Int J Technol Des Educ (2012) 22:205–226
DOI 10.1007/s10798-011-9194-0

The validity and value of peer assessment using adaptive comparative judgement in design driven practical education

Niall Seery · Donal Canty · Pat Phelan

Published online: 10 December 2011

Abstract This paper presents the response of the technology teacher education programmes at the University of Limerick to the assessment challenge created by the shift in philosophy of the Irish national curriculum from a craft-based focus to design-driven education. This study observes two first year modules of the undergraduate programmes that focused on the development of subject knowledge and practical craft skills. Broadening the educational experience and perspective of students to include design based aptitudes demanded a clear aligning of educational approaches with learning outcomes. As design is a complex iterative learning process it requires a dynamic assessment tool to facilitate and capture the process. Considering the critical role of assessment in the learning process, the study explored the relevance of individual student-defined assessment criteria and the validity of holistic professional judgement in assessing capability within a design activity. The kernel of the paper centres on the capacity of assessment criteria to change in response to how students align their work with evidence of capability. The approach also supported peer assessment, where student-generated performance ranks provided an insight into not only how effectively they evidenced capability but also to what extent their peers valued it. The study investigated the performance of 137 undergraduate teachers during an activity focusing on the development of design, processing and craft skills. The study validates the use of adaptive comparative judgement as a model of assessment by identifying a moderate to strong relationship with performance scores obtained by two different methods of assessment. The findings also present evidence of capability beyond the traditional measures. Level of engagement, diversity, and problem solving were also identified as significant results of the approach taken. The strength of this paper centres on the capacity of student-defined criterion assessment to evidence learning, and concludes by presenting a valid and reliable holistic assessment supported by comparative judgements.

Keywords Teacher education · Technology education · Holistic assessment · Comparative pairs

N. Seery (✉) · D. Canty · P. Phelan
Department of Design and Manufacturing Technology, University of Limerick, Limerick, Ireland
e-mail: Niall.Seery@ul.ie

Background

Carty and Phelan (2006) categorised Irish technical education as a "craft-oriented approach" with a possible movement towards a "design approach" in some subjects. In 2007 the philosophical shift to design based education was embraced through the design and implementation of new syllabi at senior cycle[1] which formed the basis of a sound approach to technology education. This new suite of subjects reflected a significant change in the subject area philosophy, with the driving force now being technological awareness and design. The aims of the new syllabuses changed the focus of technology education, to provide students with the skills associated with design and realisation and the ability to apply these skills by thinking and acting imaginatively and creatively.

Shifting from pre-defined craft education to design-driven technology education presents practical, philosophical, and pedagogical challenges. Identifying and understanding what to value is the primary challenge facing contemporary teachers, learners, and assessors. There is considerable debate, nationally and internationally, on the importance of identifying skills and attributes appropriate to developing a knowledge economy where curriculum and practice play a significant role. This is problematic in an Irish context where a lack of continuity between the formal and the implemented curriculum now exists due to lasting traditions of the vocational focus (Dunbar 2010, p. 39). The need to realise the new syllabi is supported by Ritz (2009) who makes a direct call for educators and policy makers to "look beyond the development of engineers, industrial technologists, or craft workers" and argues that we must take educators beyond the limits of specific professions.

The defining of teacher education as a codification of existing practice is limiting when one considers the need for a new approach to technology education, as the educational purpose and objective no longer aligns with the practice. Exploring alternative practices with teachers who are simultaneously engaging with the implementation of new syllabuses, could render observing novel approaches so complex as to be indecipherable. Instead, initial teacher education affords the opportunity and environment to control variables and facilitate an exploration into a proposed approach to technology education that aligns with a contemporary syllabus.

Within the initial teacher education programme of study, the mandated need for a defined stream of wood and metal based modules (vocationally derived and defended) becomes less apparent when embracing design as the learning medium. Setting design tasks within material-biased modules limits the potential for coherent and meaningful engagement as many of the material and process design decisions are predefined by the material discipline.

This study focuses on the year 1 students of two concurrent undergraduate teacher education programmes, which were traditionally defined by a craft-based heritage. This paper explores the impact of student determined values and peer assessment on the development of technological capability through a design based activity within non material bias parameters.

Design at the core

Identifying the contemporary values and goals that underpin a new conception of technology education becomes important when contributing to the education of the student.

[1] Four syllabi were published, two of which were implemented in 2007—*Technology* and *Design and Communication Graphics*. The remaining two syllabi, *Architectural Technology* and *Engineering Technology* are pending implementation.

Lewis (2009) claims that there are a variety of generative cognitive processes that are more likely to occur in technology education than elsewhere in the curriculum. Shifting from the provision of technical skills to a broader education agenda supports the global consensus that values problem solving, construction techniques, creativity, and design (Rasinen 2003). Achieving the educational outcomes of creativity, autonomy, and fulfilment are critical in adopting skills and aptitudes that our rapidly changing global society necessitates.

Challenging the hegemonic culture that exists with design and technology is of critical importance. McCormick and Davidson (1996) suggest that teachers of Design and Technology often see the product outcomes and associated skills as important in themselves, and, in the end, these products tend to take precedence over the process of design and problem solving. The vocational heritage of Irish technology curricula may now find teachers struggling to establish what to value in the design activity. Much of the discussion within design and technology centres on the relationship between design and make. Baynes (2010) makes the following two critical points about this relationship:

1. There is a temptation to overvalue and hence to over assess the finished product.
2. There is often a miss-match between the pupil's imaginative vision and the pupil's ability to achieve it in reality.

Davies (1996) discussing *Art and Design* argues that the focus of assessment tended to be on the artefact and not on the process of learning, resulting in students presenting solutions to please the teacher instead of developing learning heuristics to solve the design problem. These learning heuristics as opposed to algorithmic approaches are core to the design activity (Cross 2001). Kimbell et al. (2004) identified that projects which tended to be creative and innovative were generally based on and driven by ideas. In contrast they highlighted that projects found to be merely competent or adequate (but not innovative) were based on a conventional, linear, managed approach to the process of designing. Kimbell et al. (2004) identified three elements to an assessment framework based on student ideas within the design activity that are indicators of creative and innovative solutions as: *Having*, *Growing*, and *Proving* of ideas.

Promoting and developing creative and innovative endeavour as a critical feature of a design based approach requires a means by which assessment can value student ideas and capture their design evolution. When considering the role of assessment, educators must strike a balance between the impact of criteria in directing student engagement and its capacity to value innovation and creativity within the activity.

With the objective of assisting teachers and awarding bodies, many design process models have been developed. Such models give a defined and standardised structure to engage with what is a complex, iterative process (Kimbell et al. 1991; Kimbell and Stables 2007). Mawson (2003) reports on the adherence of technology teachers to a linear concept of the design process. Kimbell et al. (1991) also noted that although helpful guides to teachers, defined models were dangerous tools as they prescribed the stages that pupils needed to complete and tended to result in pedagogical practice. This holds true in the Irish context where the guide to reporting design activities dominates the classroom approach and engagement in the process. Guilford (1950) expressed the view that it is difficult to develop design based attributes due to the conforming nature of schooling. The mismatch between the rhetoric about the importance of conceptual aptitudes and the value placed on creative talent raises concern about the coherence of educational strategy.

Conforming to assessment criteria

With clear objectives for technology education and an understanding of the role of the subjects within the context of contemporary schooling, the focus shifts to assessing student competency. The importance of differentiating between the assessment *of* and *for* learning (Stiggins 2005) becomes central when focusing on the process. Kimbell (2010) adds a further level of complexity that highlights the conflict that may exist between curriculum and assessment policy, with the difficulties centring on standardisation and testing. This questions the validity of what it is we are actually measuring. Students conforming to the assessment criteria and aligning their outputs to address given criteria (regardless of meaning) facilitate the sorting nature of contemporary assessment. This assessment challenge is amplified within technology subjects as Kimbell et al. (1991) argue that the essence of the problem with design based educational activities lies in the transformation of active capabilities into passive products, this is compounded by the assessment of active design thinking through passive student portfolios. Assessment criteria that over-define the stages and functions of design can render the objective benign as exploration, experience and decision making, which are central to learning, are compromised by the rigidity of assessment. True technological capability involves self-monitoring and awareness of how and when to use particular skills and knowledge. Barlex and Trebell (2008) argue that competency develops with coherent thinking and not just as an accumulation of knowledge. The value of design based activities lies in autonomy, the context and need to acquire relevant multi-disciplinary knowledge, demonstration of capability, problem solving, communication, and synthesis.

The problem is trying to measure evidence of thinking while encouraging diversity within a system predicated on standardisation and weighted criteria. Kimbell (2007), reports that “Learners can be excellent in design and technology in dramatically different ways”. Therefore the outcomes and solutions to design problems can often involve more variables than can be represented in a sequence or loop (Williams 2000). Facilitating diversity in response to design must be supported; the difficulty lies in the inability of traditional criterion referenced assessment to accurately reflect and value the process of the learning activity. So how do we help student teachers manage uncertainty, welcome ill-defined problems and take ownership of their own learning?

The measure beyond the artefact or finished product is critical to ensuring the sustainable value of design-driven competencies. Measuring a complex, iterative process requires a flexible model of assessment that can value evidence of learning in response to individual heuristics while supporting diversity and measuring capability.

Kimbell (2007) outlines the difficult nature of judging student work against abstract criteria. But when compared with an exemplar of capability the task becomes much more meaningful when normalised for the assessor. Project e-scape (Kimbell et al. 2005, 2007, 2009) outlined a new and innovative approach to the assessment of performance portfolios developed by Alistair Pollitt and based on the *Law of Comparative Judgement* (Thurstone 1927, see Pollitt 2004, see also paper by Pollitt in this journal). The approach uses an adaptive comparative judgement (ACJ) model of assessment where comparisons of students work relies on a holistic judgement with overarching criteria used to guide the assessor to make a professional judgement. This requires the judge to have an understanding of what is better or worse in terms of the required capability while eliminating the varying standards that may exist across a group of assessors. This approach had particular significance for the group of student teachers in this study as they could focus on evidence of capability without the worry of trying to determine levels of attainment. They could make reasoned judgements based on both their own capability and that of their peers. McDowell and Sambell (1999) support this suggesting that peer assessment has the potential to increase thinking, learning, confidence, and establish the role and purpose of assessment.

Research questions

This study focused on two workshop-based courses, traditionally rooted in the development of craft-based skills (i.e. wood and metal working). The courses were 12 weeks in duration and ran over one semester of the academic year. Responding to the shift in focus from the acquisition of skills to the inclusion of design-based aptitudes; these modules necessitated a revision of teaching and learning activities. Broadening the educational experience and perspective of students within the constraints of traditional institutional structures demanded a clear aligning of educational approaches with learning outcomes. Hence, this research focused on the following questions:

1. Is it possible to structure an assessment mechanism that will encourage diversity and creativity in the work, whilst effectively measuring design-based endeavours and capability?
2. Have student teachers the capacity to define capability and derive their own values appropriate to individual design solution?

Method

Approach

To help equip the students with the background they need, the initial 6 weeks of the semester concentrated on the development of core knowledge and skills in both areas of study. The research focuses on the second 6 weeks where students executed a semi-open design task based on the initial skills developed, giving a context to the learning and resulting in an exercise in 'near' and 'lateral' transferability (Schunk 2004). Students were given autonomy to design their workshop schedule, select appropriate processes and materials to achieve their design solution. As the design task was presented at the beginning of the module, students were given a scaffolded incubation period. The completed task accounted for 50% of the academic credit for both modules. The approach taken was to set a composite design task that bridged the learning outcomes of the wood and metal based modules (limiting material bias), to monitor and support student's progress throughout the task, and to evaluate both the process of design and the outcomes of the activity in terms of capability.

Participants

Students ($n = 137$) from the Materials and Engineering Technology (Metal) and the Materials and Construction Technology (Wood) initial teacher education degree programmes engaged in the modules. Students ranged in age from 17 to 32 with a mean age of 19.13 and a standard deviation of 2.97. Table 1 shows the homogeneity of both groups.

Table 1 Homogeneity of the subgroups

	Course	Male	Female	CAO entry[a]	Mature entry
M&ET (Metal)	64 (47%)	64	0	62	2
M&CT (Wood)	73 (53%)	72	1	66	7

[a] CAO entry—refers to the standard entry route to third level education. Second level pupils apply to the Central Applications Office and are awarded entry based on performance in the Leaving Certificate Examinations (Performance calculated on the best six grades achieved)

Design and implementation

Although the focus of this paper is on the role of assessment, the method used in the study highlights the design of the student activity, the infrastructure to support quality engagement and the assessment method. Ensuring that the assessment is not undermined by poorly designed learning activities or restricted by specific outcomes that limit meaningful enquiry, the research method focused on a semi-open design task.

Design of the semi-open brief

Following the initial skills building element of the module, students were challenged with the completion of a semi-open design brief. The brief was designed to align with specific learning outcomes of the relevant modules (Wood and Metal based decorative processing techniques) and provide scope for design. Students were required to make an A4 framed pictorial scene with the composition of the scene being of the students own choosing, but portraying a dominant feeling or emotion. In addition, students were required to complete a second artefact. They were challenged to design and make a flower (without facial expression) to express or reflect the emotion or feeling conveyed in their pictorial scene.

The objectives of these design tasks were:

- To provide students with a medium to explore and develop new knowledge and skills in the production of a coherent set of artefacts, that enabled students to transfer their new knowledge and skill into a project that embodied their creative expression.
- The brief did not require students to discuss or present their designs under a series of predefined headings. Students were not given assessment criteria, but instead were encouraged to identify what they perceived to be significant about their design solution and therefore they were encouraged to present for the assessor what they perceived to be evidence of their capability and learning throughout the activity.

Design of electronic infrastructural support

Moving from the traditional deliverables of a predefined, prescribed artefact requires a defining of not only the narrative skills to communicate creative inspiration and design but also the defining of what are deemed appropriate deliverables. With the emphasis on both the mastery of craft skills, and design, it was important that the medium used to respond to the task was appropriate. Students need the capacity to capture, manage, record and order information in an attempt to communicate the design journey, therefore the approach taken was to utilize accessible technologies (the students' mobile phones) and a data repository where students could construct their evidence of capability and learning throughout the activity. The benefits of integrating this technology were:

- Students had the capacity to capture inspiration in real time. The creative activity that is core to the education objective of the design brief was no longer confined to or constrained by the scheduled environment of the module. The creative activity of designing was now an exercise in synthesising global inspiration.
- The capacity and functionality of modern mobile phones enables students to capture information through a variety of media without judgment. Information can be later evaluated and synthesised.
- The exploration of possible solutions and the development of divergent thinking require a robust and flexible management system. Flexibility in supporting inspiration and encouraging the synthesis of information is critical to developing ideas. Students need to capture, express and communicate information in the most appropriate manner. The mode of representation will vary considerably across data sources, but more importantly can capture and illustrate the idiosyncrasy of the students 'Presage, Process and Product' model of their learning process (Prosser and Trigwell 1999).

Students were encouraged to use their mobile phones to capture, manage, and store their inspiration and design decisions as they happened. This was supported in two stages (see Fig. 1).

1. Students had access to a digital repository to facilitate the seamless transfer of data from their phones (or any other electronic device) to their own diary space. Data was generally transferred through Bluetooth, memory card or hardwire upload from their phones to the digital repository. The system could support SMS or MMS upload but at a cost to the student. In this repository, students managed and ordered their data for output to their holistic portfolio. This allowed them to present, order, reflect on, and communicate their design journey through a medium they felt was most appropriate to evidencing their learning and capability. Students had the freedom to present and emphasise what they deemed to be important about their design solution.
2. Secondly, these digital repositories were linked to a non-directive, holistic software application that presented the students' electronic portfolio of their work to the assessor in a single screen on their PC. This portfolio was populated from the student data repository (Fig. 2). To facilitate the multiple file formats that were used by the

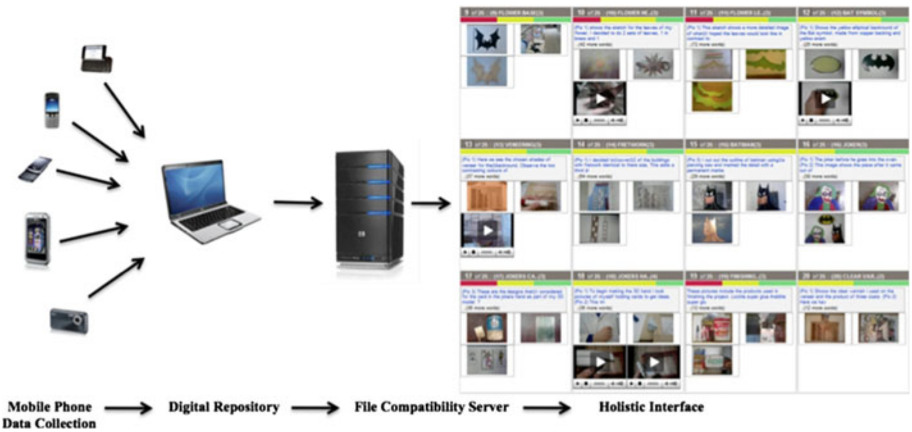

Fig. 1 Electronic infrastructural support

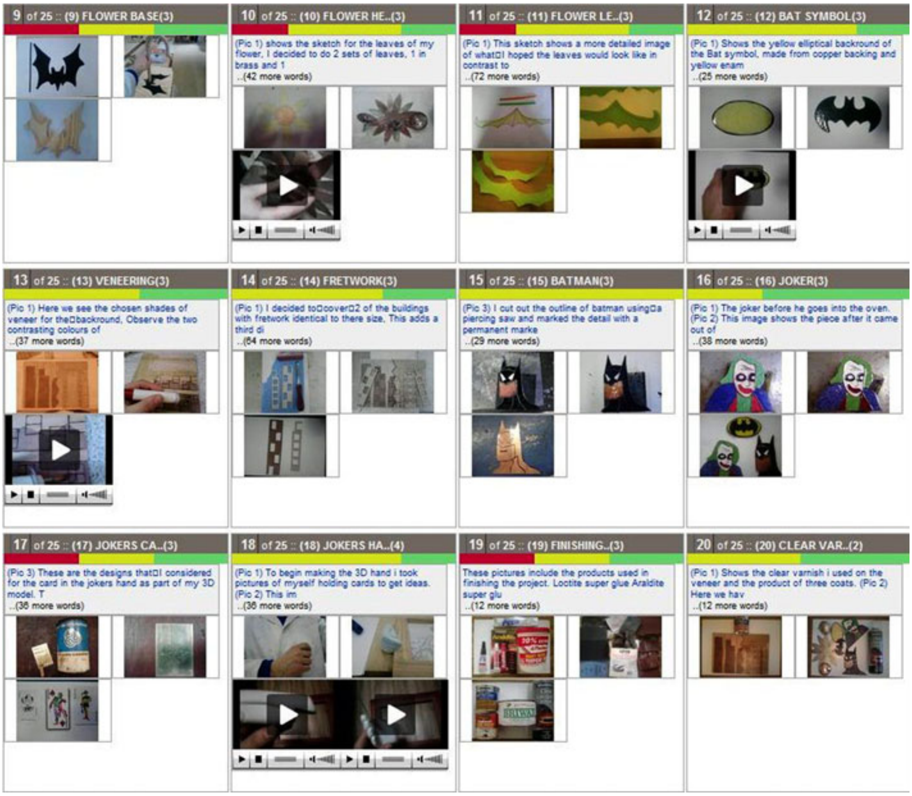

Fig. 2 Example of electronic portfolio interface

> students in the capture and compilation of their data, a file compatibility server was used to convert all data to formats supported by the electronic portfolio interface. The use of this file conversion server ensured that the students could utilise a technology device that they were familiar with (their phone) and had access to at all times. This web based portfolio together with the finished artefacts formed the evidence for assessment of capability and learning. This student-defined criterion-referenced portfolio was the blank canvas that students used to present their design journey and that, more significantly, supported the students in leading the assessor towards what they valued about their solution(s).

The portfolio comprised a number of panes (Fig. 3) that could be populated by the student from their data repository. Students made all decisions in relation to the completion of their portfolio from the number of panes required to pane titles and contents.

The construct of the portfolio was defined by each student to reflect the experience of their design process and represent the significance of their individual approach to forming their solution to the brief and showing evidence of learning. A unique feature of the portfolio was the colour coded tagging system that the student had the option to employ during the management of their data. While uploading their files to the data repository students had the option to create a tag or link between that specific piece of data to the

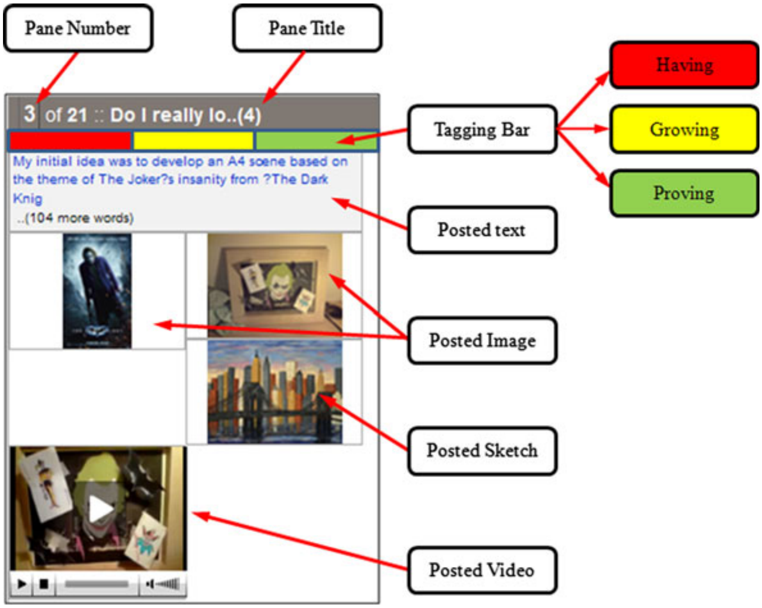

Fig. 3 Example of posts to an individual portfolio pane

criteria of *Having*, *Growing*, or *Proving* an idea. This required the student to reflect on the data being considered and to present it in terms of how it related to their engagement in the design-based activity. Tagging provides the assessor with an insight to the student perception of how they conceptualised and developed their ideas throughout the learning process. The colour coding of the student tags allows the judge/assessor to visually follow the student in one or all of the three tag categories throughout the portfolio.

Design of assessment mechanisms

The design of the research method was based on the analysis of student portfolios when capturing technological capability. The results present the relationship between what module leaders identified as capability in the electronic portfolios (independent of ACJ) and the consensus reached by the students when determining what they valued using the ACJ model of assessment.

Holistic assessment The approach supported peer assessment, as a student generated rank order of the portfolios facilitated an insight into not only how effectively students evidenced capability but also to what extent their peers valued it. The use of the holistic assessment interface facilitated the presentation of student engagement in the design activity. This was achieved by the student tagging data to the overarching criteria of *Having*, *Growing* and *Proving* of ideas throughout the activity providing an insight into their design inspiration, development and process. This supported the students presenting evidence of learning unattainable through traditional criterion referenced assessment.

The student rank order of portfolios of work was generated using an ACJ assessment method (Pollitt 2004) developed within project e-scape (Kimbell et al. 2009). The method presented a judge with two portfolios of students work. The student judge analysed each

Table 2 Portfolio analysis rubric for evidence of capability by module leaders

Skills		Knowledge		Problem solving	
Criteria[a]	Rating	Criteria	Rating	Criteria	Rating
Diversity of skills	(1–10)	Within module	(1–10)	Frequency of problems created[b]	(0–3)
Level of skill acquisition	(1–10)	Beyond module	(1–10)	Sophistication of problems	(1–10)
Justification for selection	(1–10)	Application	(1–10)	Success in solving problems	(1–10)

[a] Skills were rated based on evidence from the electronic portfolio and not the physical artefact

[b] Frequency of problems created was rated as None (0) Low (1) Medium (2) or High (3)

portfolio for evidence of capability and then decided on which portfolio was better. The judgement process was based on criteria, but these criteria were not directly scored, instead they were interpreted by the judge to form a single judgement. The aggregation of these judgements across the portfolios produced an order of the work from what was judged to be the best portfolio to the worst. This is called the "rank order" of student work. A pairs "engine" generated and managed the comparisons of the student electronic portfolios as the process evolved. To produce a valid rank order of student work, the judges made their decisions on the qualities that they believed showed evidence of capability and learning. Judging was monitored by the system and presented indicators of the emerging consensus within the judging group. Students whose judgements had a significant 'misfit'; that lie significantly outside the consensus were identified and monitored/advised in relation to their judging process.

With the student rank being generated from the electronic portfolio, it was critical to investigate if the portfolios provided the data necessary for the assessor/judge to evidence the capability of the student. A descriptive analysis of the portfolio content was completed by module leaders based on Gibson's model of capability (2008) which outlines skills, values and problem solving in the context of conceptual knowledge as the cornerstones of technological capability. Table 2 illustrates the criteria and scale used to determine the evidence of capability in the portfolios that was used by the module leaders.

Implementation and analysis

Figure 4 illustrates the elements of the research method. Statistical analysis was employed to explore the relationship between the module leader's determination of data that represented elements of capability and the student rank order of work produced by the ACJ process.

Findings

The findings are presented in three distinct sections focusing on the student's reaction to, and engagement with, the task in the learning environment, their ability to evidence capability without the aid of explicit assessment criteria, and their ability to validly assess the quality of the work produced by the class group. The evidence of capability from the student portfolios rated by the module leaders is compared with the student assessment of capability using the ACJ model of holistic judgement. The first step to ensuring the student defined rank order of work was valid was to compare the elements of capability from both

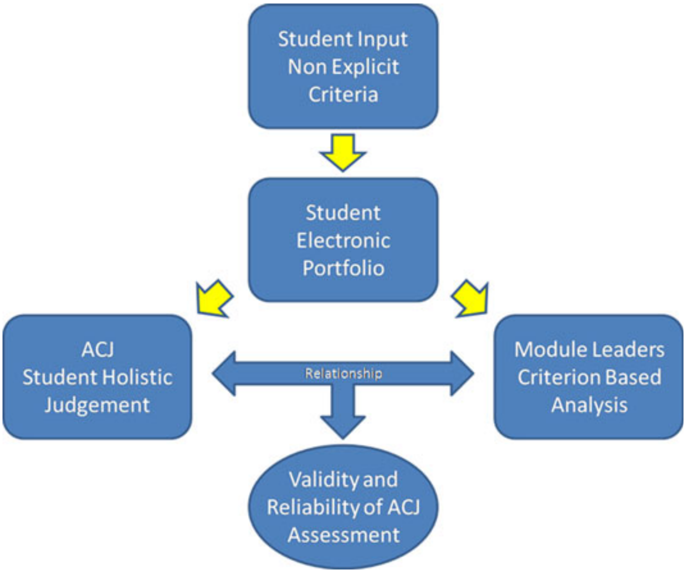

Fig. 4 Overview of research method

Fig. 5 Examples of students work

assessment perspectives (peer and module leaders) and correlate to see if the student generated rank order of work valued the critical elements of capability as identified by the module leaders. We also present here evidence concerning the capacity of the students to value their peers work through their definition of capability.

The student work was of a very high standard presenting quite a diverse range of creative solutions to the brief. Figure 5 shows 3 examples of student solutions to the design task:

Observations of behaviour

It became apparent that students' confidence in holistic assessment and the need to document their learning changed the behaviour observed in the workshops. Unprecedented lecture and laboratory attendance was recorded in both modules by comparison to previous years. The high level of attendance resulted in a dynamic and enquiring lecture series that aided in exploring the new philosophical shift in technology education. The shift in autonomy was supported on a structural level by the digital repository, as highlighted in students' comments:

> Probably the best 6 weeks we'll have in our time in college, because we got the power to make something that was ours, and that really motivated us.

From observation of the workshop activities it was noted that students were engaging in additional activity by comparison to traditional workshop engagement. Students were making decisions, all students were deciding on materials that would best represent their design, processes were selected on the grounds of desired effect and students were purposefully modelling and prototyping. Experimentation with newly learned techniques created a constructivist comprehension of the core building blocks as students attempted to transfer knowledge and skills into their conceptual solution. The following student comment supports these observations:

> I have created a model of the A4 scene to help me understand the geometries of the picture what I can and cannot include. This really helped me decide on what will and won't work even though it is made from card. The model has given me the idea of making the rod stick out from the frame. The model will also be invaluable to me when I am fabricating the parts of the scene as I can use them as templates.

Engagement with the electronic portfolio

Supporting the shift in the workshop activities, students were required to develop an electronic portfolio to evidence their capability and learning. The product outcome of the design task which traditionally dominated practical education now formed the catalyst for creative expression, problem creation and decision making, aligned with Kimbell's dialectic model of design activities (Kimbell et al. 1991). From the beginning of the design project, the workshops became a community of learning, as students positioned projects to be photographed, recorded video footage of peers work, shared Bluetooth capabilities and supported each other in achieving their goals. A culture of teamwork and team building became evident within the workshop/classroom; this is reflected in the following student's comment.

> I think the fact that we were given the freedom to make whatever we wanted and to use our imagination as much as possible enabled a lot of positivity amongst us as students, and for us as teachers it will allow students to develop their own thoughts, goals and aspirations, through our guidance and link imagination and thought with a set of workshop skills and problem solving skills

Over the course of the design task each student produced on average 50 unique electronic files on their mobile phone and submitted them to their individual digital repository. Files were produced in real-time and ranged from video to text and addressed all facets of their design activity. The following student comments are examples of real-time capture of inspiration:

> I got my idea from a flower my girlfriend had in the apartment, it was one of those fake flowers and i never really paid much heed in it until I was looking for my idea. I actually took the flower apart and realised it would be a nice design to make out of copper because it was really only one leaf and when opened out it resembled a conical shape.

Together with building a holistic interface the electronic infrastructure also supported students reflecting and tagging their files in terms of the over-arching criteria of *Having*, *Growing* and *Proving* of ideas. Although it was not mandatory to tag, over 90% of students tagged their work.

Over-arching assessment criteria

Students had complete autonomy to decide on the structure, content, and medium to present their design story. This gave the students the opportunity to present the assessor with something more than just a chronological reporting of their engagement in the design activity. The portfolio was built from their digital repository during engagement in the learning process. This facilitated the student in presenting the assessor with an insight into their meta-cognitive development throughout the activity by subtly presenting the individuals' design process model, approach, and emphasis. Figure 6 presents the mean tagging per portfolio for the entire cohort, which presents a more interrelated dynamic indicator of the activity. From the graph it is clear that on average students illustrated that they were engaged in idea generation and development throughout the entire design activity.

This gives a useful insight into the way students engaged in the learning process and shows the capacity to capture something broader than capability. The relationship between *Having*, *Growing* and *Proving* of ideas appears to align with the dialectical model of interaction between hand and mind presented by Kimbell et al. (1991).

The following section explores the capacity of the holistic interface to evidence student capability, as this interface forms the platform for the adaptive comparative judgement.

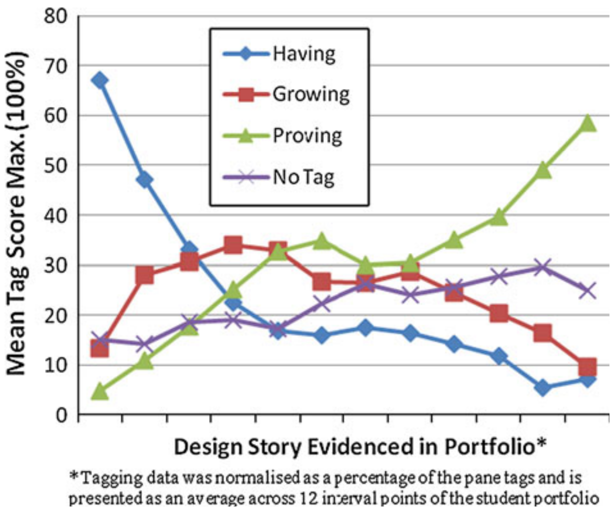

Fig. 6 Tagging by portfolio

The evidence of learning in the holistic interface

Initially the analysis centred on the module leaders exploring the set of portfolios for evidence of mastery of skill (justification of material and process selection), evidence of theoretical knowledge (its need and application), and evidence of problem solving. Both module leaders analysed the holistic interface and rated students work based on Gibson's model of technological capability (2008).

Craft skills were analysed and rated on the evidence of diversity of the selected processes, level of execution of skills and the justification for selecting the appropriate craft skill in the context of the project (Fig. 7).

A Spearman's correlation coefficient test was completed for all three categories of craft skill, indicating a significant relationship between level of skill, justification of skill and diversity of skills with a strong effect (Table 3).

This indicates that students that presented high levels of skill could also justify the selection of that skill and tended to present a broader skill-set across their project work. The objective of the study was to see comparatively if the ACJ (peer judgements) valued this capability appropriately.

Student portfolios were analysed for knowledge as defined by the module outcomes (within the module), applied knowledge (application), and knowledge that was acquired as defined by the students' needs beyond the remit of the module (beyond the module) (Fig. 8).

A Spearman's correlation test was completed for all three categories of knowledge indicating significant moderate to strong relationships in all tests (Table 4).

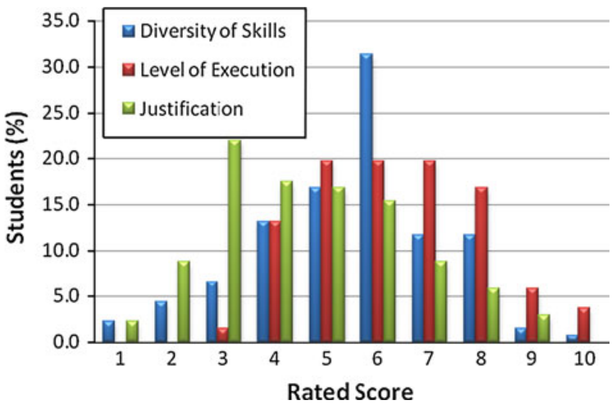

Fig. 7 Analysis of practical craft skills in electronic portfolios by module leaders

Table 3 Skill category correlations

Spearman's test	Correlation (*r*)
Between diversity of skill and level of execution	0.814
Between diversity of skills and justification of skills	0.800
Between level of execution and justification of skills	0.783

All significant with $P < 0.001$

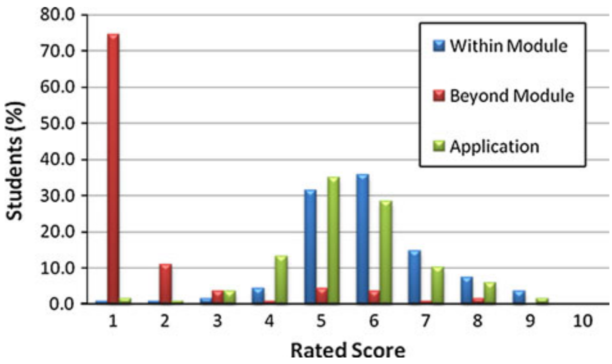

Fig. 8 Analysis of knowledge in the electronic portfolios by module leaders

Table 4 Knowledge category correlations

Spearman's test	Correlation (*r*)
Between knowledge within and knowledge beyond the module	0.618
Between knowledge within the module and application of knowledge	0.867
Between knowledge beyond the module and application of knowledge	0.530

All significant with $P < 0.001$

This also afforded the researchers the opportunity to look at students' acquisition and application of knowledge and the capacity of ACJ to value it.

When examining problem solving capabilities, two distinct categories were identified by the module leaders where problems created by the students centred on processing and aesthetic problems. Students' problem solving ability was rated ($1 =$ low to $10 =$ high) by sophistication of problems created and level of success in solving that problem. Figure 9 presents these findings.

The relationship between sophistication (processing) and success (processing) was significant with a strong effect size ($r = 0.830$). For problems associated with aesthetics, sophistication and success scores showed a similar distribution with a significant relationship between both aspects and an effect size of $r = 0.83$.

These results show that there was clear evidence that students were able to demonstrate capability through the electronic portfolio. This illustrates the effectiveness of the brief in providing an activity that enables the students to explore, develop and present capability. The results also indicate that students presented a broad and balanced understanding of capability through their project work. Strong correlations within the sub elements of capability indicate the inter-related nature of performance and capability. With clear evidence of capability displayed in the portfolios the focus turns to the capacity of the ACJ peer assessment model to value what has been identified and rated in terms of capability by the module leaders.

Holistic assessment supported by adaptive comparative judgement

This section presents the capacity of the ACJ system to measure and value evidence of capability as a result of the design task. The results focus on the relationship between the module leaders assessment of capability as presented in the student portfolios and the

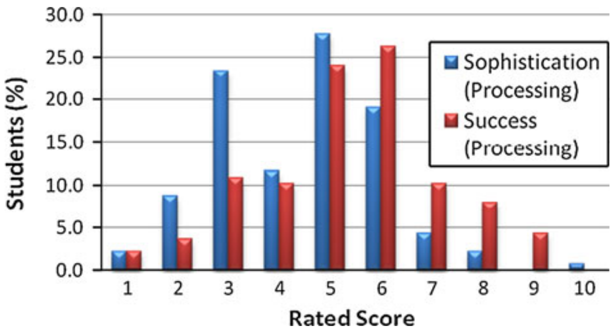

Fig. 9 Analysis of problem Solving (Processing) in the electronic portfolios by module leaders

student holistic judgement of capability using the ACJ assessment with the same set of portfolios. As the core elements of technological capability are knowledge, skills and problem solving, these must be evidenced in a valid assessment tool. To establish if these core elements were a significant influence on the judging decisions, a strong correlation between their evidence in the portfolio and their portfolio rank position is necessary.

Defining the rank order

The adaptive comparative judgement was completed by 63 students acting as judges achieving a reliability coefficient of 0.955. The paired judgements were completed over 19 estimation rounds with the judging group. The rank order of the portfolios stabilised statistically after 11 rounds with judging continuing until 16 estimation rounds were complete. Figure 10 shows the rank order of portfolios determined by the parameter values generated through ACJ. Also indicated on the graph is the error plot for each portfolio (vertical red line on each portfolio) which shows low error values across the range of portfolios on the rank order of work. This highlights the unanimous consensus on the position of each portfolio of work on the rank order.

Detailed monitoring of the judgments by the pairs engine provides useful statistics on the consistency of the judgements made by the students and the level of consensus on the position of the student portfolios of work on the rank. The weighted mean square statistic (WmnSq- calculated by Mean + 2 × SD) is important when indicating portfolio or judge misfit, any value greater than this criterion is considered to show a significant amount of misfit (Pollitt p. 73 in Kimbell et al. 2009). Portfolio statistics were analysed and Table 5 shows that there were seven portfolios that were outside the fit criterion for the rank.

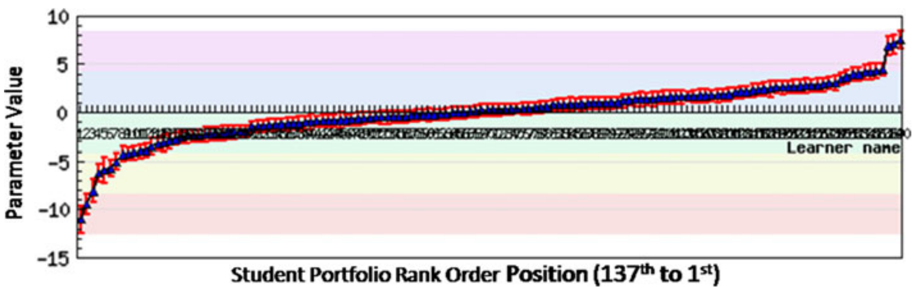

Fig. 10 Parameter value error plot of student defined ACJ rank

Table 5 Fit statistics for portfolios

Portfolio data	Rank data	Note
Portfolio WmnSq (average)	1.178	Theory predicts ave should = 1
Portfolio misfit criterion	1.69	Mean WmnSq + 2(SD)
No. of portfolios outside criterion	7	
Position of misfit portfolios on rank	1,3,4,7,13,43,127	

Table 6 Fit statistics for misfit portfolios

Misfit portfolio rank	1	3	4	7	13	43	127
Parameter value	6.72	5.62	5.61	4.5	3.32	1.38	−3.48
WMnSq	1.83	1.83	1.97	2.53	1.7	1.88	1.73
UWMnSq	21.42	7.03	3.38	22.2	17.44	8.65	2.42

Table 6 presents the fit statistics for the portfolios that lay outside the fit criterion which are referred to as misfit portfolios.

On analysis of the judgement history of the misfit portfolios identified in Table 6 it was noted that technical problems with the portfolio were cited as reasons why a particular portfolio 'lost' to a portfolio of a lower parameter value. Judging comments:

> portfolio B would have won only for the text being in computer language

> could not load portfolio A??????

Generally the portfolios in this misfit category 'lost' to portfolios that were lower but relatively close to them in parameter value. This would account for them being outside the fit criterion. Some also had unusual decisions made on them which is highlighted by the unweighted mean square value (UWmnSq). Portfolio 1 on the rank won all of its 13 judgments except for one where it was compared to a portfolio that was lower than it by 5.48 in terms of parameter value which is quite large representing general consensus that these two portfolios were not similar in quality. The judge made no comment on this judgement, and may represent apathetic engagement or simply a slip.

Analysis of the judging data shows two of the 63 judges from the group were outside the fit criteria for the generated rank. It should be noted that one of the misfit judges from the rank only made one judgement while all other judges averaged 20 judgements each. The second judge was outside the misfit parameter by only 0.01. Overall this low level of misfit for novice assessors is impressive and shows the capacity of the system to aggregate the values that defined the students' decision making.

The following comments highlight the students' reflections on specific judgements and show their response to evidence presented and capability identified.

> Very good representation of the chosen inspiration. Very interesting inspiration and portrayed very well in the project. Good description of production, material choices and production stages….

> A is the winner because the steps are more clear. The ideas are good and they adapted better to different problems. The workmanship was more technical and advanced and better executed.

Table 7 Relationship between rank and evidence of capability (holistic interface)

Elements of capability and ACJ rank order correlation			
Skills	(r)	Knowledge	(r)
Level of skill	−0.638	Knowledge within module	−0.64
Diversity of skill	−0.632	Knowledge beyond module	−0.501
Justification of skill	−0.669	Knowledge application	−0.58
Problem solving-processing		Problem solving-aesthetic	
Problems created (processing)	−0.474	Problems created (aesthetic)	−0.406
Problem sophistication (processing)	−0.592	Problem sophistication (aesthetic)	−0.536
Problem success (processing)	−0.603	Problem success (aesthetic)	−0.609

All significant with $P < 0.001$

> Barely the winner, tough decision. I gave it to him because I felt his emotion more than the other one, he was using his own emotions and own home as an inspiration and came through well.

The module leaders identified the elements of capability in the initial phase of the research. The previous section presented the students' rank order of work giving an insight into the validity and rationale for the judgements.

Having established that the aggregation of judgements concluded by the student judges produced a reliable rank, it was necessary to ascertain if their definition of capability aligned with the evidence of capability presented by the module leaders to determine if the ACJ rank order was valid. Although there is consensus within the student rank, they could have all agreed on inappropriate evidence. This stage of the findings looks at the relationship between module leaders' analysis of capability and the student defined rank order of peers work. Table 7 presents the relationship between the ACJ rank and the evidence presented in the holistic interface under the headings of practical craft skill, subject knowledge and problem solving abilities.

The statistically significant moderate to strong relationships between the evidence of capability presented in the electronic portfolios would suggest that ACJ assessment has the capacity to evidence and reward capability. This relationship is based on student's definition and critique of capability, where theoretically a skew in either or both could undermine the reliability of these relations. The negative correlations indicate the higher a portfolio scored, the higher they were ranked by the students (i.e. 1st, 2nd etc. hence the negative value).

A significant relationship was also found between the effectiveness of communication skills and the rank order with $r = -0.585$ and $P < 0.001$. However the number of panes presented by students in the interface had a weak relationship with where they were positioned on the rank. This suggests that student's judgements were based on criteria beyond the volume of data presented. This is supported by the following student who commented:

> Close call but I liked the ideas and the effects which looked very good, I know he didn't have alot of content in the portfolio compared to the other one but I liked all of the work and the finished product.

More than just capability

From analysis of the portfolios it became clear that there was evidence presented beyond the conventional focus. Students presented mistakes as evidence of learning as highlighted in the following comment:

> There was undoubtedly several mistakes made throughout the project, but fortunately there were none that I did not learn from. It is this learning and knowledge that I hope to take out with me into the future and I believe that always being aware of mistakes that you have made only makes you better as a teacher

Students also presented the weaknesses of their work and clearly articulated why they saw these as a weakness. The self reflection in the following student comment shows the capacity of the holistic assessment to support learning:

> I fixated on a dinosaur colouring book. To be honest I am not really over satisfied with my project. It is very plain compared to some of the ones submitted by my fellow classmates. I believe this is down to planning. I know having looked at all of the projects submitted and my own, my project is nowhere near the best! If ever something like this is done again more planning will need to be done to be near the best.

The process of assessment also supported critical judgement beyond capability, as students commented on the significance of peer engagement.

> Initial idea well thought out. Grew and proved ideas well. Flower is exceptional, really defiant, you can feel the stubborn cockyness of it.

> Uses the same technique as A but to greater depth as he contrasts and compares. This portfolio clearly shows a deeper research into the project with unlike A - three different design ideas.

This shows the capacity of the students' to make critical judgements when presented with a diversity of information on which to make their decisions.

Discussion

Banks et al. (2004) argues that student teachers must acquire and apply technological knowledge, master craft skills, make decisions and selections on the appropriateness of processes and materials, display problem solving capabilities, make professional judgement, develop meta-cognitively and communicate effectively. This study approached process based modules with a design-driven focus. Supported by the flexibility of the adaptive comparative judgement assessment model, students embraced the concept of removing external assessment criteria and defining the criteria that they judged to be applicable to their design solutions. The essence of this approach was based on students showing evidence of progressive enquiry within the area of study and not a reactionary response to standardised assessment. Students needed an approach to construct subject knowledge that facilitated risk taking, valued making mistakes, and developed reflection (McCormick and Davidson 1996; Barlex 2007) as the cornerstones of making meaning.

The evolution of collegiality among the class group was far in excess of that which has been observed previously—in prescribed and traditionally focused activities. One hypothesis for the students' supportiveness is that no student felt as if they were trying to compete on predefined areas of the brief. As the activity was not determined by criterion reference, students were at liberty to share, critique and support the workshop based activities. Students were confident that their projects could not be effectively measured by a generic linear application of predefined criteria and therefore were not competing on that level. Instead students were creating the criteria that showed evidence of their capability, displayed as a result of completing the project. A critical factor for the students was that the assessor (their peer) could empathise with their work having completed the process themselves. The process also encouraged students to engage in discussions on capability with their peers in an effort to broaden their concept and understanding of capability as the ACJ model sees judgements on students work made across a wide range of assessors.

While the design solution validly measures the mastery of craft and process skills, and an observation of the creative expression combined with the emotion displayed by the flower artefact gave an insight into creative capabilities, this is not the full story. The significant education value lies in the process of learning. The problems each student solved, the problems each student created and then solved, the inspiration for their expression, the synthesis of their ideas, and the emotional engagement that enabled students to personalise the experience, formed the true value of the design activity. Developing design activities within structured educational paradigms is a challenge, but can be achieved if the learner is given a proactive responsibility and not a spectator's role in their own learning.

The study illustrated how the students became the auditors and authors of their own learning experience. Learning developed from a paradigm of *'have to know'* to *'need to know'*, as students were not only creating their own solutions but also their own problems. The essential transferability of new knowledge and skills fostered a deep engagement in the learning activity. The autonomy gained during the design task and the absence of external criterion-referenced assessment reduced the anxiety of having to produce what was perceived to be required. Instead students could focus on defining and communicating what they valued in a personalised learning activity.

Student confidence in a democratic approach to the assessment formed the basis for unrestricted engagement. As a result the relationship between student and assessor (their peer) was relaxed. Students did not need to second-guess the values and preferences of the module leader. Not having to predict 'what the assessor is looking for' was one benefit of the approach taken, not being able to align your solution to what everyone wants renders the dominant 'formulaic, routinised, and predictable' (Kimbell et al. 2004) approach to design redundant. Students quickly became the co-constructors of their own meaning. What resulted—which was strikingly obvious—was the diversity in ideas and solutions in the students work. Barlex (2007) suggests that such a spread of responses is an indicator of a class in which creativity is being supported.

This study employed an adaptive comparative judgement model of assessment at the core of the investigation to develop and measure appropriate competencies in technology education. The findings of the research illustrate that the electronic portfolio supported students making critical judgements about the data that they collected and presented in their design journey. Capability as defined by Gibson (2008) identifies skills, knowledge and problem solving as the cornerstones of a competent person. The capacity to evidence these elements of capability in the holistic interface was illustrated by the significant relationships presented.

The importance of this study centres on what we describe as the 'delta' criteria. This is the ability of the assessment criteria to change and respond to the work the student presented as their evidence of capability. The meaningful engagement in design-based activities allowed students to explore and develop capability in a way that was distinctly personal. How they created meaning for themselves in this process is of critical importance resulting in something unique to the student but that was nonetheless recognised and valued by other students in the judging process. Evidencing this variable is central to the assessment model that enables it to react to the student (dog wagging the tail), rather than the current ubiquitous practice in which students shape their work according to the requirements of assessment (tail wagging the dog).

The final element to the assessment model employed was the consensus achieved by the students in generating the rank order of peer work based on evidence of capability and learning from the electronic portfolios. With valid and reliable ranks generated by the judging cohorts the corresponding portfolio and judging fit statistics present a very high level of consensus within the groups on their understanding and interpretation of capability and learning within the activity. The significant correlation between the student defined rank order of portfolios and the module leaders' evaluation of capability indicate the success of the modules in developing students that were both technologically capable and aware, but more importantly presents the capacity of ACJ to support design based activities and objectively value the process.

Conclusions

Focusing on design priorities in craft-based education challenges the practice and approach to teaching and learning. This paper highlights how assessment can be a positive, progressive influence that develops higher-order thinking and encourages innovative and creative enquiry. As a result, didactic transposition is replaced with a reactionary/responsive activity that mentors autonomous enquiry.

Not only can students engage in creative endeavours within a reduced risk model of assessment, but assessment can drive diversity. Comparative pairs supported students evidencing more diverse skills based on responding to more problems (that they created), having learned more technological knowledge.

As a group, students effectively validated the reporting structures employed by most awarding bodies. But what was significant was that students were not responding to predefined criteria, instead they created relevant headings and evidenced their interpretation and comprehension of them. Students showed the capacity to define capability and derive their own assessment criteria appropriate to alternative design solutions.

Holistic assessment enabled students to value a wide range of evidence that displayed their peers' capacity to be creative, to communicate, and to display their technological capability in this context. The strength of the ACJ model of assessment lays in its capacity to aggregate subjectivity, thus supporting diversity in designing. It is envisaged that ACJ would become a cornerstone of contemporary assessment and it is hoped that students would continue with a creative designerly approach to meeting learning outcomes, without being concerned with externally set assessment criteria and be confident in translating these values into their own professional practice.

References

Banks, F., Barlex, D., Jarvinen, E., O'Sullivan, G., Owen Jackson, G., & Rutland, M. (2004). DEPTH–developing professional thinking for technology teachers: An International Study. *International Journal of Technology and Design Education, 14*, 141–157.

Barlex, D. (2007). Creativity in school design and technology in England: A discussion of influences. *International Journal of Technology and Design Education, 17*, 149–162.
Barlex, D., & Trebell, D. (2008). Design-without-make: Challenging the conventional approach to teaching and learning in a design and technology classroom. *International Journal of Technology Design Education, 18*, 119–138.
Baynes, K. (2010). Models of change: The future of design education. *Design and Technology: An International Journal, 15.3*. The Design and Technology Association.
Carty, A., & Phelan, P. (2006). The nature and provision of technology education in Ireland. *Journal of Technology Education, 18*(1), 7–11.
Cross, N. (2001). Designerly ways of knowing: Design discipline versus design science. *Design Issues, 17*(3), 49–55.
Davies, A. (1996). Assessment and transferable skills in art and design. *International Journal of Art and Design Education, 3*, 327–331.
Dunbar, R. (2010). *Informing technology teacher education; exploring the effects of contemporary teaching and learning*. PhD thesis, University of Limerick.
Gibson, K. (2008). Technology and technological knowledge: A challenge for school curricula. *Teachers and Teaching: Theory and Practice*, *14*(1), 3–15.
Guilford, J. P. (1950). Creativity. *The American Psychologist, 5*(9), 444–454.
Kimbell, R. (2007). E-assessment in Project e-scape. *Design and Technology Education: An International Journal, 12*(2), 66–76.
Kimbell, R. (2010). The transient and the timeless: Surviving a lifetime of policy and practice in assessment. *Design and Technology: An international Journal, 15.3*, The Design and Technology Association.
Kimbell, R., & Stables, K. (2007). *Researching design learning*. Dordrecht NL: Springer.
Kimbell, R., Stables, K., Wheeler, T., Miller, S., Bane, J., & Wright, R. (2004). *Assessing design innovation*. London: Department of Education and Skills.
Kimbell, R., Stables, K., Wheeler, T., Wosniak, A., & Kelly, V. (1991). *The assessment of performance in design and technology*. London: School Examinations and Assessment Council/Central Office of Information.
Kimbell, R., Wheeler, T., Miller, S., Pollitt, A., (2007). *E-scape portfolio assessment phase 2 report*. TERU, Goldsmiths University of London.
Kimbell, R., Wheeler, T., Shepard, T., Browne Martin, G., Perry, D., Hall, P., et al. (2005). *E-scape portfolio assessment phase 1 report*. Teru: Goldsmiths University of London.
Kimbell, R., Wheeler, T., Stables, K., Shepard, T., Pollitt, A., Martin, F., et al. (2009). *E-scape portfolio assessment phase 3 report*. Teru: Goldsmiths University of London.
Lewis, T. (2009). Creativity in technology education: Providing children with glimpses of their inventive potential. *International Journal of Technology and Design Education, 19*, 255–268.
Mawson, B. (2003). Beyond "The Design Process": An alternative pedagogy for technology education. *International Journal of Technology and Design Education, 13*, 117–128.
McCormick, R., & Davidson, M. (1996). Problem solving and the tyranny of product outcomes. *The Journal of Design and Technology Education, 1*(3), 230–241.
McDowell, L., & Sambell, K. (1999). The experience of innovative assessment: Student perspectives. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 71–82). Buckingham: Open University Press.
Pollitt, A. (2004). *Let's stop marking exams*. Paper presented at the annual conference of the international association for educational assessment, Philadelphia.
Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching: The experience in higher education*. Buckingham: SRHE and Open University Press.
Rasinen, A. (2003). An analysis of technology education curriculum of six countries. *Journal of Technology Education, 15*(1), 31–47.
Ritz, J. M. (2009). A new generation of goals for technology education. *Journal of Technology Education, 20*(2), 50–64.
Schunk, D. (2004). *Learning theories: An educational perspective* (4th ed.). Upper Saddle River, NJ, USA: Pearson.
Stiggins, R. (2005). Rethinking the motivational dynamics of productive assessment. *Manitoba Association of School Superintendents Journal*, *5*(1), 8–12.
Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*, 273–286. Reprinted as Chapter 3 in L. L. Thurstone (1959), *The measurement of values*. Chicago, Illinois: University of Chicago Press.
Williams, P. J. (2000). Design: The only methodology of technology. *Journal of Technology Education, 11*(2), 48–60.