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.

RESEARCH

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.

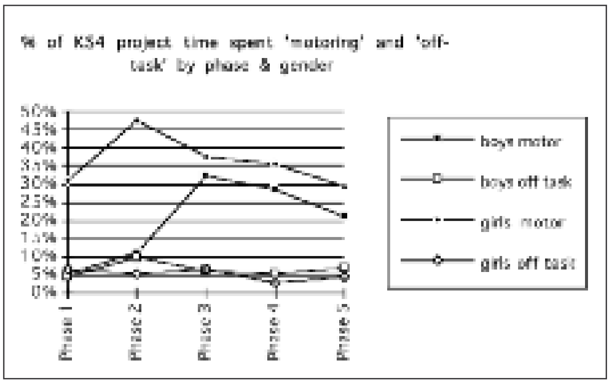

Chart 1 Differences in engagement across projects between girls and boys in KS4

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.

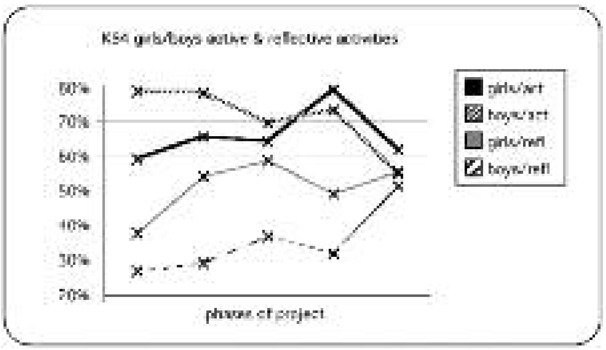

Chart 2. Differences in active and reflective modes of engagement between girls and boys in KS4

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.

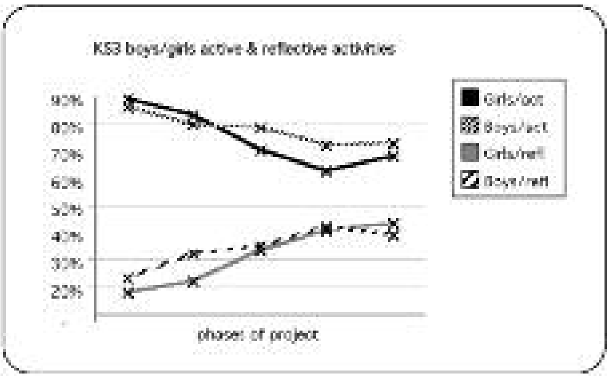

Chart 3 Differences in active and reflective modes of engagement between girls and boys in KS3

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.

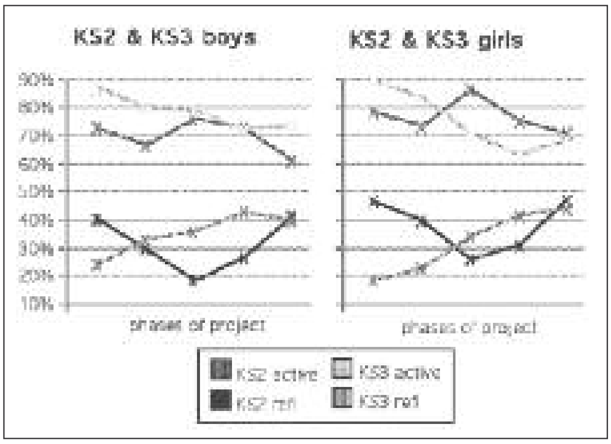

Chart 4 Differences in active and reflective modes of engagement between girls and boys in KS2

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.

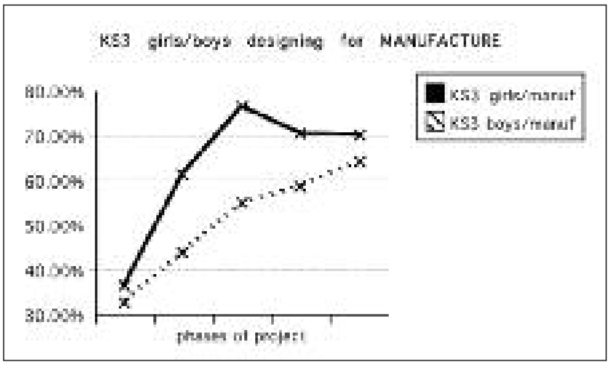

Chart 5 Differences between girls and boys focus on manufacture in KS3

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%).

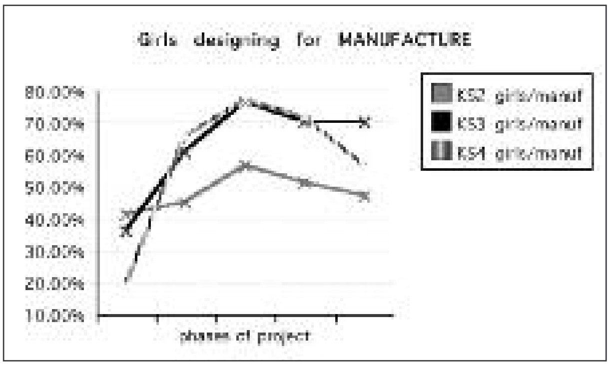

Chart 6 Differences between girls' focus on manufacture at KS2, KS3 & KS4

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

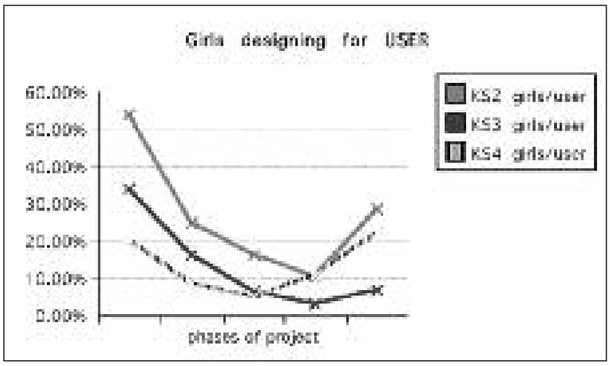

Chart 7 Differences between girls' focus on the user at KS2, KS3 & KS4

> 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

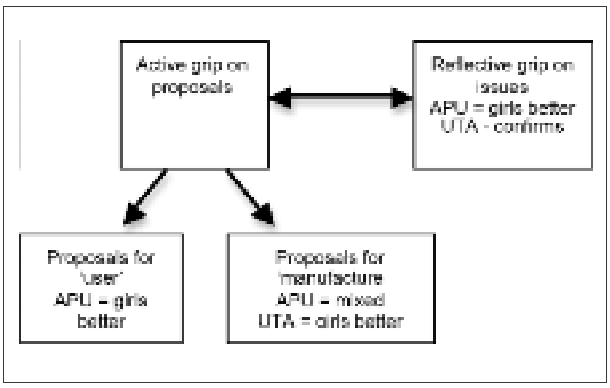

Figure 1 Gendered relationship between action and reflection, as shown by APU & UTA data

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

RESEARCH

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