

2012

# Authentic digital representation of creative works for assessment

Christopher Newhouse  
*Edith Cowan University*

---

This article was originally published as: Newhouse, C. P. (2012). Authentic digital representation of creative works for assessment. Proceedings of Australian Computers in Education Conference (ACEC). (pp. 7). Perth, Australia. Australian Council for Computers in Education. Original article available [here](#)

This Conference Proceeding is posted at Research Online.

<http://ro.ecu.edu.au/ecuworks2012/123>

# **AUTHENTIC DIGITAL REPRESENTATION OF CREATIVE WORKS FOR ASSESSMENT**

C. Paul Newhouse

Associate Professor, School of Education, Edith Cowan University

## **Abstract**

*This paper reports on the first stage of a three-year project to investigate the representation of student practical work in digital forms for the purpose of summative assessment. In a number of senior secondary courses in Western Australia, such as Visual Arts and Design, students submit creative practical work for summative assessment that takes various forms dependent on the chosen context. The collection and judging/marketing of this work has increasingly been considered to be problematic by assessors, teachers and probably many students. While it should be feasible to digitize this work and apply a similar method of marking, the digitization process and output would need to be shown to authentically represent the original work of the student. These portfolios of practical work can vary considerably in form, structure and purpose, making reliable assessment very difficult and results from previous research indicate that this problem may be better addressed through digitization. The first phase of the project involved creating digital representations of the practical creative work submitted by a sample of students in the Year 12 Visual Arts and Design courses and comparing the results of marking these with the physical forms.*

## **Introduction**

In most disciplines creativity and practical application of knowledge and skills are important facets of learning. Typically these facets are also important outcomes for society and require higher-order thinking and the application of learning processes. Many educational researchers argue that traditional assessment fails to assess learning processes and higher-order thinking skills and argue that digital technologies may address this problem (Lane, 2004; Lin & Dwyer, 2006). Further, it often means that creative practical applications of knowledge and skills are not assessed in preference for theoretical applications that are easy to be assessed by traditional paper-based tests or examinations. This argument is really about the validity of the assessment in terms of the intended learning outcomes, where there is a need to improve the criterion-related validity, construct validity and consequential validity of high-stakes assessment (McGaw, 2006). The use of portfolios has typically been offered as one means of addressing these issues but has presented some obstacles, particularly in terms of manageability and measurement reliability (Clarke-Midura & Dede, 2010). This has limited their use in high-stakes assessment. However, in many courses with a major component based on practical performance a portfolio is used for formative and school-based summative assessment. Therefore if some of the obstacles to using portfolios for high-stakes assessment can be overcome by using digital technologies and modern psychometrics then this will better align assessment with preferred pedagogy (Clarke-Midura & Dede, 2010; Lane, 2004; Ridgway, McCusker, & Pead, 2006).

The purpose of the study reported in this paper is to investigate the efficacy of digitization and the paired-comparisons method of judging of portfolios of creative practical work for the purposes of summative assessment in the Visual Arts and Design senior secondary school courses. It is designed to build upon the success of five-years of research conducted by some of the chief investigators from 2006 to 2010, with the WA Curriculum Council (Newhouse, 2010), collaboration with researchers in the British e-Scape project (Kimbell, Wheeler, Miller, & Pollitt, 2007), and build on the work in tertiary education by Dillon and Brown (2006). The Centre for Schooling and Learning Technologies (CSaLT) has worked with the Curriculum Council of WA to investigate using digital technologies to improve the assessment of student performance in courses with a major practical component (Jones, Penney, Newhouse, & Campbell, 2009). One of the outcomes of this research has been the successful application of a paired-comparisons method of marking to the assessment of digital portfolios (Newhouse & Njiru, 2009). These digital portfolios were either created by the students themselves or captured directly as students performed.

The new study addresses the problem of effectively and efficiently assessing student portfolios that include artefacts (e.g. paintings, sculpture, drawings) and process documents and are submitted for high-stakes summative assessment purposes. The aim is to evaluate the feasibility of digitizing these portfolios and judging them using online tools and the comparative pairs method of marking. The study connects with three major areas of knowledge: portfolio assessment; psychometrics, and computer-supported assessment.

Portfolio assessment is not new and is regularly used for low-stakes or formative assessment purposes. Its use for high-stakes summative assessment has been considerably less prevalent and problematic. For example, Koretz (1998) analysed the outcomes of four large-scale portfolio assessment systems in the USA school systems and concluded that overall the programmes varied in reliability and were resource intensive with “problematic” (p.309) manageability. It is part of what Messick (1994), calls “performance-and-product assessment” (p. 14) that can be traced back at least to the 1960s. He distinguishes between a performance, which concerns processes and procedures, and a product that is a remaining outcome. He goes on to argue that the extent to which each should be included in a performance assessment depends on the extent to which task procedures may be clearly determined and varied. In the Visual Arts course the focus of the portfolio assessment is on the product whereas for the Design course the focus is on the processes and procedures.

Psychometrics is the field of measurement of psychological (mental) attributes (processes) (Barrett, 2003). As such it is concerned with quantifying mental processes that are typically considered to be qualitative in nature, which as Barrett discusses, makes it a problematic and controversial field. It is a critical field of research for summative performance assessment, particularly in the arts where assessment necessarily relies on subjective judgements. In fact as Barrett discusses there is a sense where there are perceived to be dichotomies between psychometrics, scientific measurement and creative human activity (as occurs in education). Humphry and Heldsinger (2009) discuss this dilemma in the use of rubrics for analytical marking in performance assessment and the application of Rasch modeling. Pollitt (2004), a psychometrician, argues that it is unlikely that such a process of adding up scores on a set of criteria will accurately measure a student’s “performance or ability” (p. 5) and that more holistic judgements of performance are required. While Pollitt (2004) describes the paired comparison judgement method as “intrinsically more valid” and better than the traditional system, he believes that without some ICT support it has not been feasible to apply due to time and cost constraints, and he does suggest that further research is required to determine the appropriateness and whether “sufficient precision can be achieved without excessive cost” (p. 16).

Computer-supported assessment encompasses a range of applications of computers from the whole assessment process to assisting in one aspect of assessment (e.g. marking) (Bull & Sharp, 2000). Much of the published research in the field relates to higher education (e.g. Brewer, 2004). In the school sector there has been some use of portfolios for assessment but usually a physical portfolio and most often in the creative arts (e.g. Madeja, 2004). There has been increasing international interest in the application of computer support to improve assessment. It was the focus of a recent keynote address by McGaw (2006) in which he argued that with computer support and modern psychometrics summative assessment could be better aligned with intended curriculum outcomes and preferred pedagogies. The proposed study has a focus on digital representation of particular forms of creative expression for the purpose of summative assessment. Dillon and Brown (2006) have identified issues and developed protocols in the use of ePortfolios in a number of areas of the creative arts. They began with the question concerning what “constitutes knowledge in the discipline” (p. 430), then consider how this “knowledge can best be represented in media” (p. 430) before determining technical requirements such as file format, size and sensory quality.

There is a critical need for research into the use of portfolio assessment of performance on complex tasks that is feasible within the constraints of school systems (Clarke-Midura & Dede, 2010). This study investigates supporting portfolio assessment through digitisation to deliver high levels of reliability and manageability and the capacity to be scaled-up for statewide implementation in a cost

effective manner. The study builds on concerns that in courses where portfolio assessment is used judgements are not comparable between contexts, are not reliable due to the subjectivity of assessors, and are not cost-effective for large groups of students spread across large jurisdictions. Traditional modes of assessment for visual arts in Western Australia have relied heavily upon the physical examination of artefacts (visual diaries and artworks). The limitations of this approach includes the challenges associated with physical handling of bulky diary artefacts; the time constraints for receiving, sorting, processing the materials into examination bundles for scrutiny by examiners; issues related to storage of materials during the assessment period; absence of raw source material at later points for confirmatory purposes; and the potential for damage (or loss in transit) of critical objects prior to the physical examination.

## Method

The study is evaluative in nature set within an ethnographic framework in that activity must be considered to occur within learning environments where the characteristics of teachers and students and the culture created are critical to an understanding of all aspects of the curriculum and pedagogy, including assessment. The research design is an ethnographic action research methodology using interpretive techniques involving the collection of both qualitative and quantitative data. It involves at least two development-evaluation cycles that actively involve teachers in each cycle or phase and collect data to analyse the perspectives of the key groups of participants (teachers, assessors and students). These data are compiled into case studies within a multi-case approach in which each case is defined by the course.

The first one-year *Development and Pilot* phase of the study was completed between July 2011 and June 2012. In this phase, methods were developed for representing as digital files the range of portfolios typically created by students in the two courses. The efficacy of the digital representations created was evaluated, including a comparison with the marking of the original physical portfolios and interviewing teachers and students. An online digital system was used to facilitate the marking of a stratified sample of about 80 portfolios per course by analytical and comparative pairs methods of marking. The study involves samples of practical work drawn from both courses that are representative of the main types of submitted. This was determined initially through conducting a situation analysis that included a review of syllabus requirements and a review of portfolios submitted in the previous year. Purposeful sampling was used to select teachers for each of the courses who are experienced in teaching the courses. Many were also assessors.

The second one-year *School-Based Implementation* phase of the study has commenced to investigate the feasibility of creating the digital representations in schools for upload into a central digital repository for marking. It is suggested by Dillon and Brown (2006) that more accurate representation of creative arts portfolios is achieved where the portfolios are developed by students. In this phase a large representative sample of student work in Year 11 will be assessed using the same online digital marking systems as in Phase One. In the second year a stratified sample of portfolios will be selected from the two courses based on an understanding of the main types submitted. This will involve working with teachers in schools to identify appropriate portfolios and methods of digitization.

A range of types of quantitative and qualitative data was collected from each group of stake-holders in Phase One. These included surveys, interviews, reports, and student work output. These data are being analysed within a feasibility framework of the four dimensions based on the British e-Scape project (Kimbell et al., 2007): Manageability; Technical; Functional (Validity and Reliability); and Pedagogical. It is appreciated that there is a tension between these dimensions, in particular as Stobart (2008) explained that with the 'one-handed clock' improvements in one dimension come at a cost to one or more of the others.

## Results of First Phase of Study

The first phase of the study was conducted between July 2011 and June 2012. At the time of writing some of the preliminary analysis had commenced but had not been completed. A sample of 11 Visual Arts classes and 6 Design classes and their teachers were involved in the first phase of the study. From these classes a total of 82 students from the Design classes and 75 students from the Visual Arts classes were involved in the study. The practical work of these students, submitted for tertiary entrance examination, were digitized by the research team at the location to which they were delivered, prior to the work being officially marked.

Prior to the digitisation of the student work a number of working meetings were held with the research team and discipline experts to develop a set of procedures and guidelines. These were relatively simple for the Design course submissions that were presented as a set of about 14 A3 pages that needed to be colour scanned into a PDF file. However, for the Visual Arts course submissions they were relatively complex and as a result they were successfully piloted with the work from a class of Year 11 students. The final set of procedures and guidelines used for digitization in the first phase of the study are provided in the table below.

Table 1

*Procedures and guidelines for the digitising of the practical submissions.*

Type	Requirement	File
<b>Design</b>	Colour scan of all A3 sheets	PDF
<b>Visual Arts</b>		
2 Dimensional	ID number visible in each photo/video	
	Photo of 'Artist Statement' and proposed installation if provided	JPG
	Full size photo (Hi-res 300dpi) + match-box for size. Four megapixels for all photos giving 300dpi at a reasonable size. 72 dpi for on-screen viewing.	JPG
	4 x close ups - extracted from main photo	JPG
	All photos combined into one document	PDF
	HD Video (pan & zoom) - 10 secs	WMV
	'Amateur' snaps using mobile	
3 Dimensional	ID number visible in each photo/video	
	Photo of 'Artist Statement' and proposed installation if provided	JPG
	Full size photo + size object such as a match-box	JPG
	At least 4 x angle photos (L, R, top, bottom)	JPG
	4 x close ups - extracted from main photo	JPG
	All photos combined into one document	PDF
	HD Video (pan & zoom) - 10 secs	WMV
	3-D Animation for some works	MOV
'Amateur' snaps using mobile	JPG	

Due to severe time constraints and limitations of space it was not possible to fully implement the intended procedures and guidelines for digitizing the Visual Arts work, however, the best attempt was made to follow these as closely as possible. It was not possible to set up lighting equipment or backdrops, much of the 3-D work was too large or delicate to move, most of the 2-D work needed to be placed on easels with a slight lean backwards, and in some cases not all of the pieces of a students submission could be found. Further, due to time constraints placed by external authorities (less than one day to record 75 submissions) typically photographs and videos could not be checked and retaken, and colour balancing could not be changed to suite individual pieces of work. However, for the scanning of the Design portfolios there were only minor issues such as not being able to use the sheet-feeder for some (photographic, delicate or laminated paper), but eventually all were able to be converted successfully to PDF files.

After the initial digital capture there was some editing required to prepare the files for marking. For the Design PDF files the only editing required was that some pages had to be rotated. For the Visual

Arts files most needed some cropping and some rotating. For each submission four close-up images from the photographs were created using Photoshop. The Visual Arts specialist on the research team selected the four areas for the close-ups by viewing print-outs of the main images. All the original photograph files and the close-ups were combined in a single PDF file using a Powerpoint slideshow template as an intermediary. The video files were converted to both WMV and MOV formats to accommodate the computer systems markers were likely to be using. A folder was created on the server for each student into which all the files for his/her submission were placed using standard file names.

Eventually there were 75 Visual Arts and 82 Design student folders on the server. Each student's folder was copied to a USB Flash Drive and given to the student at school to view on a computer while they completed the survey on paper to illicit their perception of the representation of their work, their attitudes towards the process and their experience and skills in using computers. The Design students were generally positive about the concept of digitally representing their portfolios and would prefer to submit them digitally. In fact many students had created the portfolios in some digital form originally. The Visual Arts students were less positive with many concerned that the digital representations did not adequately represent their artwork. Some believed that there was value in creating digital representations in addition to the physical objects.

The folders for all the students for each teacher were placed on a CD-ROM for each teacher for them to view at their leisure. A week or so after delivery the Visual Arts teachers were interviewed face-to-face and the Design teachers were sent similar questions via email. These questions followed similar purposes to those asked of the students. Teachers of the Design course were generally enthusiastic about the concept of students submitting digital representations of their portfolios. Teachers of the Visual Arts course were generally negative about the concept of submitting digital representations of the students' work and were very enthusiastic about maintaining the current system of submitting the physical artworks to a central location. It should be noted that there were no teachers from country schools involved in Phase One of the study.

Experienced assessors were selected to mark the digital representations of the students' work using an analytical method through an online system. The system was developed in-house using the *Filemaker Pro* database system. This could be accessed through an Internet browser, once logged on an assessor could select each student, viewing the digital representations of their work, and click on buttons to record marks for each criterion. Three assessors marked all the Visual Arts work and two assessors marked all the Design work. The same criteria and marking scheme, through a rubric, was used as was used for the tertiary entrance examination for each course. Assessors initially came to the University for an hour or so to be trained and supported in the use of the system. They then completed the marking at home or work. Teachers were also asked to send a class ranking for the work submitted by their students.

The digital representation files for each student were also uploaded into the 'Pairs Engine' comparative pairs online marking system provided by TAG Learning. For Visual Arts each file had to be initially uploaded into the MAPS online portfolio system as journal entries whereas for Design the PDF files could be simply copied to a folder on the 'Pairs Engine' server. The teachers involved in the study, the analytical markers, some of the research team, and some officers from the Curriculum Council were invited to participate in the comparative pairs marking. For each course an initial half-day workshop was held at the University to train the assessors in the use of the system and to allow the assessors to discuss the criteria and basis for their judgements with the aim of developing a more common understanding between them. After this workshop these assessors completed their judgements using the online system from their homes or places of work. At the time of writing the analysis of the results of marking had not been completed.



## Conclusions

In Australia the agenda of developing a national curriculum has been building, particularly to provide comparability of assessment. This has implied the need for initiatives and strategies designed to meet standards and improve the consistency and authenticity of assessment across the country. This study is beginning to provide guidelines and examples to assist curriculum authorities and educators in developing strategies to digitize portfolios for the purpose of summative performance assessment across a range of areas of the curriculum. The outcomes of this study will: improve the manageability of assessment processes (e.g. marking from anywhere, less use of physical space and time); increase the validity and reliability of the scores from marking; readily maintain an enduring record of student work; improve comparability over time and between media contexts and course contexts; provide knowledge of assessor perceptions and training requirements for assessors; and with the inclusion of developmental input (e.g. photos, videos, audio annotations) further improve the authenticity of Visual Arts and Design courses portfolios.

## Acknowledgement

The theory discussed in this paper and the research upon which it is based are as a result of the work of a research team organised by the Centre for Schooling and Learning Technologies at Edith Cowan University. The team was led by Paul Newhouse and includes researchers Jeremy Pagram, Lisa Paris, Mark Hackling, Christine Ure, Alun Price, Martin Cooper and Alistair Campbell, and a number of research assistants and advisors.

## References

- Barrett, P. T. (2003). Beyond Psychometrics: Measurement, non-quantitative structure, and applied numerics. *Journal of Managerial Psychology*, 3(18), 421-439.
- Brewer, C. A. (2004). Near Real-Time Assessment of Student Learning and Understanding in Biology Courses. *Bioscience*, 54(11), 1034.
- Bull, J., & Sharp, D. (2000). *Developments in computer-Assisted Assessment in UK Higher Education*. Paper presented at the Learning to Choose: Choosing to Learn., Queensland, Australia.
- Clarke-Midura, J., & Dede, C. (2010). Assessment, technology, and change. *Journal of Research on Technology in Education*, 42(3), 309-328.
- Dillon, S., & Brown, A. (2006). The art of ePortfolios: insights from the creative arts experience. In A. J. a. C. Kaufman (Ed.), *Handbook of Research on ePortfolios* (pp. 420-433). Hershey PA: Idea Group Inc.
- Humphry, S., & Heldsinger, S. (2009). *Do rubrics help to inform and direct teaching practice?* Paper presented at the Assessment and Student Learning: Collecting, interpreting and using data to inform teaching.,
- Jones, A., Penney, D., Newhouse, C. P., & Campbell, A. (2009). *Digital assessment in high stakes physical education practical examinations*. Paper presented at the 26th ACHPER International Conference, Brisbane, Queensland.
- Kimbell, R., Wheeler, T., Miller, A., & Pollitt, A. (2007). *e-scape: e-solutions for creative assessment in portfolio environments*. London: Technology Education Research Unit, Goldsmiths College.
- Koretz, D. (1998). Large-scale portfolio assessments in the US: Evidence pertaining to the quality of measurement. *Assessment in Education*, 5(3), 309-334.
- Lane, S. (2004). Validity of High-Stakes Assessment: Are Students Engaged in Complex Thinking? *Educational Measurement, Issues and Practice*, 23(3), 6-14.
- Lin, H., & Dwyer, F. (2006). The fingertip effects of computer-based assessment in education. *TechTrends*, 50(6), 27-31.
- Madeja, S. S. (2004). Alternative assessment strategies for schools. *Education Policy Review*, 105(5),

3-13.

- McGaw, B. (2006). *Assessment to fit for purpose*. Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23(2), 13-23.
- Newhouse, C. P. (2010). Aligning Assessment with Curriculum and Pedagogy in Applied Information Technology. *Australian Educational Computing*, 24(2), 2-5.
- Newhouse, C. P., & Njiru, J. (2009). Using digital technologies and contemporary psychometrics in the assessment of performance on complex practical tasks. *Technology, Pedagogy and Education*, 18(2), 221-234.
- Pollitt, A. (2004, June 2004). *Let's stop marking exams*. Paper presented at the International Association for Educational Assessment Conference, Philadelphia.
- Ridgway, J., McCusker, S., & Pead, D. (2006). *Report 10: Literature Review of E-assessment*. Bristol, UK: Futurelab.
- Stobart, G. (2008). *Testing Times, The USes and Abuses of Assessment*. Abingdon: Routledge.