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

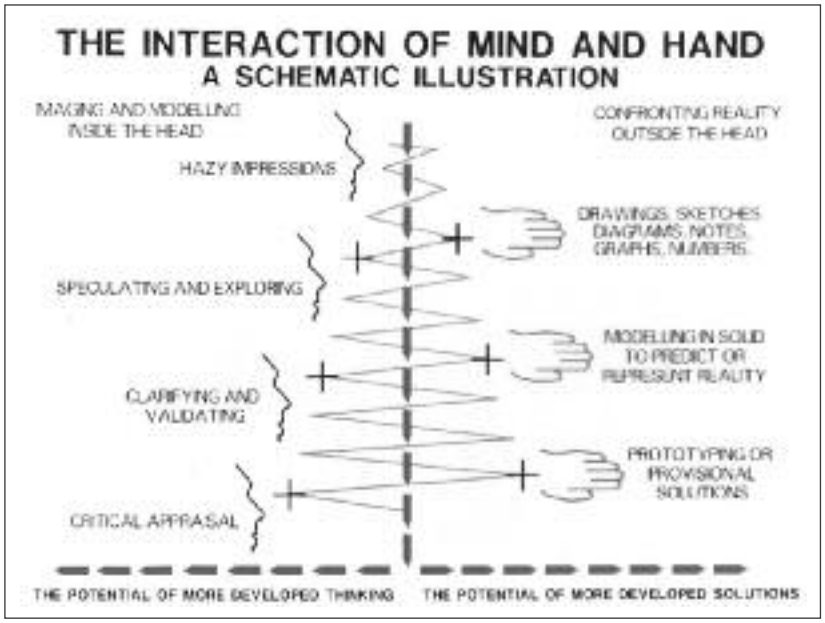

Figure 1. The APU Design and Technology Model

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.

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

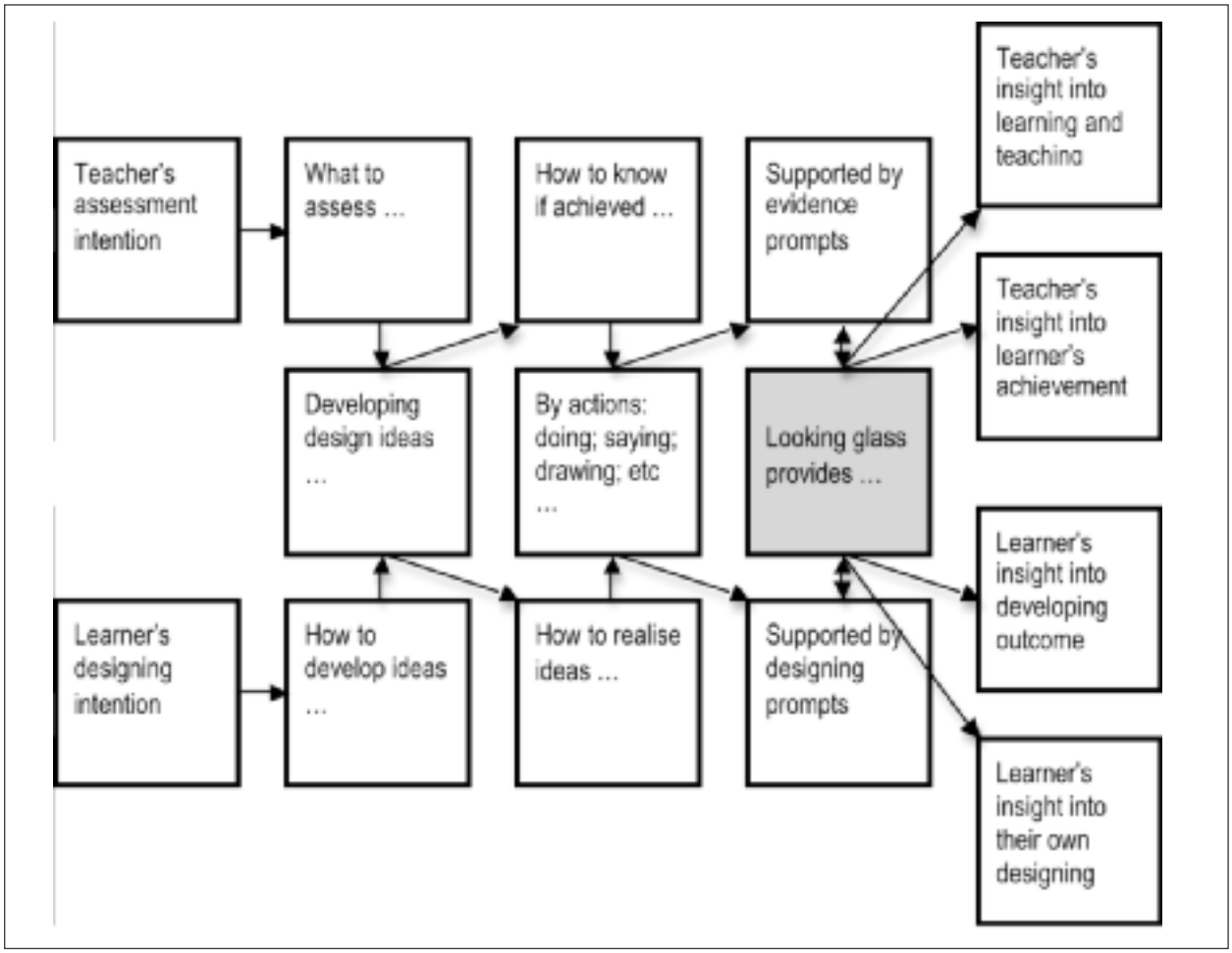

Figure 2. The mirror effect of effective evidence prompts

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

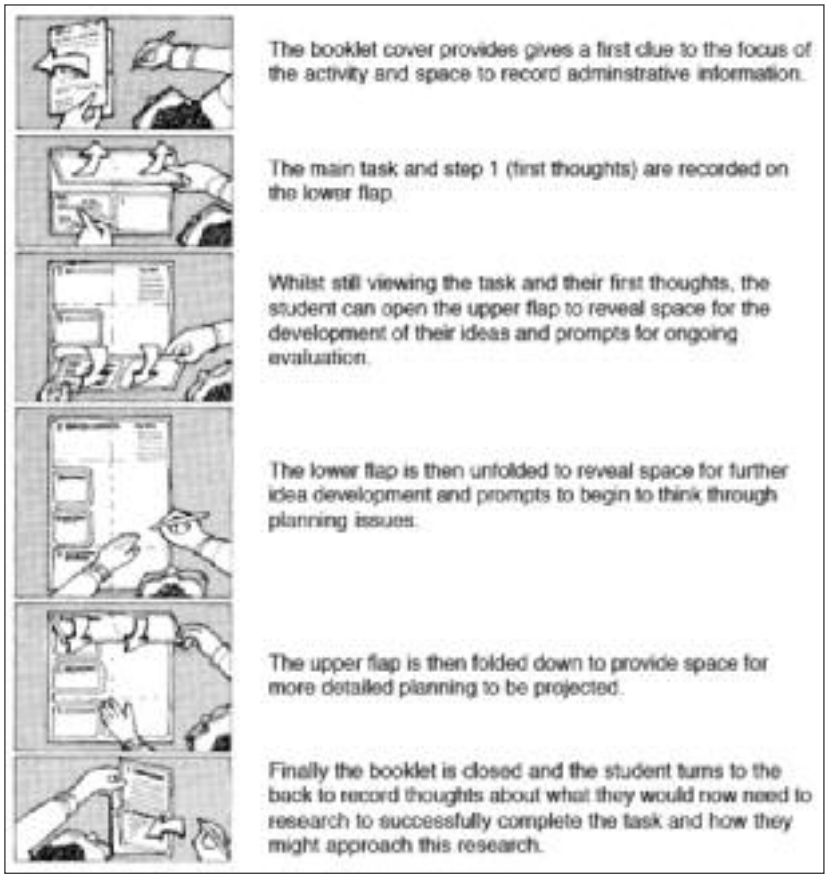

Figure 3. The unfolding booklet of the "unpickled portfolio"

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)

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

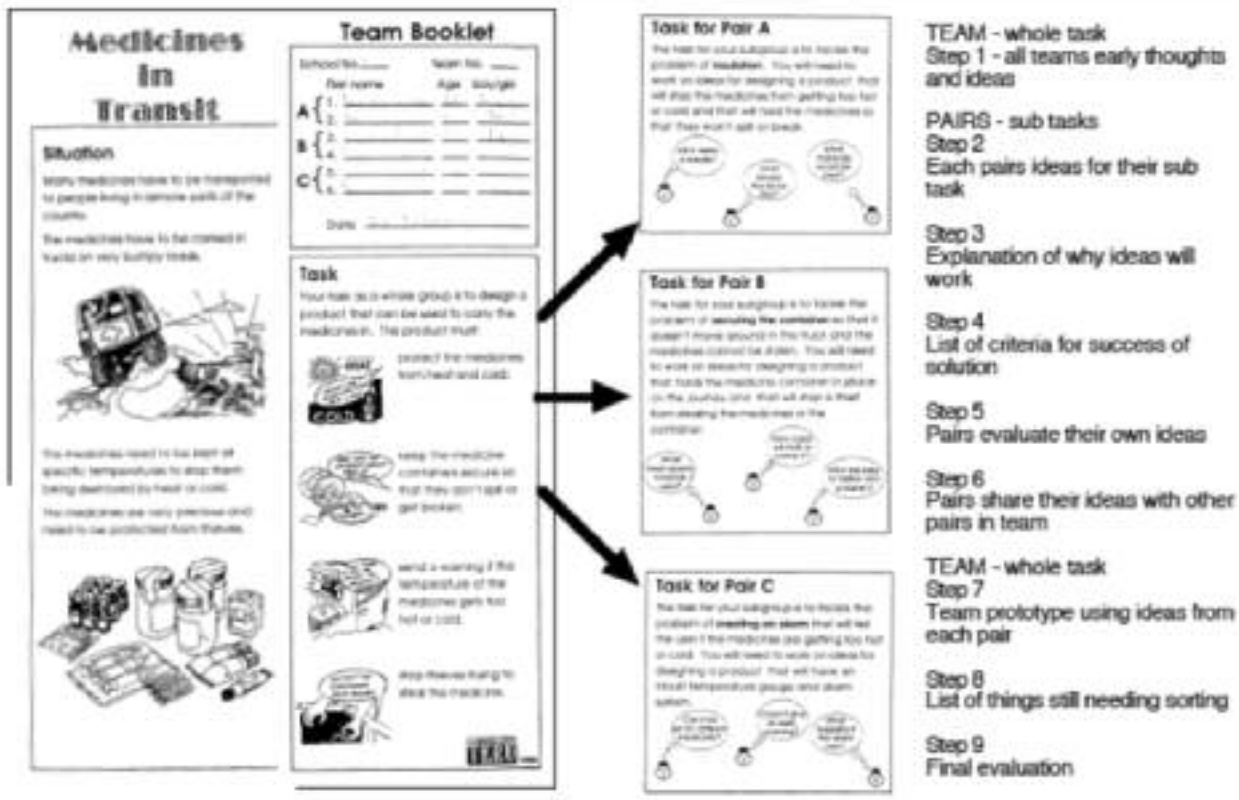

Figure 4. The challenge for the NWPTEP assessment task.

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

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

'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

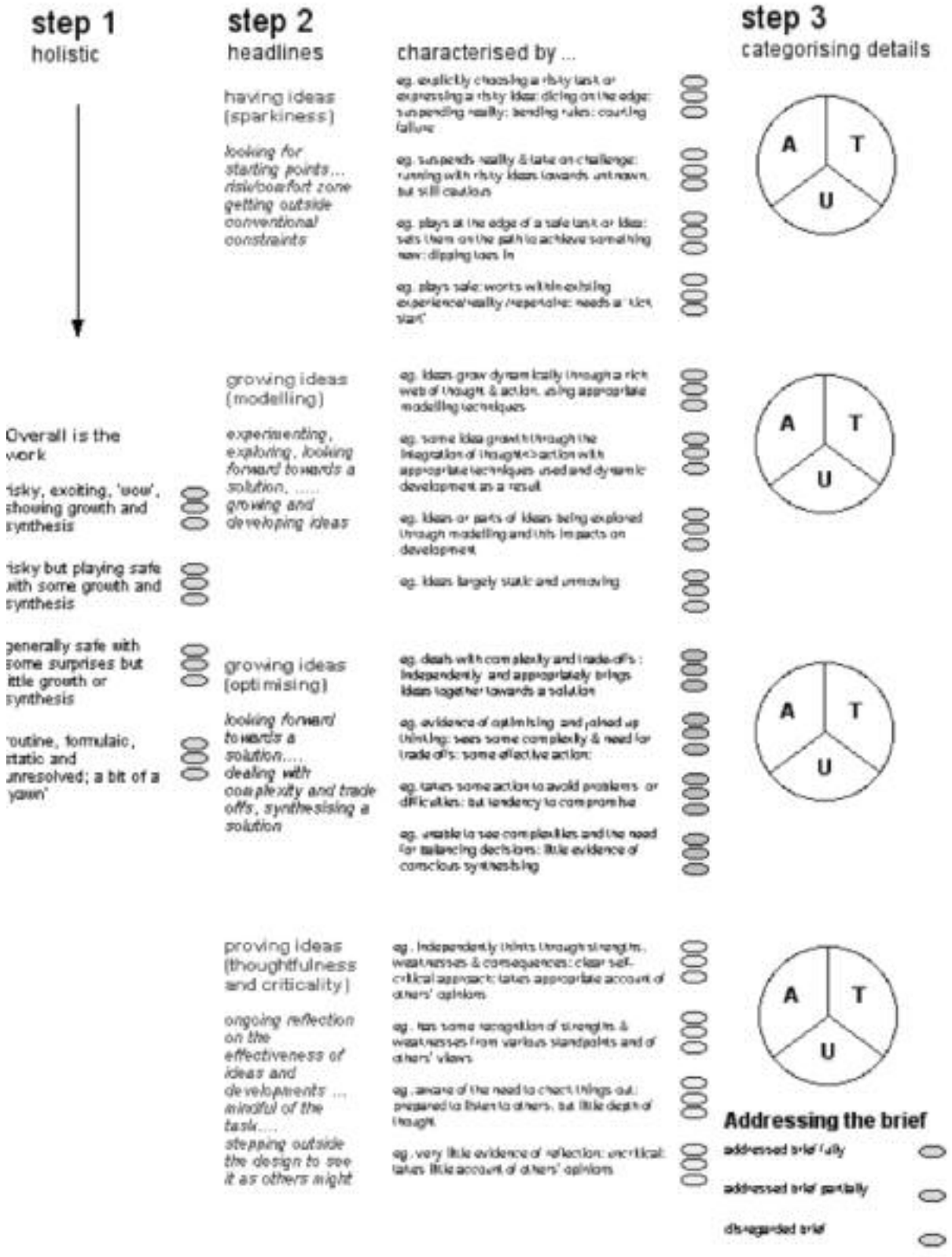

Figure 6. The assessment rubric from Assessing Design Innovation

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.

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

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.

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.