

## Futurelab conference transcript

### **Beyond the Exam: Innovative Approaches to Learning and Assessment**

19-20 November 2003

Watershed Media Centre, Bristol

#### **Performance assessment: assessing the inaccessible**

Professor Richard Kimbell, Goldsmiths College London

##### **Richard:**

Good evening and welcome from Goldsmiths College. This morning, from Paul Black, you got the serious bit of London University. Goldsmiths College is the lunatic fringe of London University. I am from there. Goldsmiths College is where Mary Quant first designed her boobey traps, where Damien Hurst learnt to stuff animals and Tracey Emin learned to make her bed. So to give you a flavour of what Goldsmiths is like; it's a very creative place, a very exciting and creative place. But we also have a certain discipline and my concern today is coming from the design department, and specifically I'm interested in what happens when kids design things in schools, in the subject that we in the National Curriculum call design and technology, and how we assess that performance, and the problem with it is that designing is essentially a process and a lot of that process is inaccessible. So because it's inaccessible it raises interesting problems for assessment. That's what we're going to talk about.

The key to performance assessment is performance evidence and typically for it to be good assessment the evidence has to be tangible, shareable, comparable, assessable and believable. That raises some really serious challenges in the world of designing, because designer technology is defined as a process of design and development, and when you go to the distinctive contribution statement from the QCA, what it says about design and technology, amongst other things, is that it prepares pupils to intervene creatively to improve the quality of life, to become autonomous and creative problem-solvers as individuals and members of a team. So in design and technology all pupils can become discriminating and informing users and can become innovators. That is a hell of a mission statement to try and do that lot.

So what is the evidence that's going to count for assessment purposes - and, typically, the assessment that we use is the assessment of performance on task. We give students tasks, we watch them do it, they record their work, develop their work, make their pieces and that's what we look at for assessment purposes. We tend not to use examinations. We certainly don't use tests. We tend to use performance evidence, and also, of course, reflections. The students' reflections on their performance as evidence. This area is typically housed in a portfolio, which is typically drawings with writings and jottings, that reflects upon the process as it evolves. So that's the kind of package of evidence that we're looking at typically.

I can show you loads of interesting projects with kids doing things at various Key Stages. The Key Stage 1 classroom is exhausting, I can tell you. Kids are going ninety to the dozen and designer technology tends to emerge from other areas of curriculum, particularly from stories. This is a project that came from *The King Who Has No Clothes*, and so the kids said, well we ought to make him some clothes. So this pair decided to make him some shoes. Now that is a fairly serious challenge, making a pair of shoes from nothing. So these kids set about making a pair of shoes and you can see what they're like, bottom right there. They were real and they wore them. the kids wore them around the class and they were very proud of the level of the shoes that they'd managed to create.

Key Stage 2 tends to be a bit more technical. The idea of electric motors, mechanical movements and in this case, Coke cans, and you can see this tank idea of this trundler thing that trundles up any hill, however steep, and if it gets too steep it just flips over and keeps going, so it's a nice little technical project.

Key Stage 3, this is a project that we did with a group in a very short time and this particular example was developed using CAD. The project was based in an art gallery and the owners of the art gallery wanted little light portable seating units that people could pick up as they came in the door they and take around the gallery with them, so if they wanted to sit and look at something, they could put the seat down and watch. This was a solution developed, at least to model form, by one student and you can see it's a very elegant solution to the idea of a collapsing seat developed using a CAD system.

Key Stage 4, this chap kept fish but wanted to go on holiday without the fish dying, so he designed that system in order to feed the fish on a regular basis. So in there is a kind of revolving cylinder with a little notch out of it, and in there you put all the food and as the notch goes round once a day it collects some food at the top and then it drops it out the bottom and then comes back; it goes for a fortnight, so he can go on holiday without his fish dying.

That kind of project is not untypical of those Key Stages and supplementing it all is the portfolio of drawings, models and all the things that explain the process. Now there's a long tradition of portfolio drawings and these are little snippets that I picked up from the original drawings of Edison and Bell when they were trying to develop the telephone. I'm sorry about the quality of this drawing but there's a wonderful concept model there of how the speech gets echoed into this chamber here, which makes that membrane vibrate, which puts a needle up and down into this liquid, which changes the voltage around this. It's the original concept model of the telephone.

Now the important point about these drawings - you can see, obviously, these are much later on drawings, that's another fairly early one, but these are much later and much more developed drawings - the point about those drawings, Gorman and Castle were two psychologists who examined this archive of stuff and they said the process of innovation is characterised as a recursive activity in which inventors move back and forth between ideas and objects. They have an idea, they reflect it in objects; they have another idea and they reflect it in the object, and that was their interpretation of this archive of materials and, in fact, our view of that modified it somewhat because it seems to me that that definition doesn't take into account the drawing.

What Edison and Bell are doing here. this is not just a mental model, it's a drawing of a mental model. So yes, they're making these as mechanical objects in the end but they're initially making them as drawings and it seems to us that the key thing about the portfolio is that the act of expression pushes ideas forward. Having drawn that, which was an initial concept, the very drawing of it enables Edison and Bell then to take the next step, to say "I wonder what would happen if we did this?" So the drawing itself enables us not only to see our ideas but also to be clear what they are and what their consequences might be, and this is the important thing about design portfolios.

But when design portfolios come up against assessment there is a different kind of problem. What we have done, particularly at GCSE, AS and A2, is we have formularised the design portfolio. We've said, it goes through these stages, it starts by doing this and then it does that, and then does that, and then it does that, and there are marks given for this bit and that bit, and the other bit and the next bit, and kids learn to jump through these hoops and it becomes a very artificial process. If I could just demonstrate this for you. The idea that ideas come in the middle here is complete nonsense. Ideas drive designing. I'm going to give you a design project and you're going to design it in ten seconds. Here is your problem. Alice is 8 years old and loves riding horses, and she's fallen off her horse and she's injured her back and she's lying in traction in a hospital bed, flat out in a hospital bed for three months. She loves doing jigsaw puzzles. How can we design something that helps Alice, lying on her back for the next three months, to do jigsaw puzzles? You could have a magnetic jigsaw puzzle, you could use Velcro and you could use mirrors. You could have a software jigsaw, could you not?

Can you see how, within ten seconds of giving you that task, you're having ideas? You're not allowed to have ideas yet; you haven't done your brief and you certainly haven't done any research. What I'm trying to demonstrate is that this kind of formulaic description, which is what the examination boards and QCA require for designer technology, is complete bollocks. It's not how designers work, but we've been forced into this position for the purposes of assessment. Right, what are we going to do about this, because the result of this is that too frequently, particularly GCSE, AS and A2, the portfolio is typically not creative, it's formulaic; it's typically not done in teams, it's done individually; it's typically not innovative, it's safe. It's high stakes assessment and teachers will not take chances with kids in high stakes assessment for very good reasons. I'm not getting at teachers here. It's not happening because teachers are idle but because we're systematically rewarding students for producing work that is formulaic, individual and safe. That's what the criteria almost encourage. We reward those qualities and teachers can't be blamed for making sure that students produce them.

So for example, let me give you another illustration of the nonsense. One of the criteria somewhere says because it's good to relate kids designing to the real world of manufacturing, you've got to send off a letter to a manufacturer and make sure you get a reply back from them to put in your portfolio, and the kids might very reasonably say, what do I need to write about? It doesn't matter about that, just make sure you get your letter to a manufacturer and make sure you get a reply, because that counts as evidence of consulting the real world. So we kind of artificialise the process of designing in order to assess it and then we end up assessing something that's not real designing. It's a weird world.

I'm going to show you a case study where we quite deliberately broke the rules. I finally persuaded a few people at DfES and QCA that this was the problem and that we are systematically rewarding the wrong things for the very best of reasons and they oddly enough gave me some money and said OK, go away and solve this problem. We said, what would happen if we set up some design challenges in schools in such a way that allowed teachers to be genuinely creative in ways that they know how to be, and took away the constraints of the formal portfolio.

Here is one such case study, and one important point about this is that it was done in two days, from start to finish; two whole days, five hours a day, ten hours work altogether. Normally GCSE projects will last more than six months, often for a lot of formulaic nonsense. This was two days. It's a textiles setting, the task set by the teacher was to design a protective summer hat - this was in June we did this, June of this year - the problem of sunburn and all that, and we're going to be following Faye. Faye was one of the girls and they had a group that was for consultation purposes but basically the work was individual.

So Faye starts with some early discussions in the group about sunburn and hats, and you can see quite early on there's a kind of watery feel to a lot of this. There's seaweedy stuff and seashells. so the idea of water emerges quite early on, and she comes up with this concept, because they had a conversation in their group and the conversation was along the lines of what happens when you get sunburn at home, and one of their group said, "well my Mum's got this stuff in the freezer" - what they meant was one of the freezer blocks - "she took it out the freezer, wrapped it in a towel and put it on my sunburn". So that was a starting point. Faye thinks to herself, maybe I can take that idea and develop it into a hat. So there is the concept, which is two sheets of plastic glue-gummed together leaving a cavity in the middle, which can be filled with water. It then changes and develops into what becomes the top of a top hat, which is re-made, filled with water. subsequently re-made again; this is the third one in about ten minutes and made not only just to be filled with water but to look like water, and even in there you can see it's got goldfish as well. Just to be completely persuasive about the fact that it's water and that there's some real goldfish in there. There are also a few other bits of wildlife lurking in the depths of this

thing.

But then, if that's the top, it's got to be held in some kind of structure, and she then starts thinking about the brim of this top hat thing and works through that. Then she has to link the brim to what is effectively the crown with a section around the middle, which, again, has a watery film. you can see she got involved in this project. There was no messing about, she got stuck in, and it's got a very watery feel to that structure as well and in the end she then had to assemble it, and she's got three components; the brim, the sides and the crown bit. How do you fit that lot together, and it's a real problem. So lots of glue-gum comes in handy; glue-gum fixes everything, and what she ends up with is a prototype and this prototype hat lives in the freezer so that when it's a really hot day you can go and get your hat out of the freezer. Isn't that an interesting idea, and you can plonk it on and it'll keep you cool for hours until the heat gradually softens all the stuff that's in there.

So what we had there is a concept, an idea, that led to a prototype and it's what - there's a fabulous website called ido.com, IDO is a design consultancy in London, well all over the world actually - and they talk about their innovation process and they describe it as fast, furious prototyping. They get an idea, they try and make it, then they try and make it again and then they try and make it again. The point is that what Faye did here was produce this prototype, which is not a bad working representation of what the thing would become, she did it in ten hours, start to finish in two days.

Now that could be left at that point. You could leave it there or it could be the jumping off point for a development project in which the whole thing became much more sophisticated and much more developed and much better made, and all those kinds of things. You could take it either way, but whatever you have to say about that process is it was real designing, real engagement with the task, real ideas driving it. So how does it measure up to the assessment problem? What we've established first of all is that innovation doesn't lie in these kind of methodological steps of any process. It lies in ideas that drive designing, students' designing. Originally that just came from a casual conversation between two of the girls about this freezer block.

So what have we got for assessment purposes? We did this experiment not just in this school; we did it in lots of schools and we got evidence that was assessable, that enabled us to get hold of the ideas that students were engaging with in their product development, and we were able to put them on a scale, and the scale, broadly speaking, is a scale of innovation. We were interested in the risk of the innovative and it was remarkably reliable to be able to ask a group of people who knew what they were looking at to put things on a scale. We could go further than that and we could break the ideas into three kinds of categories. We could look at having ideas - just generating the idea is one thing, growing the idea is a different thing again. The process of growing an idea is the heartland of what designing is and it typically has two components. The business about modelling to explore and develop, that's part of it, and then optimising the controlling bit; I can't get everything of what I want, if I want this I've got to trade off some of that, so you've got to optimise one thing against another, and in the end you've got to prove that it's going work. So it's having ideas, growing ideas and proving ideas.

Now what we had at the end of their piece of work, at the end of that project, which we did in about eight schools, we had loads of end product prototypes and not much else. So the evidence that we had for making reliable assessments was a bit iffy and we needed that evidence to make sure that the performance evidence is that. That brings me to essentially the focus of what I'm talking about this afternoon.

In the process of this project, what we've done is taken a process that normally lasts a long time, like anything up to six months for a GCSE project, and we fast-forwarded it into a very intense short time, but then all we've got evidence of is the end product and what we need is evidence of the growing process in between, and that takes me back to a project that we did in the late '80s for the Assessment and Performance Unit. I don't know how

many of you will remember the Assessment and Performance Unit but it was an interesting body, a sub-branch of the DfES, trying to monitor standards across all schools, and what we did for the APU is we developed what we called 'fast folios' where we fast-forwarded kids through a process with a set of prompt questions to make sure they were trying to answer a particular question at a particular time within their development process. So in this case the kid was developing some kind of wheeled vehicle, wheeled toy, and at some point in that process the teacher would intervene with a question that said, please can you put down anything to do with the materials you're thinking about for this product; just jot something down, typically in red, about what kinds of materials you might think of using and why, and it was just a little inserted probe, if you like...

[gap: tape change]

... at the end it rings a bell, and the idea behind this is that there's a light in the middle and this band of light moves around to reflect. a really nice idea, and this was developed in a very short time. But then the activity stopped and the teacher said could you just put down anything about materials that's appropriate to the next development of this thing. What did the kid do? The kid wrote, grippy stuff. So in other words, when my hand goes on here and moves to dial time like that, it needs to be grippy; not, it needs to be a polycarbonate de dah de dah - it needs to be grippy, which is a designer response. That's what a designer would say, not knowing the material, they would identify the key quality of the material and subsequently find out about what material would be appropriate.

Using this fast-forwarding process we demonstrated a kind of fairly dynamic development of ideas through this paper-based assessment system that we developed in the late '80s. The idea with this project is taking this idea of a 'fast folio', almost a scripted activity that engages students in a design task and then drops in certain questions at certain times to see what they're thinking about these questions. We're adapting that into the ten hours of prototyping activity that you saw about the hats, to see whether we can design some activities that not only give us interesting outcomes and prototypes but also along the way, tell us something about the kids' thinking within this process that they're engaging in.

Any kind of assessment has to confront at least three of the horsemen of the Apocalypse and Susan talked about two of them about this morning, the problem of validity and the problem of reliability, so I'll pass by them. They are important, if this assessment is to give us a measure of something, it's got to be a measure of design and technology capability, not a measure of something else, not the measure of the ability to do a test. It's got to be a real capability and it's got to give you reliable data so that if I give the piece to one teacher and then give it to another teacher, they'll give you the same result. It's got to be repeatable but it's also got to be manageable. Susan didn't mention this, this morning. Remember the first round of National Curriculum SATS in '91 or '92, they were more or less valid, more or less; they were more or less reliable but they were utterly unmanageable and the result was that, in the end, the Minister had to resign. I was in the DfES just about the time when John Patton was under severe scrutiny about this stuff and it struck me as a real irony. There he was sitting in his office in a building that was called Sanctuary House and I thought, it probably doesn't feel like a sanctuary to him.

So whatever we do has to be manageable but we are in the process of trying these activities now. We're developing the very idea of the fast, furious prototyping project linked to the kind of worksheet that gives you pinpoints of process-based evidence as the thing works, as you go through this six or ten hours, and we're trialling the first of those next week down in Cornwall and then subsequently for the next few months. But even if they work, two big questions remain. The first one is can the fast folio be an e-folio? How much of this process, which is typically at the moment paper-based, can be screen-based and is it better if it's screen-based, what bits of it are better and what bits of it are worse? Looking for concept models of what an e-portfolio might look like we've come up with several models to think about. You can think of the e-portfolio as a kind of e-bucket, which is just a big virtual space where you pour everything in that you've got, and it's just a collection

point. It won't be an e-sponge. If you talk about the internet to kids they'll just say, oh that's where I'm doing my research. They view anything from cutting up the Argos catalogue to surfing the internet and picking up ideas as research, so it might have a bit of that in it. That misses the point completely. Even if you had the biggest e-bucket in the world and the most absorbent e-sponge in the world, that would not help you with the design portfolio, because the point about the design portfolio is that it has to be a tool that empowers not just the recording of work, but it's got to promote and develop performance.

The point that I'm making here, it's like you're saying for the active expression, it pushes ideas forward, the expression allows us to see ideas and to be clear what our ideas are. So in e-terms, whatever the e-portfolio is, it's got to operate in such a way that by doing whatever you do within it, it actually facilitates and develops; it doesn't just become another recording device or another storage device. It's got to be a development tool and that's a much bigger challenge. So we are, hopefully, going to be looking at that kind of dimension of the portfolio.

The other part of the challenge that remains is the problem of group work. At the moment in our next trials we're only looking at group work as a vehicle for promoting individual performance. How can kids use the group to promote their own work, and I recognise that's only a very watered down version of what I would really like to spend time on, which is looking at how groups operate to be effective in designing. You can play all kinds of games with this. You can tell so much about those two kids just by looking at them. You can see who's running it and who's not, and who's following.

So the problem of assessing designing is that designing is about process performance and, therefore, you're trying to access something that's fairly inaccessible. You have to work hard to get the evidence to assess. So assessing the inaccessible is about accessing the inaccessible. You've got to dig for it and develop strategies to promote the evidence that you need in order to make good assessment. Moreover, you have to do it without perverting the performance and without crushing the creativity. Now those are serious challenges that we are currently engaged in, in the context of design and technology. But I would like to suggest that insofar as we get this to work, it works in any domain. You can substitute designing for any other label you want where process is the key concern as opposed to the assessment of bodies of knowledge or collections of skills.

So that's a quick run through. We've got about ten minutes for questions.

**Q:**

It's a very small point, but you added manageability to the problems of assessment and I know that you were involved in a complex examination system back in the '70s for A-Level Oxford design and one of the things that killed that was not just manageability, it was cost. It had a number of components to it. Perhaps you could say a little bit more about that, the number of components involved and why that assessment system would have just been too costly to extract right across the whole country?

**Richard:**

The one you're referring to is the Oxford A-Level in Design, which, as you say, I was an examiner for in the early '80s. It was a fabulous A-Level and it really allowed kids to do dramatically different things, and to have their work assessed in all kinds of interesting ways, so that not only were a core of criteria identified but teachers could add more criteria and then teachers could bounce the relative weights of those criteria. But the whole thing hinged around a viva, between each individual candidate and an examiner. Every kid that sat that A-Level got an individual viva. As the exam boards got more and more competitive and more and more squeezed, that just went out the window and therefore that whole system became very compromised. One of the things that appeals to me about the e-portfolio is the idea of whether one could kind of re-construct the idea of a viva online and that's kind of interesting. The idea that you could have a student's work and portfolio in front of you, as an examiner sitting in some exam room, and ask questions, interrogate the

candidate about their design, which is not the candidate using a viva in the sense we saw earlier on of trying to answer questions that are known, it's like saying that the candidate can do a presentation without knowing quite what the examiner is going to ask about. So it will be like a real viva, one where you can't predict the question but you've got to know your work well enough to be able to formulate an answer to whatever the examiner asks you. So I think here there's the case where the technology could re-facilitate a system that we've lost because it was just too expensive and too difficult in the paper domain. It might be possible in the electronic domain.

**Q:**

I think you may have already answered my question, but I still want to reflect on it. It is a reflection relating to the whole day not just to your presentation, which in many ways has pulled things together. We started off this morning talking about assessment for learning. Other times today we've talked about digital-based assessment and I've been waiting all day for the two to meet. My feeling about what you were talking about in your presentation was that we were back with Paul Black and formative assessment, and that's fine, but I'm looking to see how we can pull together the two, so far, very separate strands of today. I think you've done it - but do it a little bit more.

**Richard:**

That's a really good question and you've articulated something that was lurking in the back of my mind as well. Yes, the formative assessment stuff, assessment for learning is a powerful part of what we do in design portfolios because we're concerned with the process, not just as demonstration but we use that process in the classroom to develop students' progress. We use the portfolio as a development tool, not just as an assessment tool. So that's a core part of what we're concerned about. So how does that link to the new technologies bit, and my first instincts about most new technology stuff - most of which I've seen in the States - really frightens me because it's so mechanistic and simplistic. It's becoming more sophisticated and it's becoming more interesting, and I think there is potential now. I can't quite see how it's going to work but there is potential to take this electronic medium and use it to support this kind of concept of a learning portfolio, and do it in a way that paper just can't do now for reasons of bulk, portability and examiners. In the electronic domain it would be possible to reconstruct some of the things we've lost because of the difficulties of the paper domain, but it's not an easy thing to visualise what it would be like. I know what it wouldn't be. I know it wouldn't be an e-bucket and I know it wouldn't be an e-sponge. It's got to be some kind of e-spanner, so that the electronic environment is one in which the process of engaging with the portfolio doesn't just record ideas, it actually prompts and develops ideas to grow back to the formative bit. But it's got to be a very clever bit of software to do that because it's got to enable kids to take their project in any direction they choose, which is one of the nice things about designing and you'd have to say how would Faye with her hat have done that in an electronic domain. One of the things we're looking at in our next trial is just using the simple technology of the digital camera, and one of those views that we've installed across what has become a six-hour activity is that effectively every 45 minutes every kid's work gets photographed, whatever it is, and we've got a system where we can whip out the photo card, plug it into a printer, print off the thumbnail images, stick them back on the kids' portfolios so they've got, within a minute, a little image of the object they're working on, the model they're working on and by the end of the activity they'll have a storyboard of electronic images, which show how their idea developed through that six hours. That's one example of how the relatively simple technology can show kids their own process, which is part of this business about the process being a learning process, so they become meta-cognitive, they become aware of their own process by seeing it unfolding before their very eyes. But we're still based on paper. Moving that into the digital domain completely so they're working on screen rather than on paper would be a different level of challenge; we are struggling to formulate what a proposal would be like and when we've worked out what that proposal will be like, we'll be going to people for some serious money to develop it.

**Q:**

You pre-empted part of my question and that was about self-reflection and I think that's just what you're talking about there. Another concept that was spoken about today, which I thought was really interesting, particularly with the primary children, and I would really like to hear your reflections on, is the idea of pure evaluation. Because with design, one of the things that we all expect in the real world when we're designing anything is to do a huge amount of user testing, to know is there a market for what I'm doing here, and what you were saying is that the process of evaluation currently for designer technology isn't about the real world; it's some sort of a construct.

**Richard:**

I think one of the really powerful things about designing and one of the things that really liberates kids' meta-cognitive awareness, is their ability to stand outside themselves and look in on their own processes. The thing that facilitates that is that designing is almost always for somebody else. You're designing something for somebody else to use, so my evaluation of it might be one thing but the question is what would this person think about it. So the kid has to empathise with what the other person would think. Now it can help to have other members of your group, of your peers, looking in on that and thinking about it, but, of course, the real strength is where youngsters are able to stand in the shoes of the other person and imagine what the other person is going to say. So it would be typical to have peer evaluations not really for evaluative purposes but as ways of prompting thoughts about what the real user might think. So you'd have a discussion in your little group about what would they really value, what would they really want, how would it really work, and the aim is for the group of students to have a really strong idea of how the ultimate user is going to feel about this. So designing for somebody else is one of the things that makes that a very pointed and powerful technique, and the value of the peers is just to give another set of perspectives on what the user would think. But I'm ashamed to say that in this early trial we're not really using peers in a fully empowered way. We're using them merely to encourage individual development. What I'd also like to do is a project where we really let loose the power of groups as creative groups, because that's how industry works. You go to IDO and you watch them developing a new product. They don't do that as individuals. They spark things off each other and they work almost always as a group. Now that is a serious challenge for QCA and for national standards. So I'm backing off that just a little bit and saying can we, artificially maybe, use the power of the group to see if we can enhance the individual performance, which is a kind of halfway house.