with the assessment approach that we introduced but expressed frustration that external recognition came through more the more standard approach of matriculation examinations, that excluded problem-based learning areas such as technology education.

Linking learning and assessment through activities that learners feel have relevance to their own lives and their own ways of learning transcends national and regional contexts. I hope that the experience and insights we offer from our research has added some value to ways this can be achieved.

References

Beets, P., & van Louw, T. (2005). Education transformation, assessment and ubuntu in South Africa. In Y. Waghid (Ed.), *African(a) philosophy of education: Reconstructions and deconstructions* (pp. 175-189). Stellenbosch: Department of Education Policy Studies.

Beets, P., & van Louw, T. (2011). Social justice implications of South African school assessment practices. *Africa Education Review*, 8(2), 16.

Biggs, J. (2003). Aligning teaching and assessment to curriculum objectives *Imaginative Curriculum Project*. Learning and Teaching Support Network, Generic Centre.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 26.

Buchanan, R. (1995). Wicked problems in design thinking. In V. Margolin & R. Buchanan (Eds.), *The idea of design* (pp. 3-20). Cambridge, Massachusetts; London, England: MIT Press.

Costa, A. L., & Kllick, B. (1993). Through the lens of a critical friend. *Educational Leadership*, 51(2), 49-51.

Donald, D., Lazarus, S., & Lolwana, P. (2002). *Educational psychology in social context*. South Africa: Oxford University Press.

Isaacs, D., & Hollow, D. (2012). *The eLearning Africa Report 2012*. Germany.

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

Kimbell, R., & Stables, K. (1999). *South Africa: North West Province Technology Education Project - an evaluation*. London: TERU, Goldsmiths College University of London.

Kimbell, R., & Stables, K. (2007). *Researching design learning: issues and findings from two decades of research and development*. Berlin: Springer.

Kimbell, R., Stables, K., Wheeler, T., Wozniak, A., & Kelly, A. V. (1991). *The assessment of performance in design and technology*. London: SEAC / HMSO.

Kimbell, R., Wheeler, T., Stables, K., Shepard, T., Martin, F., Davies, D., et al. (2009). *e-scape portfolio assessment: a research & development project for the Department of Children, Families and Schools, phase 3 report*. London: Goldsmiths, University of London.

Lawler, T. (1999). Exposing the gender effects of design and technology project work by comparing strategies for presenting and managing pupils' work. In P. H. Roberts & E. W. L. Norman (Eds.), *IDATER 99: International Conference on Design and Technology Educational Research and Curriculum Development*. Loughborough: Loughborough University of Technology.

Lawler, T. (2006). *Design styles and teaching styles: A longitudinal study of pupils' ways of doing designing following complementary re-grouping and teaching*. Paper presented at the TERC 2006: Values in Technology Education, Gold Coast, Australia.

Lawson, B. (2004). *What designers know*. Oxford: Elsevier.

McLaren, S. V., Stables, K., & Bain, J. (2006). *Creativity and Progression in transition through Assessment for Learning in Design and Technology Education (CAPITTAL-DT): A pilot project Report to Determined to Succeed Division.* Glasgow: University of Strathclyde.

Pollitt, A. (2012). Comparative Judgment for assessment. *International Journal of Technology and Design Education*, 22(2), 14.

Seery, N., Canty, D., & Phelan, P. (2012). The validity and value of peer assessment using adaptive comparative judgement in design driven practical education. *International Journal of Technology and Design Education*, 22(2), 21.

Stables, K. (1992). The assessment of technology at key stage 1. In C. Gipps (Ed.), *Developing assessment for the National Curriculum* (pp. 42 - 50). London: Kogan Page.

Stables, K., & Kimbell, R. (2000). The unpickled portfolio: pioneering assessment in Design and Technology. In R. Kimbell (Ed.), *Design and Technology International Millennium Conference* (pp. 195-202). London: The Design and Technology Association.

Stables, K., & Kimbell, R. (2001). Technology Education in South Africa: Evaluating an innovative pilot project. *Research in Science Education*, 31(1), 71-90.

Stables, K., Kimbell, R., & Molwane, O. (1999). Technology Education in S. Africa: An evaluation of the impact of an experimental high school curriculum with particular reference to teacher pedagogy and student group work. Paper presented at the IDATER 99: International Conference on Design and Technology Educational Research and Curriculum Development, Loughborough University.

RESEARCH

Stables, K., & Lawler, T. (2011). *Assessment in my Palm: e-scape in Israel Evaluation of Phase 1*. London: TERU: Goldsmiths, University of London.

Stears, M., & Gopal, N. (2010). Exploring alternative assessment strategies in science classrooms. *South African Journal of Education*, 30, 14.

World Bank. (2008). *World Bank working paper No. 128: Curricula, Examinations, and Assessment in Secondary Education in Sub-Saharan Africa*. Washington, DC: World Bank.

The National and the Local: Conflicting requirements in the assessment of learners' performance

Prof Richard Kimbell, Goldsmiths, University of London

Introduction for the 2015 DATA Special Edition
This paper originates in a Keynote presentation at the Technology Education Research Conference (TERC2014) Sydney, Australia. Nov 2014. It arose as an invitation from the conference team to consider the tensions that arise from the very different concerns of formative and summative assessment. Specifically in this case the organizers were aware of the digital tools that we had developed for assessing learners' performance – and that we had shown these tools and approaches to be highly reliable. They were interested to hear how these tools and approaches might fit with the culture of learning in the classroom. Are they seen as an externally imposed discipline – or do they somehow contribute to an enhancement of the culture of the classroom?

Introduction
The formal assessment of performance in schools is typically undertaken by recognized authorities in assessment. In the UK this is sometimes private Awarding Organisations like AQA or Cambridge Assessment (eg for 16+ or school-leaving certification) and sometimes it is Government bodies like the Standards and Testing Agency (e.g. for National Curriculum Assessments). Arguably this latter is more about testing schools than children – using children merely as a lever to gain some purchase on schools' performance. In either event the priorities informing these assessments will be national. The tests must be deliverable and manageable nationally and the data must produce an articulated and reliable national standard. But the vast majority of assessments in schools are made by teachers, and they typically have other, more local, classroom concerns. Of course teachers are interested in how they measure up nationally, but they principally want to know how they can help their individual students to improve their performance. What are the strengths and weaknesses of the individuals and how might the teachers modify their classroom practices to enhance their learners' performance?

The assessment of performance is one of those fields where technical national requirements (for reliability and standards) meet local cultural practices (of pedagogy and individuality). And the meeting is frequently uncomfortable and unsatisfactory. In this paper I outline an approach to assessment that reconciles local cultural practices with national requirements.

	0	1	2	3	4	5
Situation & brief						
Investigation						
Solution						
Realisation						
Testing						
Totals						

Figure 1 Assessment framework for the 1970 Certificate of Secondary Education in Design Studies

The atomization of assessment
In the 40 years from 1970-2010 the process of assessment became ever more atomized. Whilst global judgements of quality were – at one time – seen as adequate and normal, it progressively became necessary to drill down into such judgements and provide more detail. As an illustration of this, in 1970, the assessment of a student design portfolio for the 16+ Certificate of Secondary Education 'Design Studies' was based on 5 judgements (each out of 5) summed to a single figure (out of 25). In 2010 a similar portfolio submitted for a National Curriculum assessment would be subjected to analysis against 150 'Statements of Attainment' that then have to be amalgamated through a complex set of rules to arrive at a final 'NC level'.

There were two reasons for this progressive atomisation. First was the belief that by identifying *elements* of, or aspects of performance, the final judgement might be more justifiably and more reliably decided. Second was the pedagogic priority, that identifying the *elements of performance* that are praise-worthy or inadequate makes it possible to see how the teacher can help the learner to improve his/her performance. The first we might call *atomization for reliability*, and the second we might call *atomization for improvement*. (See Kimbell, 1997; Kimbell & Stables, 2007).

The articulation of assessment *criteria* became a fixation in the 1980 and 1990s; they proliferated into every facet of performance assessment. Along with the tendency went the belief that we were somehow transforming assessment from a personal view into an objective science. And this was despite the warnings of countless writers. Angoff, for example, observed that behind any criterion, there lurks a set of norms (Angoff, 1974), or Persig who argued that quality must be understood without definition; a direct experience independent of and prior to intellectual abstractions (Pirsig, 1991). Wiliam (1998) went so far as to suggest that most summative assessments were interpreted not with respect to criteria (which are ambiguous) nor with respect to norms (since precisely-defined norm groups rarely exist), but rather by reference to a shared *construct of quality* that exists in well defined communities of practice.

But against the in-rolling tide of criterion-creators, these were voices in the wilderness and the tide continued to roll in. However generously we might wish to judge the motives of those responsible for this trend towards atomisation, the *effects* of it – the outcome (at least in England and Wales) – has been utterly disastrous. And the scale of the disaster can be judged by reference to two events. In 1992/3 the burden on teachers (person hours and paperwork) of National Curriculum assessments had reached such a level that there was a completely unprecedented national boycott of all assessment by teachers. It was so absolute and so widely supported (including by heads and parents) that in the end the Minister responsible was sacked. Some modest fiddling at the edges followed, but no change of any real significance resulted. So a broken system was patched up and hobbled on. Then in 2006 the new Minister (we had 19 between 1970-2010) decided that the assessment of student portfolios of coursework (e.g. for 16+ GCSE design & technology) was so unreliable that the whole process had to be abolished. Coursework assessment was no longer an acceptable mode of examination.

I should make it clear that there were plenty of other disastrous events accompanying the atomizing trend, but I have chosen to identify these two because they provide an interesting juxtaposition with the motives underpinning the trend. Atomization for the purposes of classroom improvement generated an entirely opposite effect; an absolutely solid boycott from classroom teachers. They wanted nothing to do with it, claiming that (i) it was massively burdensome and (ii) the assessment told them nothing they didn't already know. Atomisation for the purposes of improved reliability resulted in another entirely opposite effect. It generated such chaotic unreliability that the Minister banned coursework assessment. Forty years of progressively atomized assessment created uncountable hours of hard labour for teachers and hopelessly unreliable outcome statistics. By any standards, the end result of this atomizing trend was catastrophic.

When you find yourself in a hole…stop digging. It is surely time to change direction and explore new and less flawed models of assessment.

Re-thinking assessment

In 2004 we had been awarded a new research project (e-scape) in TERU at Goldsmiths – to explore the possibility of on-line portfolios and digital assessment. It proved to be a 6 year venture through three phases amounting to an investment of approximately £1m. By the time we got the contract, the writing was already clearly on the wall about the existing discredited model of assessment (long lists of criteria, all scored and added-up by the students' own teachers) – and we understood that this new project gave us a license completely to re-think what assessment might be like in a digital world.

Within the project 15–year-old learners constructed digital portfolios of work (in design & technology, science and geography) in response to authentic, extended tasks. These tasks were conducted in normal design studios, science labs, and on geography fieldwork. In design & technology, learners designed and developed products using PDAs as digital sketchbooks, notebooks, cameras, and voice recorders (NB this was in 2005/6, well before ipads and other tablet devices became available). Their work was automatically and simultaneously sent through a wi-fi connection to a secure web-space in which their virtual web-portfolio emerged. At the end of the national trials, we had 350 design & technology portfolios, 60 in science and 60 in geography. (See Kimbell et al. 2009, Kimbell & Stables 2007)

From the outset of the project we realized that the web-based nature of the portfolios enabled us to explore a radically different approach to assessment. And the approach was informed by three big ideas.

Three big assessment ideas

1. Absolute or comparative judgement

School-based assessments typically use numbers on a scale. Judge the portfolio against this criterion on a scale of 1-20 or 1-8 (depending on its perceived importance). Assessment is on an absolute scale, and – theoretically – if I award 7/20 to student x, then that work is exactly the same standard as student y in another school where another teacher has also awarded 7/20.

But, judging on absolute scales is VERY difficult. How warm is your current room (in degrees C)? How heavy is the book you are reading (in grams)? How fast are you driving (in mph)? We typically do not hold a standard against which to measure these judgements – so unsurprisingly we are more often wrong than right.

> When someone comes to make a judgement in the everyday world, the point of reference is most often taken from past experience. Different people have different accumulations of past experience and for that reason make different judgements about the same issue. We call this difference 'a point of view'…. All judgments are comparisons of one thing with another …the judgment depends on what comparator is available.
>
> (Laming, 2004, p.17)

> When we try to judge a performance against grade descriptors we are imagining or remembering other performances and comparing the new performance to them. But these imagined performances are unlikely to be truly representative of performances of that standard, and very likely to vary in the minds of different judges.
>
> (Pollitt, 2004, p.7)

What we can do…. VERY reliably…is to weigh one thing against another. As Laming says, 'All judgments are comparisons of one thing with another' and if I am asked to compare the temperature of two adjacent rooms I can immediately tell you which is warmer even though I can't tell you the 'real' temperature. Or two books…which is heavier. Or two portfolios…which is stronger. Comparative judgement is easy and accurate.

2. Judging parts or wholes

I have already spent a while deploring the trend towards atomization in assessment, but it is worth thinking for a minute about the *reverse* of atomization. Imagine that you are a biologist presented with a new species of plant/animal that you have been asked to identify and classify. And you have been provided with a set of instruments including a microscope, a hand lens and a ruler. What would be your procedure?

I'm not a biologist, but simple common sense suggests that you start with the naked eye 'its 25 mm long with a body in two parts and 6 legs'. You might then pick up the hand lens to get a better look at what appear to be the eyes. Then you might need a microscope to see how the scaly surface of the body is composed. What I would definitely *not* do is to ask for 150 microscope slides of bits of the specimen and – on that basis – try to identify what it is.

We start from big pictures and - progressively - drill down for more detail. We do not start with a box full of details and try to build up a big picture. So why on earth do we do that when we are trying to assess a student's performance? Our first instinct should be to say 'this is a great piece of work' or 'this is really weak' and then - progressively – drill down into it to find out why.

Holistic judgement has long been understood to be important in design & technology. Indeed, in the 1988 Interim Report of the D&T National Curriculum Working Group (a year before the full report was published) they commented as follows:

> 'These considerations point to the conclusion that, because Design and Technology activity is so integrative the approach to the assessment of pupils' performance in this area should ideally be holistic'
>
> (DES, 1988, para 1.30)

It was a matter of some astonishment therefore when the full National Curriculum report was published a year later and proposed an assessment regime involving ticking boxes (or not) against 150 atomised Statements of Attainment. All the available evidence advised the reverse approach.

When we were running the APU Assessment project at Goldsmiths for the Department of Education and Science (1985-1991) we had about 120 teachers involved in the assessment of the student work that was generated in the 1988 testing programme. We asked these teachers to make an initial *holistic* judgement (on a 6 point scale) and then follow it with a series of increasingly detailed judgements of elements of the work.

> Of all the judgements markers made, they felt more confident and were more reliable when assessing holism.
>
> (Kimbell et. al., 1991, p133)

3. Sorting networks

A sorting network is a mathematical approach to sorting a sequence of numbers. Sorted data (e.g. in a computer where files are sorted by file-size or date) is much easier

Figure 2 Computer Science Unplugged YouTube 2005 (https://www.youtube.com/watch?v=30WcPnvfiKE)

to work with than unsorted data, so mathematicians have spent a long time working out protocols for sorting. It's difficult for me to explain the working of the sorting network that I'm interested in for this paper, so the best way forward is to watch a short you tube video (only 2 minutes) that outlines the approach. It shows 6 children sorting a set of numbers into order by following lines on the floor and (when they meet another) going left if their number is smaller, and right if its bigger. Magically, the numbers sort themselves into order.

Assessment is a sorting problem. We start (as a teacher) with a pile of essays and end up with a sorted pile (best to worst). Awarding Organisations start with a random mass of candidates and end up with a sorted candidate list. Once the work to be sorted can be readily distributed and accessed (which they can when its web-based) then enough people can get involved to undertake the sorting process.

Assessment in project e-scape

Having established project activities in design & technology, science and geography, and having derived the web-portfolios, we turned out attention to the problem of assessing them. And after a series of experiments we embarked upon a completely new approach to assessment that used the three big ideas outlined above (comparative judgement, holistic judgement and sorting networks). Wiliam would describe the approach as 'construct-referenced' assessment (Wiliam, 1994) in which performance is not defined in advance as a set of learning outcomes, but rather the *construct of quality* that underpins assessment judgements is sufficiently understood and shared by a community of practice.

In practice, comparative judgement requires that scripts (portfolios) are sent to judges in pairs, and the judges simply report which one is the 'better' in each pair. They make this judgement informed (in our case) by five headline criteria. But they don't judge the criteria separately. They are asked to hold these criteria in mind as they make their holistic judgement. Whilst current 'marking' approaches require only that each portfolio be scored once, comparative judgement needs each portfolio to be seen several times in different pairings.

> The essential point will be familiar to anyone grounded in the principles of Rasch models: when a judge compares two performances (using their own personal 'standard' or internalized criteria) the judge's standard cancels out. …A similar effect occurs in sport: when two contestants or teams meet the 'better' team is likely to win, whatever the absolute standard of the competition and irrespective of the expectations of any judge who might be involved. The result of comparisons of this kind is objective relative measurement.
>
> (Pollitt, 2004, p.6)

In addition to the reliability benefit of the canceling out of judges' individual bias – a related benefit was immediately clear. Conventional marking is by one marker of one portfolio (at a time). The whole process is individualized. With comparative judgement – using web technologies – it is possible to have whole teams of judges sharing their judgements about the whole sample of portfolios: a collective process that also contributes to the improvement of inter-rater reliability.

In the final phase of the e-scape project we automated this paired judgement process by developing the 'adaptive comparative judgement' (ACJ) engine, a Rasch modelling algorithm that identified the portfolio-pairs to be judged next, and which judge they should be sent to. It is an *adaptive* algorithm; it learns about the portfolios as it accumulates judges' responses. So at the outset a judge might be sent two portfolios that are randomly chosen from the sample, and if one was pretty good and one fairly weak it's an easy judgement to decide which is stronger. But gradually the engine works out an approximate rank-order for the portfolios, so it can send judges a pair of portfolios that are much closer together in the rank. This refines the rank very rapidly.

In the 2009 national trial – with 350 portfolios and 28 judges – we rapidly arrived at a rank order with a reliability statistic of 0.95. This is an astonishing statistic. Absolute reliability about a set of multi-media portfolios that portray creative designing activity by 350 learners. Never before

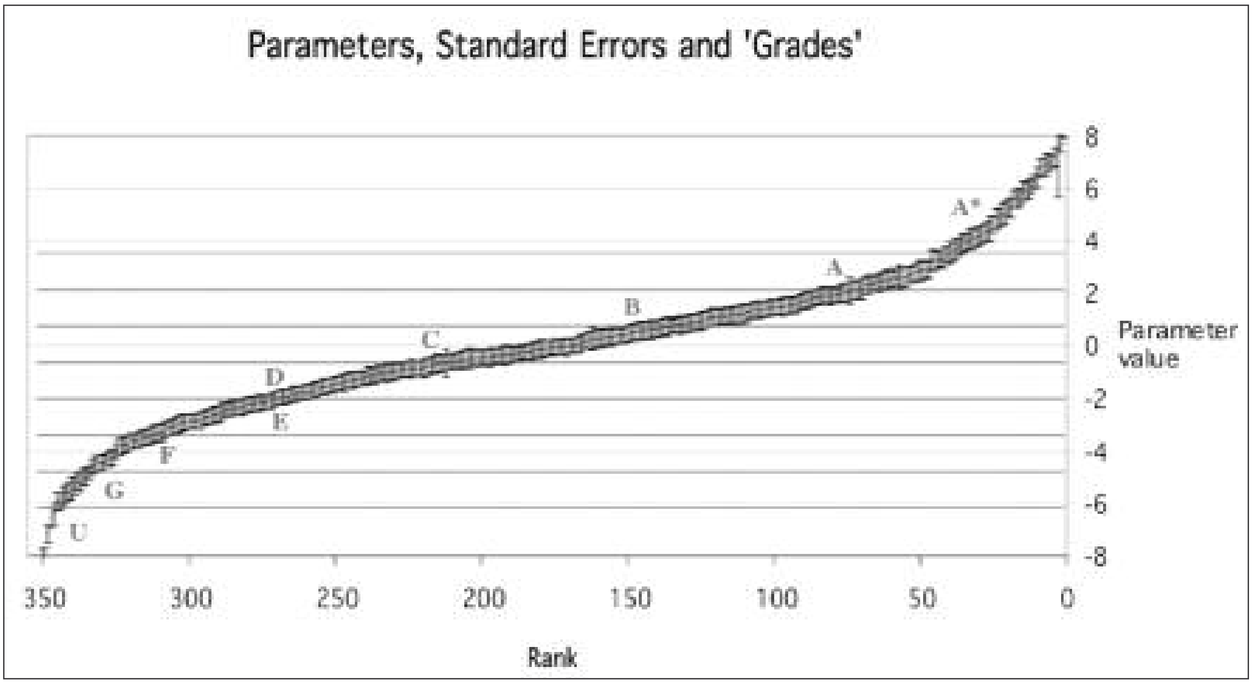


Figure 3 The rank created by Adaptive Comparative Judgement in the 2009 national trial.

has it been possible to produce this level of reliability with such data. And all the conventional paraphernalia of assessment was gone. No extended scoring sheets…no allocation of marks and painful calculation ofoverall scores…no 2nd markers and disagreements…no moderation.

All the judges had to do – in relation to each of the pairs of portfolios sent to them – was to say 'this one is better than that one'. End of story. Our teacher/judges thought it was wonderful and were delighted that they had contributed to such an astonishingly reliable outcome.

Why is this model of assessment (ACJ) so reliable?

There are at least five reasons why the assessment of performance (using comparative judgement, holistic judgement, and sorting networks) is so much more reliable than conventional marking.

1. Because it is collaborative judgement

In a normal marking context, teachers are responsible for marking the portfolios of their students. So teacher x in his school marks all the x portfolios and teacher y does the same in her school. What is the shared understanding of teachers x and y? Do they hold a common standard? Sometimes they will and sometimes they won't. And that is just not a good enough basis for deciding which students should pass and which should not.

With the comparative judgement process using ACJ, all the teachers submitting students for the examination become judges. All the portfolios are held in a big national pot (actually in a server-farm under Canary Wharf in London). So all the portfolios are mixed up together and not held at the school level. Judges are sent a pair to decide upon – and then another pair – and then another and so on. In the 2009 trial, each of the teacher/judges made approx 120 paired judgements and that was the end of the process. On average, judges took 4 mins to make a judgement, so 8 hrs in total. And most reported that it was a shorter time than they would normally have spent on marking their class of portfolios in the normal way.

Critically however the teachers were not judging their own class work – they were contributing their judgements to the whole national pot. So those teachers – for the first time – could see what the national standard of work was really like. We asked for feedback on the process….

> The judging system feels to be fair; it doesn't rely on only one person assessing a single piece of work. It removes virtually all risk of bias.... It feels safe knowing that even if you make a mistake in one judgement it won't significantly make a difference to the outcome or grade awarded to the student as other judges will also assess the same project. Also knowing that the system automatically checks the consistency of the assessor's

> judgements again reinforces the feeling of fairness that this process brings. (DW) much, much faster…less scary (re individual marker impact on individual learner life chances)…get a whole view much more readily (RW)
> (Kimbell et. al., 2009, pp.69-72)

Figure 4. Teachers marking a class set of portfolios

2. Because it is comparative not absolute judgement
For the first time in a national assessment process, teachers were not being asked to stick a number against a set of criteria. They just had to look through both portfolios – consider the basket of criteria we have trained them to identify – and then make a single holistic judgement. Is it portfolio x or portfolio y? The overriding reaction of the teachers was astonishment at how easy it all was

> Easier assessment; no need to calculate grades and points (RM)
> Speed of judging (VG)

It's worth pausing for a moment to consider the contrast with their normal process of assessment, for we were surprised how readily they took to the idea of comparative judgement. When we discussed it with them it all became clearer. Normally they start by laying out their portfolios (best-worst) on the desks in a room and then they go round the room (often in teams) filling in the forms to get the final marks. And this becomes a comparative judgement process. "We've given Julie 7 for that, and John is definitely weaker" They are using the benchmark they set for Julie as a means for deciding on the mark for John. It might look like criterion-based judgement, but it's also comparative.

And the big difference with ACJ is that what emerges is not a mass of different school-based standards, but a single national standard to which every teacher has contributed.

3. Because it is holistic not atomized judgement
The teachers where absolutely unanimous about the importance of holistic judgement – and its clear advantage over the atomized approaches with which they have become so wearily familiar.

> GCSE marking relies heavily on a tick box assessment of a pupil's work. It can be frustrating when confronted with an excellent piece of designing and making that does not meet the exam board's criteria. Too often the linear pattern of coursework requires the assessor to jump back and forth to find the marks that a student deserves. The e-scape judging is so simple in comparison. (AM)
>
> It gives more appropriate results than atomised approaches which can lead to inaccurate overall assessment especially when the overall attainment is more than the sum of the parts. This often happens when the various elements of a designing process come together in a successful outcome that outstrips the quality of work in any (or all) the parts of the process. (DP)
>
> One of the major strengths of holistic judgements I see is its flexibility…in which you can give credit to students for what they have actually done rather than whether they are able to "tick the boxes" to match a set of assessment criteria. (DW)
>
> Making holistic judgements meant that I was not forced to give credit to an apparently well-designed project that was completely unrealistic in terms of being an actual product. (VG)
> (Kimbell et al 2009 pp 69-72)

4. Because the algorithm underlying ACJ is very efficient
Given 350 portfolios and the principle of comparative judgement, one might think that every portfolio has to be compared with every other one. That is 350 x 350 judgements! In reality the algorithm does a lot of the work for us and it works on the idea that if A beats B which beats C which beats D…then A will probably also beat D. And it works out a probability for that. Imagine a matrix of 350 x 350. The boxes in it are where judgements are made (yes or no – based on which wins). And the trick with the algorithm is how many of the boxes can remain empty (those two portfolios have not been directly compared) and yet produce a reliable outcome.

In the initial ACJ prototype for the 2009 trials we worked on the notion that each portfolio would need to be compared with 20 others. But in the event – after only 16

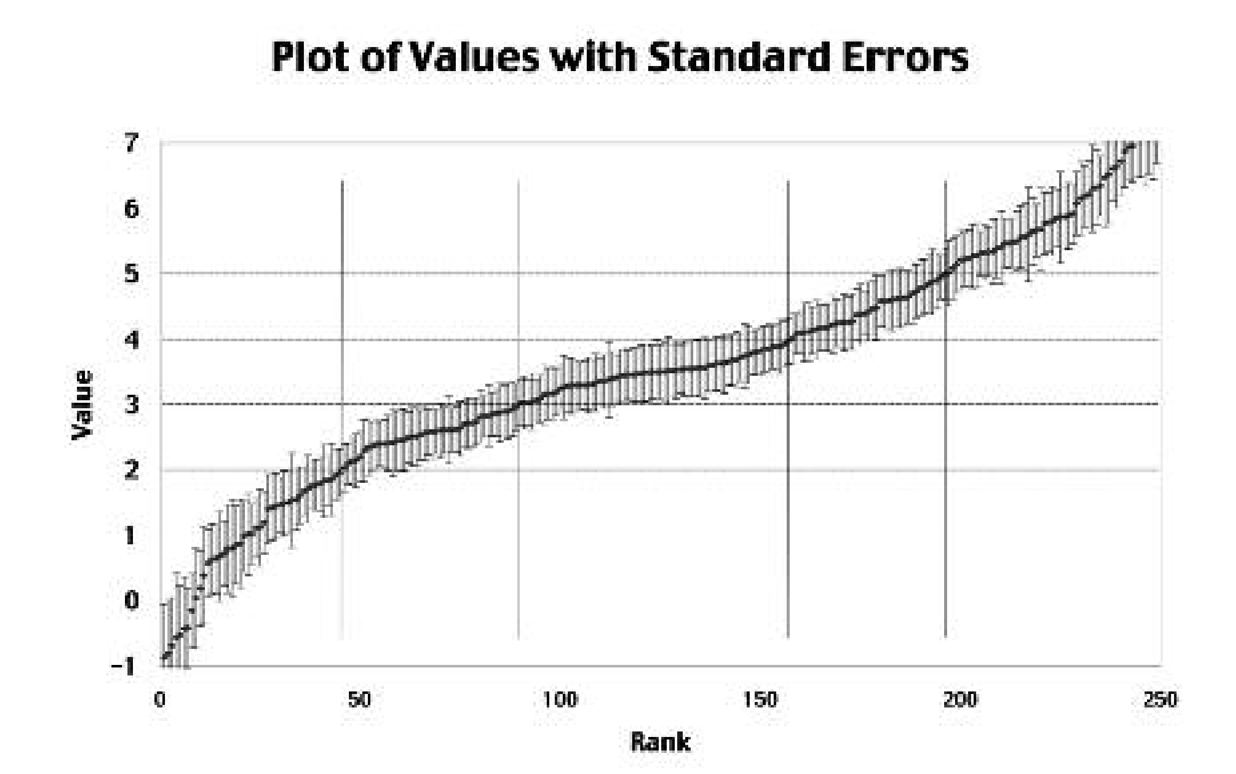


Figure 5. The standard error 'tails' show us which portfolios are causing disagreement

rounds – the rank order didn't change however much more judging we did. Then, with some further refinement, the algorithm produced a solidly reliable rank after 11 rounds. And currently (Oct 2014) it requires 10. And remember that the reliability is far higher than anything that can be generated using existing marking processes.

> ...the portfolios were measured with an uncertainty that is very small compared to the scale as a whole…. The value obtained was 0.95, which is very high in GCSE terms. Values of 0.9 or so are considered very strong evidence of validity for the test.
>
> It is worth noting that the average standard error of measurement for a portfolio was 0.668, which is just less than half the width of one of these "GCSE" grades. It is unlikely that many GCSE components – or any that are judged rather than scored quite objectively – could match this level of measurement accuracy.
>
> (Pollitt, in Kimbell et. al., 2009, p.81)

5. Because the problems encountered are made explicit
When teachers make judgements about portfolios there are two potential sources of problems for the judgement. First, the teachers may make random or inconsistent judgements and second, any given portfolio might (for some non-obvious reason) cause judges to disagree. In both cases the ACJ engine collects the data to decide what we might do about it.

All the judges generate a 'misfit' statistic that tracks the consistency of their judgements against those of the whole judging team. It might be thought of as a 'consensuality' measure. If a judge is making judgements that are way out from those of the rest of the team, we need to know – and to understand why. They may be right and the rest of us wrong – but we'll never know if we don't check. So the misfit stats accumulate for every judge and during the 2009 trial – with 28 judges – only three approached anywhere near to a misfit score that required intervention.

As for the portfolios – they too accumulate a misfit score that shows as the 'standard error' attaching to each portfolio. The blue dots that make up the blue line are the 'real' position of the portfolios, but the grey 'tails' either side indicate the size of the standard error on each. Some have bigger tails than others and if they become too big they can be pulled out and subjected to a separate moderation process.

So if the teachers and the portfolios might be sources of error – both are covered by the internal processes of the ACJ engine.

Conclusion

I suggested in the introduction to this piece that the assessment of performance is one of those fields where technical national requirements (for reliability and standards) meet local cultural practices (of pedagogy and individuality). And I suggested that the meeting is frequently uncomfortable and unsatisfactory. One has only to see the increasing number of appeals by schools against decisions by Awarding Organisations (at least, in England) to gauge the extent of the misfit between the concerns of teachers and those of national assessment agencies.

> **Sharp rise in appeals against primary school exam results**
> Rising numbers of primary schools lodged official complaints over marking in SATs tests this year amid fears children may have been given the wrong grade, it emerged today. Some 5,537 reading and maths papers were sent for review in the summer – an 88 per cent increase in just 12 months. The process resulted in 1,255 exam scripts being marked up.
> (The Telegraph 31st Oct 2013)

I promised at the start of this paper that I would outline an approach to assessment that reconciles local practices with national requirements, so it is time to make good on that promise.

In 2010, the national assessment Standard Assessment Tasks (SATs) (in England) were managed by The Qualifications and Curriculum Authority (QCA), who were already alarmed at the rising number of appeals against the judgements made in their SATs for 11 yr olds. They were particularly alarmed in the 'writing' tests where appeals had sky-rocketed. Because the QCA had been responsible for monitoring the progress of project e-scape, they knew of our work with ACJ. Indeed only the previous year (2009) we had submitted to them our phase 3 report – detailing the process and the result. So they asked whether it would be possible for us to use the ACJ methodology for the assessment of pupils' writing. When we said that we could, they provided a sample of 1,000 scripts. Each was of a piece of free story-writing (between one and two sides of A4) on a given theme. We adapted the ACJ interface to take the written text and recruited 54 primary teachers to do the judging. The result was as successful as we had expected and the teachers' response was also as predicted.

> The overall reliability of the assessment was 0.961, meaning that this assessment was considerably more reliable than any other assessment of writing that we can find reported in the national or international literature.
>
> When the judges were asked for their opinions about the method, they listed these main advantages: speed, the holistic nature of the process, increased fairness, professionalism, and a positive impact on teachers and schools.
>
> Every respondent described it as Fine, Easy, or Very easy.
>
> When asked if they would prefer to use the Comparative Judgement method or return to Marking, 25 chose Judgement, 0 chose Marking, and 2 voted for both.
> (Pollitt, Derrick and Lynch, 2010, Summary)

Moreover, in the section of the report where we invited teachers to feed back their comments, we received observations that were almost identical to those resulting from the 2009 e-scape trials.

> Each script being judged by many professionals instead of a child's fate resting on one marker
> Fairer with many assessors
> Reduces subjectivity in marking as it isn't based on just one person's opinion
> It takes the pressure off being the sole person responsible
> Allows scripts to be considered in their entirety without individual features assuming priority because of a mark scheme
> Judge the whole piece
> It feels natural and fair
> (Pollitt, Derrick and Lynch, 2010, Sect 3.1)

But I would particularly draw readers' attention to the teacher comments that centred on the professionalism of teachers and the extent to which the approach would make a positive impact on teachers and schools. There were many comments of this kind.

> Allows for professional judgement
> Uses our years/ decades of experience
> This system makes more sense – making a general judgement as to the level of a piece of work is what most teachers do anyway before they go through the criteria to prove what they think
> As a teacher, I felt I was able to make a better judgement in terms of the child's overall approach to texts and it excites me to think we could actually teach

> children the overall value of texts rather than subject them to judged deconstruction of a text.
> (Pollit, Derrick and Lynch 2010, Sect 3.1)

These comments – about professionalism, normal classroom practice, and exciting teaching opportunities – are not the kinds of responses one expects to hear from teachers just emerging from an extended bout of marking. But this was not marking. This was Wiliam's (1994) community of practice articulating their shared construct of quality.

So reconciliation is possible between good, professional, teacher expertise (the culture of the classroom) with the needs of national assessment (reliability and standards). The one does not exclude the other. And moreover I would leave readers with a final observation.

The Awarding Organisation that initially marked the writing SATs had the normal extended hierarchy of subject officers, examiners, chief examiners, moderators and senior moderators. And still they managed to produce such a suspect result that thousands of schools appealed the outcome. In our ACJ trial of the very same writing SATs, not only did all the teachers collaborate in arriving at a *common* standard, but moreover the process was judged to be professionally worthwhile for them. *And there was no hierarchy of examiners*. We had one expert analyzing the misfit statistics and checking the reliability as it emerged – but the entirety of the judging itself was in the hands of the community of practice; the classroom teachers. As it should be, since they are the people who taught the children to write their stories in the first place.

Do not underestimate the significance of this. If this democratised model of construct assessment were to be adopted nationally and internationally, it would dramatically empower classroom teachers – enabling them to develop and share their constructs of quality in learners' work. And at the same time it would equally dramatically improve the reliability of national assessments.

References

Angoff, W. H. (1974). *Criterion-referencing, norm-referencing and the SAT. College Board Review*, 92 (Summer), 2-5, 21.

DES/WO. (1988). *National Curriculum Design and Technology Working Group Interim Report*. London: Department of Education and Science and the Welsh Office.

Kimbell, R., Stables, K., Wheeler, T., Wozniak, A., & Kelly, A. V. (1991). *The assessment of performance in design and technology*. London: SEAC / HMSO.

Kimbell, R. (1997). *Assessing technology: international trends in curriculum and assessment*. Buckingham, UK: Open University Press.

Kimbell R & Stables K (2007) *Research Design Learning*, Dordrecht, Netherlands, Springer.

Kimbell, R., Wheeler, T., Stables K, Shepard T Martin M Davies D. Pollitt A Whitehouse G (2009). *e-scape portfolio assessment Phase 3 Report*. to SEAC. London: Technology Education Research Unit, Goldsmiths University of London.

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

Pirsig, R. M. (1991). *Lila: an inquiry into morals*. New York, NY: Bantam.

Pollitt, A. (2004, September). *Let's stop marking exams*. Paper presented at the IAEA Conference, Philadelphia.

Pollitt A, Derrick K, Lynch D, (2010) *Single Level Tests of KS2 Writing: the method of paired Comparative Judgement*. TAG Developments. Unit 3:11, Canterbury Ct. Brixton Rd London SW96DE

The Telegraph on-line (31st Oct 2013) *Sharp rise in appeals against primary school exam results* http://www.telegraph.co.uk/

Wiliam D (1998) *Enculturating learners into communities of practice: raising achievement through classroom assessment*. Paper presented at European Conference on Educational Research, Ljubljana, Slovenia, September 1998.

Wiliam, D. (1994). *Assessing authentic tasks: alternatives to mark-schemes*. Nordic Studies in Mathematics Education, 2(1), 48-68.

End-piece

Prof Richard Kimbell and Kay Stables, Goldsmiths, University of London

RESEARCH

Introduction for the 2015 DATA Special Edition
The majority of this paper first appeared as the 'concluding reflections' that Kay and I put together for the 2007 Springer book in which we documented the research undertaken in TERU since 1985 (Kimbell & Stables [2007] ch 15 *Researching Design Learning*. Springer). The paper is based on a personal overview of our thoughts about the 'so-what' of our research. Was it all worth it? How is the world changed? For whom is life better?

There are two additions to that 2007 material – and they both appear at the end of the piece. In the first '*The final curtain*', we outline our reasons for calling time on TERU; announcing its formal closure. In the second '…*or maybe not*' we suggest a new direction for research at Goldsmiths – building on the wider network of design researchers here who are principally engaged with the Design programmes.

Throughout any designing process – and not least at the end – there are some questions that it is always worth asking of the designer. In schools, learners are familiar with this process, including on occasions presenting their work to their peers for critical review. When we have been responsible for choreographing such sessions, the hardest question we often ask is "so what?' In fact we have elevated this somewhat crude question into a formal research tool.

"Doing a 'so-what'" is a nice shorthand way of probing into the questions that lie beneath the surface. OK so you have developed a new chair / calendar / baby-feeder. So what? How is the world changed? For whom is life better and richer? And who ends up worse off? And why? Such questions challenge budding designers to think of themselves as something more than merely developers of more stuff. We have repeatedly drawn attention to the similarity that we see between designing and researching, so perhaps we should turn our question on ourselves. 30 years of technology education research…25 of them as TERU: so what? How is the world changed? For whom is life better and richer?

We might offer all sorts of answers to this challenge, and – on reflection – they fall into three categories that, taken together, act as a satisfactory conclusion to this work. The superficial answer would be to claim that we have always managed to answer the research questions we set ourselves (or the ones that were set for us). So we now know a lot more – about designing, learning, pedagogy, capability, assessment and the rest – than we did when we started. Moreover, the fact that this research is frequently cited elsewhere might be taken as evidence that the work has some value in the educational world. But that is to take a somewhat limited view of things – not unlike the designer pointing to the new chair as if that – by itself - is sufficient justification for all the hours, weeks and months of labour. Nonetheless, we would not wish to diminish the importance of these practical extensions to the stock of knowledge and understanding that collectively informs the educational game.

Perhaps a more significant 'so-what' argument might be made for the impact we have had on others' work. One of the biggest impacts of the *APU* research that started us off, lies in the huge circle of people that were directly and indirectly touched by the project. The research team of course, but also the very eminent Steering Group and the team of civil servants who oversaw the process; the teachers who administered the tests; the team of markers and so on. We have frequently bumped into them in the subsequent years – in various parts of the world – and they often point to the significance of the experience for them as growing professionals. And what goes for *APU* has subsequently applied in equal measure to all the projects, and for all the colleagues that we have interacted with in the process. We do not exclude from this circle the learners themselves, who so often find themselves at the uncomfortable cutting edge of one of our experiments. One of the values that has driven our research and development activities has been that the outcomes should always be such as to empower and enliven learners and their lot in school. It has been one of our greatest sources of satisfaction to see these learners – sometimes the strugglers rather than the stars – enjoying themselves and growing in confidence and capability. A comment that will live with us from the *Assessing Design Innovation* project was made by a teacher in South Wales as she handed us the evaluation sheets from her group who had taken part in the first version of the 6 hour activity.

> One of the remarks that I recall from the project review sheet was "…it shows what I can do in a positive way." – this was written by a pupil who is school phobic and finds school work difficult (Teacher Database A.M-J)

Another group that has inevitably been touched by our work has been the research students we have supervised or otherwise interacted with. Sometimes they found themselves recruited as researchers, but more frequently they were used as a critical sounding board partly for our benefit but also partly so that they could view their own work through a different lens. But sounding boards are not inert – they vibrate at the same frequency as the sound and their creative vibration is sustained beyond the life of the original stimulus. They have all gone their various ways – sometimes within and sometimes not within research-like jobs – but they carry with them more than just their thesis and their beautiful robes. We would like to believe that they also carry some of the values and beliefs that have informed our work, welded to the skills and understandings that they developed through exposure to it. This second category therefore amounts to a rather bigger and more significant 'so-what', for quite beyond the substance of the research we have conducted and the findings we have published, our effects on the multiple circles of people with whom we have interacted could probably, justifiably, be described as substantial.

Which brings us to the third and final category of 'so-what'. And it is personal. Through the research projects outlined in this book we have tackled some tricky problems and dealt with some tricky clients; we have floated some whacky ideas and burned an awful lot of midnight oil to get them to work; we have argued endlessly with ourselves and with many others; we have run short of money on some projects and been grateful for the 'beer-float' that was gradually accumulating in TERU from the small surpluses on others; we have shared our ideas with others throughout the world and sought to understand its significance for them as well as for us. In the process (which has for the most-part been hugely pleasurable and satisfying) we have ourselves grown. So that is the final 'so-what'. It was deeply enriching and it was great fun.

The final curtain

So now it is 30 years since we started out with the APU project, and 25 years since we formally established TERU at Goldsmiths as the research and development vehicle through which we would operate. In that time we have won research grants in excess of £6m…and maintained only a modest beer-float. Richard has now retired from Goldsmiths and Kay is reducing her time commitment – though we both continue with some teaching, with research consultancy, and with writing.

But research is a hard, professional game that is not suited to part-timers. And for us, other attractions are waving from the wings. So after a good deal of heart-searching we have decided that the time has come to draw a line under TERU: To celebrate what has been done and to acknowledge that we have come to an appropriate end-point. So TERU will be formally wound-up at the end of this academic year.

…or maybe not

The end of TERU does not signal the end of design education research at Goldsmiths. TERU has always existed within the Design department and over the years the department has grown a number of research strands. TERU was the first research unit - predominantly concerned with design & technology as a vehicle for learning (and mainly in the context or schools) - but others have close connections. As examples, the Pi Studio (Prospect and Innovation Studio) and the Interaction Research Studio are both research units with distinct agendas but with unifying threads of innovative design practices and methodologies.

Design education is a core concern of the department, and our design research has grown significantly with the expansion of Masters and PhD programmes. And perhaps inevitably, staff and students are increasingly asking questions about ways in which design capability is best nurtured. So, as TERU takes a bow and moves off the stage, we suggest that readers stay alert to the possibilities of Goldsmiths design education research appearing in many and various forms in the future. And – being Goldsmiths – you can bet it will be interesting.

Richard and Kay

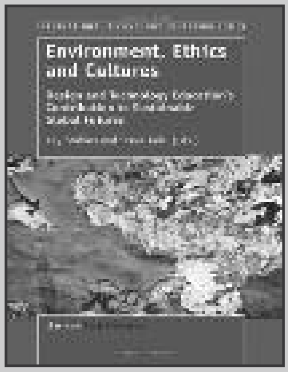
…and in case anyone would like to make contact with us about anything in this edition or maybe about other things, please use our emails that will continue to live on:

r.kimbell@gold.ac.uk
k.stables@gold.ac.uk

Review

Environment, Ethics and Cultures
Design and Technology Education's Contribution to Sustainable Global Futures

Title:	Environment, Ethics and Cultures Design and Technology Education's Contribution to Sustainable Global Futures
Authors:	Kay Stables and Steve Keirl
Publisher:	Sense Publishers
Price:	Paperback $54.00 Hardcover $99.00
ISBN: **ISBN:** **ISBN:**	978-94-6209-963-4 (paperback) 978-94-6209-937-1 (hardback) 978-94-6209-938-8 (e-book)
Reviewed by:	Professor Tim Lewis



Following the tremendous effort shown by the design and technology (D&T) education profession to maintain its position in the UK national curriculum this book is a timely reminder to re-establish the debate about the values embedded in our subject. It is a book skilfully put together by Kay Stables and Steve Keirl both eminent international education researchers and authors, however both have considerable experience of classroom teaching and this is evident in their chapters. They have brought together the work of other eminent D&T and technology international researchers, evident from the list of author biographical notes, to give us a publication with a truly international perspective. It is this variety that makes it such an engrossing book. In doing this review I have resisted the temptation to name individual authors but have used chapter titles as many are unusual and convey significant messages about the content of this book. For me it was a re-education about issues such as sustainability which are, or can be, embedded in D&T but seem to have been forgotten as we've dealt with pressing issues concerning standards, targets, assessment criteria and the like. Inevitably the style of writing varies but this adds considerable interest and readability with issues conveyed in the title dealt with appropriately but also with a smattering of humour and unusual ways of using text to get serious messages through to the reader.

Who is this book for?

The obvious answer is education researchers, particularly D&T but additionally researchers in any design discipline will find a wealth of interesting material. Additionally there is no doubt that any practicing D&T teacher will find it opens up new thinking and opportunities they may have not considered possible. However it is about D&T's contribution to wider issues which may not be seen as being part of any particular subject's remit therefore I suggest it would be beneficial reading for teachers engaged in cross curricular developments, for example teachers of science and geography. No doubt curriculum developers in schools are likely to find it helpful particularly in the newly formed UK academies with the freedom to develop their own curriculum.

I have reservations that the title may not attract this wider audience. '*Environment, Ethics and Cultures*' may appeal to those interested in these issues but my concern is that '*Design and Technology' and 'education*' are not sufficiently prominent to ensure D&T teachers in particular and educationalists in general are attracted. For example those teaching in university design departments may not see that it could be of interest to them. A further point is that the UK Engineering Council's Accreditation of Higher Education Programmes: UK Standard for Professional Engineering Competence includes the '*Economic, legal, social, ethical and environmental' context of engineering and there are several chapters that could be useful to engineering lecturers*. Hopefully the publishers will promote this book as being of importance to all educators concerned with design, technology and manufacturing and so capture this wider audience

Format of the book

Section 1 starts with an excellent introduction by the editors followed by three chapters dealing with what are best described as broad based design issues set within a global perspective, several with hints of a political

Review

Environment, Ethics and Cultures Design and Technology Education's Contribution to Sustainable Global Futures

overview. However it was the introduction that set the pattern for the way I read this book. It introduces the work of contributing authors in such an enthusiastic way that I found myself turning to and reading chapters written by these authors. I'm sure the editors did not intend this to happen but who could resist turning to a chapter with the title 'KATOGRIFA IN-FLUX: A PEDAGOGICAL TOOL TO CHALLENGE EUROCENTRISM IN POST-COMPULSORY EDUCATION FOR SUSTAINABLE DESIGN' – I really did want to know what this was about and the author documents a particularly enlightening research methodology carefully implemented with clear discussion and conclusion. I wasn't disappointed. Similarly when I turned to the chapter titled '*IN(DI)GENEITY IN DESIGN AND TECHNOLOGY EDUCATION: ANIMATING AND ECOLOGICAL CROSS-CULTURAL CONVERSATION*' I was alarmed to find myself reading a sort of cartoon script but quickly realised this was unusually creative way of conveying very serious messages about issues in the title. I found it humorous and entertaining as well as educational. So the scene was set - I jumped about this book selecting each chapter by the attractiveness of titles. It became a sort of 'coffee table' book to be picked up and read in short doses. On reflection this was entirely appropriate as I found that to fully appreciate each chapter it was necessary to reflect on the messages being conveyed by that author. Read it and think about it.

Throughout the book authors are vigilant in referencing their work and each chapter concludes with comprehensive references. The breadth of international material accessed is impressive as several authors extend their research into government and international documents as well as texts beyond those usually found in education focused books. The total of these references can only be seen as an exceptional resource for future researchers.

Section 2 focuses on global D&T education issues but within the context of education where the learner is at the centre with ownership of their education. For me this was a real eye opener as I've become steeped in UK D&T and not fully realised that D&T in other parts of the world may be very different. For those D&T people concerned about making as part of designing there is a delightful chapter headed 'AGENCY AND UNDERSTANDING' where the author uses a heading 'MAKING AND BEING HUMAN' with a quote that '*development of our brain was crucial, but, without the capabilities of the hand, the brain was an agent without an actor*.' This is followed by a short discussion of the work by of A. N. Whitehead an eminent philosopher from the 1920'/30's. Whitehead's work has informed researchers since then so inevitably I diverted my attention to re-read some of his essays (available on the internet), many of his statements being particularly relevant in current educational debates. Worth looking at. Again an element of humour creeps in when the same author asks the question '*Have we been providing the 'McDonald's' version of D&T experience*' – I'll leave you to discover the authors answers to this! In the same section another researcher discussing the D&T curriculum identifies '*what the curriculum is not*' and then uses a heading '*DESIGN AND TECHNOLOGY'S CURRICULUM PLAY*' followed by several thought provoking questions that sets the scene for discussion exposing possible answers. Initially I found myself at odds with this author concerning discussion about a body of knowledge for D&T and his concerns about the' *hreat to quality D&T by the so-called STEM agenda*'. Surely there is an opportunity for science and D&T to collaborate on issues such as sustainability! Reading on however I found myself agreeing with many of the points made and came to the conclusion that it is about getting the D&T curriculum balance right. This is exactly what this type of book is about – exposing the reader to different points of view so they can formulate their own informed conclusions.

Section 3 is my favourite as I like to read about what D&T teachers are doing in their classrooms and workshops. The authors provide a varied collection of research and development case studies from around the world. Again I was drawn to chapters with unusual titles – how can any D&T educationalist not home in on titles such as "*WE HAVE TO CREATE A WAY TO CATCH FLASHES IN ORDER TO GET ELECTRICITY*" - all about pupils perceptions and ideas about climate change. This author's conclusion and proposals are prefaced with phrases starting with '*Helping children to*' an example being '*Helping children to develop creative ideas*.' Several innovative pedagogical practices are documented. '*SUSTAINABILTY + FUN = A CHANGE IN BEHAVIOUR*' is another example of a researcher in school using a case study approach to develop and trial '*Project F*' (F for fun) based on sustainability and re-cycling issues. For teachers interested in the more entrepreneurial, commercial and consumer aspects of D&T the chapter titled '*THE SHOE SHOW*' documents a D&T project with sufficient detail that, in the hands of a skilled teacher, could be replicated. The discussion about methodology indicates that this could be adapted for other products and projects. The international nature of this book is

REVIEWS

exemplified by a chapter setting out the case of D&T in Botswana. This author points out that D&T based on the culture, history and philosophies of the Euro-Western world does not necessarily embrace sustainability and concludes that indigenous knowledge, materials and technologies should be explored and embedded in D&T education in Botswana. The author's conclusion is that D&T education should be focused on sustainable development. This message is reinforced by other contributing authors and effectively broadens what international D&T education is about. This resulted in me reflecting on what we have been doing in the UK and how we possibly need to reappraise our situation. This theme is explored further in the chapter '*OPENING UP FOUR WALLS*'. Again this author uses the word '*fun*' in a D&T context and I immediately found myself at one with this author as he explains that the basis of the case studies presented' *was a belief that learning how to learn, which questions to ask and where to seek appropriate information are important skills that need to be explicitly taught and nurtured through teaching programs, as is the capacity to learn from and with others.*' In his description of the pupils' activities he uses the word '*critiquing*' as a design methodology, a word not regularly used in school D&T in the UK. Perhaps we should look at adopting this and exploring the possibilities. This well detailed chapter is supported by photographs unfortunately not well reproduced making it difficult to appreciate the outcomes of this study.

Throughout this book authors consistently use the word 'learner', with one author producing a two column table (page 95) headed '*Learning targets and expected outcomes*'. A second column headed '*Competence/Does-education enhance learners' capacity for*:' has entries such as '*Learning to learn, Learning to do, Learning to be*' - plus several more. The '*Expected outcomes*' column is full of educational outcomes that bring together what D&T education is all about. Have a look at this table and I'm sure you will be impressed.

Conclusion
Reading this book definitely put me in the position of being 'a learner' – I can assure you that I learned a great deal that certainly enhanced my understanding of the breadth of D&T as an international subject and raised my awareness of opportunities that I have not fully appreciated.

Index of Research Articles by Titles

Volume 20

Index of Research Articles by Titles

Volume 20

Guidelines for Contributors

Journal overview

The mission of the new international Journal is to publish high quality research, scholarly and review papers relating to design and technology education. There will be three issues each year and five or six substantial articles in each issue. Submissions are welcomed relating to the primary, secondary and higher education sectors, initial teacher education (ITE) and continuous professional development (CPD). Contributions to the on-going research debate are encouraged from any country. The expectation is that the new Journal will publish articles at the leading edge of the worldwide development of the subject area. The final edition each year will include published versions of the keynotes from the D&T Association's International Research Conference.

The normal word limit for articles is 5000 words, although up to 8000 words will be permitted in exceptional circumstances. Visual illustrations are encouraged in keeping with design and technology's ethos and practice. It is the Journal's policy to positively encourage the submission of articles based on action research by practitioners, which has been the bedrock of the subject's development for several decades. The Journal would also welcome the opportunity to publish substantial literature reviews in order to consolidate contributions which have been made to the subject, and ensure that they are accessible to current researchers and teachers.

Articles for consideration by the Editorial Board should be emailed to neil@data.org.uk. If you wish to submit an article by post, please supply one paper copy plus an electronic copy on disc. Please address to: The Design and Technology Association, 16 Wellesbourne House, Walton Road, Wellesbourne, Warwickshire, CV35 9JB. Some detailed notes on the preparation of research articles are provided at the end of these guidelines.

The articles will be 'blind' refereed by three members of the Editorial Board, who will remain anonymous, and authors will receive feedback through the Editor.

Editorial Board

The Journal is edited by Prof Eddie Norman of Loughborough University (E.W.Norman@lboro.ac.uk) with the support of Co-Editor Dr David Spendlove from the University of Manchester (david.spendlove@man.ac.uk). Potential contributors are welcome to contact them in order to discuss issues they may have.

Developing conference papers for submission to the Journal

The conference paper, such as one written for the D&T Association International Research Conference, would have been subject to a word count restriction (typically 2500 words) in order to ensure that it is possible to present its content in 15-20 minutes. Clearly, this constraint is removed for a Journal article, and replaced by a larger word limit. The word count restriction on a conference paper normally means that the agenda it addresses has been correspondingly restricted and the paper has perhaps also been deliberately slanted towards the conference theme. For a Journal article, it might be possible to put the research more firmly into context by looking more extensively at prior research by others on the topic, or by looking at wider implications of the research work completed (beyond the conference theme). This might involve some further literature-based research, but it does not necessarily mean that any further data collection needs to be carried out. Conference delegates might have suggested other authors whose work the researcher should review.

A further area that can be developed within a Journal article is a discussion of the research methods employed. It is possible that other methods were considered and rejected (but not reported). The methods chosen might have known strengths and weaknesses which have not been fully explored in the conference paper. It may be possible to make comparisons with studies which use similar research methods, but which were not discussed in the conference paper, perhaps because they were too loosely related to the conference theme. Again, conference delegates may have suggested other studies with which the author might compare their work.

It is likely that the discussion at the conference will, in any case, have moved the research thinking on and this can be captured, or the data re-examined in the light of new suggestions.

The intention is obviously not to republish the conference paper, but to consider the possibility of publishing a related 'deeper, richer' account of the developing research. Conference papers are often thought of as 'stepping stones' towards more thorough research accounts, and perhaps the process is best thought of in that light. The author could consider how the subject matter has revealed itself during the process, their aims and how close they are to realising them. It is hoped that the conference paper presentation and subsequent discussion with delegates might contribute some new understanding.

Developing research assignments for submission to the Journal

Many teachers undertake CPD programmes that incorporate research elements such as literature reviews or action research studies. Some teachers also undertake MPhil and PhD degrees or other higher qualifications, such as EdD. Although assignments undertaken for such programmes will initially be formatted in accordance with the submission requirements, such postgraduate students might consider the possibility of submitting their work to a research conference or journal. A poster or PowerPoint presentation at the D&T Association's International Research Conference provides the most straightforward first step to getting feedback on your work from the research community, however conference papers have been successfully written based on such action research programmes (e.g. Alison Hardy's Questioning Styles: observations of differences in practice at Key Stage 2 and Key Stage 3 at the 2004 International Research Conference. A copy of this paper can be downloaded from the members section of the D&T Association website: www.data.org.uk).

Of course, it might also be appropriate to submit an article to the Journal based on the research undertaken and such submissions would be welcomed.

Some detailed notes on research articles

Research articles would be expected to make an original contribution to design and technology education research. Such contributions should be based on evidence (e.g. newly acquired data, historical records, published work).

The article should have a title, name(s) of author(s), their titles and affiliations. It should have an abstract of between 200 and 250 words. It should have six keywords for reference purposes. The paper should not be more than 5000 words in length, unless there are exceptional circumstances. All submissions must be typed in good English, they must have been spell-checked and include a word count (excluding the abstract and title). All pages must be numbered and double-spaced in order to facilitate the refereeing process. All references should use the Harvard Method of referencing and details of this are given in the following section. Any information relating to authorship including affiliations should be confined to a removable front page and should be free of clues such as self citations, e.g: 'in our previous work…'

Guidelines for Contributors

Example of an Abstract
The teaching and learning of technology for design

Eddie Norman, Senior Lecturer, Department of Design and Technology, Loughborough University,
Co-Director IDATER

Abstract
This paper presents a non-linear model of design and technology that illustrates the concept of technology for design (or technology for the purposes of those engaged in designing). The model shows technology for design as the summation of the knowledge, skills and values employed in design decision-making. Technology's relationship with science is discussed and research evidence concerning the emergence of new technology for a polymer acoustic guitar is described. This is a fully-documented case study, which demonstrates the existence of knowledge, skills and values that are derived from designing and making rather than science. Learning 'by doing' and teaching 'by showing' and their pedagogical implications are discussed. Examples are given from the author's teaching of undergraduates, which demonstrate the teaching of technology for design through designing i.e. where design activities are a teaching and learning strategy (e.g. injection moulding and the use of recycled materials). Examples are also given of technologies for which research evidence has indicated that there is a need for them to be taught prior to designing if they are going to be employed effectively (e.g. structures). These ideas are further illustrated using examples from a resource pack on kite design and technology for KS3. The paper shows the importance of sustaining designing and making as a teaching and learning strategy in order to promote innovative and creative designs in the next century.
(226 words)

Key words
models, technology, design, teaching, learning, guitar

The Harvard method of referencing
The work is referred to in the text by stating its author's name and the year e.g. (Pacey, 1983). If there are two authors, then both their names should be given e.g. (Roberts and Archer, 1979). If there are more than two authors, the surname of the first author should be given, followed by 'et al' e.g. (Benson et al, 2002) If there is a need to be more specific, then the page number should be stated immediately after the year e.g. (Mockford and Torrens, 1997:164). Some examples of references using the Harvard method are given below.

Books: the author(s) and year of publication should be given followed by the publisher e.g. Eggleston, John (ed) (2000), *Teaching and learning design and technology: a guide to recent research and its applications*, Continuum, London.
Journals: the author(s) and title of the article should be given followed by the title of the journal (in italic), the volume, issue number and pages e.g. Spendlove, David (2001) 'Gender issues: assessing boys as underachievers'. *The Journal of Design and Technology Education*, 6, 3, 202-206.

Papers within conference proceedings or collected work with editor: the title of the proceedings or the whole book is in italic followed by the publisher e.g. Norman, Eddie (2000) 'The teaching of technology for design'. In Richard Kimbell (ed), *Design and Technology International Millennium Conference 2000*, The Design and Technology Association, Wellesbourne, 128-134.

Hope, Gill (2001), 'Participant research in design and technology education from the perspective of a design and technology participant researcher'. In E.W.L.Norman and P.H.Roberts (eds), *IDATER 2001*, Department of Design and Technology, Loughborough University, 47-51.

When referring to more than one document published by an author in the same year, these are distinguished by adding lower case letters (a, b, c...). At the end of your paper the references should be given in alphabetical order ('a' will refer to the first source quoted, 'b' to the second etc.).

Deadline dates for Volumes 21.1, 21.2 & 21.3

Submissions to Design and Technology Education: An International Journal are accepted throughout the year, however, the copy deadline dates for Volume 20.1-20.3 (2015) are as follows:

Volume No.	Submissions deadline	Publication date
21.1	*2nd December 2015*	*February 2016*
21.2	*20th March 2016*	*June 2016*
21.3	*7th August 2016*	*October 2016*

Response to papers
The key purpose of the journal is to support the on-going conversation which will be the foundation of future progress in design and technology education. To this end, the Editors will consider publishing in subsequent issues, short responses to any of the points raised in this issue. However, please bear in mind that more substantial responses would be better submitted as papers for publication in the journal in their own right.