1-1-2013

Digital representation of visual artworks for high-stakes assessment

Christopher P. Newhouse
Edith Cowan University

Newhouse, C. P. (2013). Digital representation of visual artworks for high-stakes assessment. In Proceedings of the 21st International Conference on Computers in Education (pp. 832-839). Bali, Indonesia. Uhamka Press. More information on conference here

Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Digital Representation of Visual Artworks for High-Stakes Assessment

C. Paul NEWHOUSE
Associate Professor, School of Education, Edith Cowan University
p.newhouse@ecu.edu.au

Abstract: The collection and marking of student artwork across a large jurisdiction such as Western Australia is challenging where the work is submitted to a central location to be marked by experts and returned to students. An alternative approach would be to submit digital representations of the artworks online for marking. However, to give a valid and reliable measure the representations would need to be of adequate quality. Further, judgements of artworks are necessarily subjective giving concern about the reliability of marking for high-stakes assessment. The comparative pairs method of marking lends itself to addressing this problem and is feasible where the work to be marked is in digital form. This paper reports on one component of a three-year study to investigate the representation of student practical work in digital forms for the purpose of summative assessment. This study set out to determine whether the digital approach was feasible and adequate fidelity could be achieved in order to use the comparative pairs method of marking. The first phase of the project involved the researchers creating digital representations of the artwork submitted at the end of secondary schooling by a sample of students in the Visual Arts course and comparing the results of marking these with the physical forms. The second phase involved a sample of students creating digital representations of their own work and submitting them through an online system for marking. The study found this process was feasible, and the results were acceptable, but it lacked support from teachers and students who wanted the original artworks to be assessed.

Keywords: Visual Arts, summative assessment, portfolio, digital representation

1. Introduction

Creative expression is difficult to assess largely due to the subjective nature of the judgements made (Dillon & Brown, 2006). Where these judgements contribute to high-stakes assessment there are concerns about the reliability of resulting scores. Recent research has suggested that a comparative pairs method of marking may be the best approach to counter the subjectivity of judgements (Heldsinger & Humphry, 2010; Kimbell, Wheeler, Miller, & Pollitt, 2007; Newhouse, 2011; Pollitt, 2012). This method is based on making comparisons between pairs of artefacts or representations of performance and combining the results of these comparisons using a Rasch dichotomous model (Pollitt, 2012). However, to effectively apply this method to a large sample realistically requires that the work being judged is in a digital form so that computer software and networks can be used to access the work, enter judgements and calculate scores. For the results to be valid these digital representations of student work or performance must have an adequate level of fidelity for the purposes of the assessment. At the same time the creation and submission of these digital representations must be manageable and at minimal cost. Can this be accomplished for senior secondary school high-stakes summative assessment for a highly creative area of the curriculum?

This paper reports on a component of a three-year study that sought to address this question. This study used the practical work in two senior secondary courses in Western Australia (W.A.), Visual Arts and Design, however, this paper only considers findings associated with the Visual Arts course. The study built on five-years of research into the use of digital technologies to support high-stakes summative assessment, which had included investigation of the comparative pairs method of marking in collaboration with researchers from the British e-Scape project (Kimbell et al., 2007; Newhouse & Njiru, 2009). The main purpose was to determine the efficacy of digital representations

of student artwork for the comparative pairs method of marking for the purposes of high-stakes summative assessment. The study built on three main intersecting areas of knowledge: assessment; psychometrics; and digital representation. The processes and resulting artefacts (e.g. paintings and sculptures) of student activity needed to be represented in digital form to measure their performance for summative assessment purposes.

The assessment of practical creative expression, such as for Visual Arts, has typically been done by students submitting a portfolio of work that includes created artefacts and process documents (e.g. Madeja, 2004). This portfolio is then judged by an expert against a set of criteria and awarded a score or grade. This approach has been reasonably effective for small-scale formative assessment purposes such as with a teacher and her class. However, when the scale is larger and the stakes are higher, management and measurement reliability become obstacles (Clarke-Midura & Dede, 2010). This study sought to address management obstacles through using digital technologies and reliability obstacles through applying modern psychometrics (Clarke-Midura & Dede, 2010; McGaw, 2006; Ridgway, McCusker, & Pead, 2004).

Using portfolios for assessment is part of what Messick (1994) refers to as "performance-and-product assessment" (p. 14) where a performance concerns processes and a product is a remaining outcome. In the Visual Arts course in W.A. the existing focus of the final assessment was on the product, with the process represented in a minor subsidiary form, although some would argue it should be as important as the product (Dillon & Brown, 2006). Dillon and Brown point out that in the visual arts while the product may be "tangible" the "meaning may not be clear or literal" (p. 421) so this needs to be captured in representations of the process. They also highlight the assessment problem of creating "representational and evaluative" frameworks that help identify differences in both "technical and expressive ability". That is so the assessor is provided with adequate evidence to make a balance of judgements between the technical and expressive quality of the work submitted as representative of the student's "artistic" knowledge. That is why in our project we used a combination of photographs, video, audio and text to represent the student's work, with the artist's "voice" (p. 419) in the audio and text providing some explanation of the process and meaning.

Traditionally portfolios of creative work have been assessed for summative purposes using a range of analytical techniques that aim to quantify identifiable aspects or qualities of the work. This is much easier to accomplish for technical qualities than expressive (creative) qualities that are difficult to describe alphanumerically and detached from the assessor (Dillon & Brown, 2006). Psychometrics is the science that focuses on measuring such mental processes, quantifying the qualitative (Barrett, 2003). An analytical technique used by psychometricians is to describe each assessable quality as a criterion with a set of quantified levels of performance or achievement, often represented in a rubric, and then use some form of Item Response Theory such as Rasch modeling to generate a score or grade (Humphry & Heldsinger, 2009). Although this is preferable to just adding up the scores on the criteria, psychometricians such as Pollitt (2004) argue that this will not accurately measure a student's "performance or ability" (p. 5) because the nature of the performance is holistic and therefore a holistic method such as the paired comparison judgement method is "intrinsically more valid". However, until recently this has not been feasible for large-scale assessment but may be with the use of purpose built software, computer networks and digital representations of performances (McGaw, 2006; Pollitt, 2012).

Dillon and Brown (2006) argue that, "Digital media and information systems present the opportunity to capture, store, and manage multiple forms of evidence about artistic products and processes" (p. 420). The question is how can the knowledge and skills being assessed best be represented in digital media to allow comparative judgements to be made? Then the question is whether this can be accomplished in typical schools with their accompanying constraints? Our study aimed to address these two questions.

2. Method

Our study addressed the questions of adequate digital representation and scalability to schools in two phases. In the first phase we digitised student work that was submitted for high-stakes summative assessment, and in the second students digitised their own work to be submitted online for external assessment. In the first phase we compared the results of marking the digitised artwork portfolio with the results of marking the physical artwork portfolio. In Western Australia (W.A.) for many years

Visual Arts in senior secondary schooling has been assessed through analytically judging a physical portfolio that at one stage was more a representation of process, termed a 'visual diary', but more recently is largely the physical artworks themselves supported by a document containing an 'artist statement' and photographs of intended presentation of the artwork. In a large jurisdiction such as W.A. this approach provides many logistical and management challenges to add to the limitations of measurement and the lack of an adequate enduring record for confirmatory purposes.

The study employed an action research design with the two development-evaluation phases that involved the collection of a range of data analysed from the perspectives of students, teachers and assessors. Students were surveyed and interviewed, teachers and assessors were interviewed, and the scores from marking were analysed and compared. The analytical marking criteria provided for the course were used by the study and initially the technical specifications for digitizing the portfolios were determined through an analysis of syllabus requirements and a review of portfolios submitted in the previous year.

The first one-year *Development and Pilot* phase occurred from July 2011 to June 2012, and involved 75 Visual Arts portfolios from 11 schools being digitised by our research team and marked using both an analytical and a comparative pairs method. The efficacy of the digital representations was interrogated through interpreting the responses of students, teachers and assessors, and through a comparison with the scores awarded to the original physical portfolios (Note: The awarding body provided us with the raw scores from the official marking for the students in our sample.). The sample was purposefully selected to ensure all teachers were experienced in teaching the *Visual Arts* course having taught the course for a few years; many were also experienced external assessors.

The second one-year *School-Based Implementation* phase occurred between July 2012 and June 2013, and involved 138 students from the penultimate year of 13 secondary schools. This sample of schools was initially selected to ensure a representative range of typical schools were involved including two from country areas and some from each of the three school systems (government, Catholic and independent). Researchers supported teachers to facilitate students in digitizing their own portfolios and uploading these to an online repository. As Dillon and Brown (2006) suggest it was likely that with students digitizing their own work a more accurate representation would be formed.

3. Phase One – digitisation by researchers

In this phase of the study the artworks submitted by final year secondary students in the sample for final examination were represented digitally by the researchers to allow online scoring. This was a separate set of processes from the official scoring of the artworks for the purposes of graduation and tertiary education entrance. Prior to digitization a set of procedures and guidelines for digitising were drawn up after consultation with experts to review the course syllabus requirements and examples of student two or three-dimensional (2D or 3D) artworks. These were then tested with work produced by students from a Year 11 class at a local school and refined to give those listed in Table 1. Three teams of researchers were trained in using SLR and digital video cameras to digitise the portfolios at the central location to which all artwork portfolios for the state of W.A. had been delivered. Unfortunately it was not possible to fully implement the procedures and guidelines as intended because we were only given one day to access the work and the location had very little space available. Therefore backdrops and lighting could not be set up, most 3D artworks could not be moved, and time did not permit proper colour balancing and multiple attempts at photographs and videos. However, despite the constraints a full set of digital files was created for each artwork portfolio. Each portfolio included either 2D or 3D artwork, almost all included a one-page artist statement and most included a photograph of the intended presentation. A substantial number of the artworks comprised multiple pieces with the maximum number being 11 interconnected two-dimensional paintings.

Table 1: Procedures and guidelines for the digitising of the Visual Arts portfolios by the researchers.

Artwork type	Digitisation Requirement	File type
2 Dimensional	ID number and match-box visible in each photo/video	
	Photo of 'Artist Statement' and proposed installation if provided	.jpg
	One full size hi-resolution photo of 4 megapixels giving 300dpi at a reasonable size. Additional photos for multi-piece artworks.	.jpg
	4 x close ups - extracted from main photo(s)	.jpg
	All photos combined into one document	.pdf
	HD Video (pan & zoom) - 10 secs	.wmv
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 seconds	.wmv
	3-D Animation for some works	.mov

Prior to uploading the digital files into the online repositories to be accessed by the assessors some simple editing was required that included rotating and cropping, and for the videos a change of file format to WMV. The four close-up images were created from the original photographs based on advice from a Visual Arts teaching specialist. In addition a single PDF file was created using all the original photographs and the close-up images. All the files for each student were copied onto USB flash drives and given to them to review prior to completing a questionnaire. This process also checked that the correct files were associated with each student prior to all files being copied to a server for analytical marking, and uploaded to the MAPS online portfolio system for comparative pairs marking.

Three experienced assessors were employed for analytical marking of the digital representations and they were augmented with the teachers and some officers of the awarding body for comparative pairs marking. The marking criteria for analytical marking of the digital representations were those used for the official marking of the physical artwork portfolios as shown here (range of scores points are in parentheses).

C1: Creativity and innovation: Artwork/s is outstanding, showing exceptional creativity and innovation and the emergence of a distinctive style. (0-6)

C2: Communication of ideas: Ideas are skilfully realised and powerfully communicated in sophisticated and highly coherent resolved artwork/s. (0-5)

C3: Use of visual language: Extensive and sophisticated application of visual language in the artwork/s. Complex and highly resolved visual relationships are evident. (0-12)

C4: Use of media and/or materials: Highly discerning selection and refined use of media and/or materials demonstrating sensitive application and handling. (0-5)

C5: Use of skills and/or processes: Extensive and sophisticated selection and application of skills and processes. (0-12)

From these criteria one holistic criterion to be used in the comparative pairs marking of the digital representations was generated at a workshop involving all the assessors.

Judgement about performance addresses students' ability to creatively use visual language, materials and processes to skilfully communicate an innovative idea in a resolved artwork.

Both methods of marking the digital representations were facilitated through online systems accessed through an Internet browser. For analytical marking the research team developed a database system using *Filemaker Pro* that displayed the digital representations and the rubric, and allowed scores to be recorded using on-screen buttons. For comparative pairs marking the *Adaptive Comparative Judgements System* (ACJS) associated with the MAPS online portfolio system was used. This system, as described by Pollitt (2012), facilitates all aspects of the comparative pairs method. Assessors were trained in the use of the system at the workshop where the holistic criteria was generated.

Teachers were asked to view the digital representations of their own students' work, provide a ranking, and then answer some questions. Overall they were negative about using digital representations to replace marking of the physical artworks. They believed that assessors needed to be able to 'touch' the artwork and get a feel for the 'size' of the work. Further they believed that some artwork would not photograph well due to the materials and techniques used and that three-dimensional work would not be adequately represented. From the survey of students it was also clear that they did not think the digital representations were adequate to mark in place of the original artworks. However, an analysis of the resulting scores showed a high significant correlation between scores from both methods of marking the digital representations and the official scores awarded for the original physical forms ($r=0.86$ for analytical and $r=0.74$ for comparative pairs, $p<0.01$). In fact these correlation coefficients were much larger than those between the three analytical assessors of the digital representations ($r=0.51$, 0.54 and 0.56, $p<0.01$).

4. Phase Two – digitisation by students

For the second phase students were assisted to create a digital representation of their own work, in a similar manner to that employed by the researchers for the first phase of the study. They used the technical specifications shown in Table 2, with the only major change from the first phase being the opportunity to explain their work to the assessors as they recorded their video recording. They were instructed to have a 2D or 3D artwork ready and they were guided on how to use the SLR digital and video cameras prior to digitisation. They were able to do this in the close-up photos and the video. To effectively digitise the artworks a backdrop support and lighting kit was used with the cameras. The ePhoto kit (ePhoto Inc., 2012) used was inexpensive at less than $150 per kit and was purchased online (Figure 1). It included two light stands and umbrellas, two 45w day light bulbs (5500K/2 x 32"), one black and one white muslin backdrop, and one background support system. As the system was used in Australia, the four light bulbs that came with the kit were replaced with 105W 5500K bulbs and light holders. The backdrop and lighting system was easy to store, transport, set up and use. Students could easily be instructed on how to set up the system and their work for digitisation.

Table 2: Technical specifications provided for Visual Arts students and teachers (abridged version).

Component of digital representation	File type
ARTIST STATEMENT	
• 300 words containing the rationale for the conceptual and material development	.doc
DIGITAL PHOTOGRAPHS	
• Backdrop stand and sheet is to be used for the digital photos and the video supported by provided camera/video lighting.	
• Room natural or artificial light, ensuring that there are no shadows.	
• Photos need to match the colour of the artwork.	
• Ruler needs to be visible to indicate size of the work.	
• Photographs need to be in focus.	
• High resolution, 4 megapixels to give 300dpi at a reasonable size. 1,000 pixels on longest side and/or up to 5MB in size.	
MAIN PHOTO – 2D	
• 2D artwork needs to be completely upright.	.jpg
• 1 full size main photo of the work using a ruler as a size guide (photo 1).	
• If the photo does not capture the work, a second photo can be taken.	
MAIN PHOTO – 3D	
• 1 full size photo taken from the front of the work including a ruler as a size guide	
• 4 full sized angle photos (left, right, top, bottom)	
CLOSE UP PHOTOS	
2D – 2 photos of own choice to highlight two aspects of the work	.jpg
3D – 4 photos of own choice to highlight four aspects of the work	
VIDEO – 2D & 3 D	
• Up to 12MB of HD Video (pan and zoom) - 20 seconds with audio annotations	.avi

Figure 1: Backdrop, stand and lighting kit system (ePhoto Inc., 2012).

Students completed a questionnaire about the creation of the representations of their artwork, and their attitudes towards, and perceptions of, the digital representation of their artwork. It was clear from their responses that they had little previous experience in representing their artwork digitally with 86% indicating little or no experience. Further, 79% indicated that previously none of their work had been submitted in digital form for assessment. In fact, very little of their work had been originally created on computer but 55% indicated that they needed little or no help to digitise their artwork. They were split evenly on confidence in digitising but 70% thought it was easy to do the digitisation and 81% believed the requirements were easy to follow. The majority (62%) felt that the photographs and video represented their artwork well. However, 55% would have preferred someone else to do the digitising, 87% would still prefer an assessor to mark their original artwork, and 80% believed that in doing so they would receive a higher score.

There were open-response items that allowed them to identify the best and worst things about creating the digital representations of their work. Most stated that the process was “easy” and “fun” and they thought it was good to have digital copies of their work. Some had a number of pieces that made up one artwork and they found it helpful to capture all of the pieces together. They also liked being able to take close-up photographs and talk on their video to help them foreground details of their work. Most were happy with the quality of the photographs even to the extent that some thought they could make the work look better than it really was. Many comments also indicated that they thought the digital versions would be easier to mark and would save on transportation, storage and damage. However, there were a number of things about the digital representations that they did not like such as that the video was too short and that they did not enjoy talking on the video. Some argued that their original work would have “more impact” and look more “impressive” and that digitisation wasn’t a true representation. They also felt concerned that digitisation did not show “fine details, texture and doesn’t capture the essence” of the original artwork.

Overall the teachers were still not in favour of digital representations being marked in place of the original artwork but some made positive comments about the outcomes of the digitisation by their own students. One teacher stated: ‘I was very against it [digitisation] but after some terminal damage to work for external examination I can see the benefit from a wear and tear point of view’. They were more impressed with the quality of the photographs but still felt they missed some “subtle nuances”. Generally they believed that the original work would score higher because the viewer could “interact” with the work, the scale of the work was more obvious, and that the ‘digital representations do tend to flatten artwork and work seems to lose its tactile quality’.

6. Conclusions

At this stage our study has shown that visual artworks can be adequately digitised for the purposes of summative assessment and that students can do this using relatively inexpensive equipment, systems and software. However, teachers and most students are not persuaded of the adequacy, although if faced without the choice of centralised marking of the physical forms of the work this may change. The lack of experience of students, probably because there is currently no need to represent their work digitally, is a likely explanation of the negative perceptions of many students. In Australia as the trend continues towards a national curriculum with accountability requiring comparability of assessment it is almost certain that online judgement or marking systems will need to be used to be cost-effective. This approach will improve manageability (e.g. marking from anywhere, less use of physical space and time), increase the reliability of the scores from marking, maintain an enduring record, and provide knowledge of assessor perceptions. Using various forms of digital portfolios this can probably be achieved in any curriculum area, even difficult areas such as visual arts. Our study is demonstrating the viability of the technology so now political will and informed community attitudes are needed to make the key decisions to move all aspects of assessment into the digital age.

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 (CSaLT) at Edith Cowan University. The team was led by Paul Newhouse and includes researchers Jeremy Pagram, Lisa Paris, Mark Hackling, Alun Price, Pina Tarricone, Martin Cooper and Alistair Campbell, and a number of research assistants and advisors. The research was conducted with the Schools Curriculum and Standards Authority in Western Australia and partly funded by the Australian Council for Research.

References

Barrett, P. T. (2003). Beyond Psychometrics: Measurement, non-quantitative structure, and applied numerics. *Journal of Managerial Psychology, 3*(18), 421-439.

Clarke-Midura, J., & Dede, C. (2010). Assessment, technology, and change. *Journal of Research on Technology in Education, 42*(3), 309-328.

Dillon, S. C., & Brown, A. R. (2006). The art of ePortfolios: insights from the creative arts experience. In A. Jafari & C. Kaufman (Eds.), *Handbook of Research on ePortfolios* (pp. 420-433). Hershey PA: Idea Group Inc.

ePhoto Inc. (2012). ePhoto photography and portable film video lighting kit and support system. Retrieved 5th January 2013, from http://www.ephotoinc.com/photography-lights-studio-video-lighting-portable-video-lighting-kit.html

Heldsinger, S., & Humphry, S. (2010). Using the method of pairwise comparison to obtain reliable teacher assessments. *The Australian Educational Researcher, 37*(2), 1-20.

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.,

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.

Madeja, S. S. (2004). Alternative assessment strategies for schools. *Arts 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. (2011). Comparative pairs marking supports authentic assessment of practical

performance within constructivist learning environments. In R.F. Cavanagh & R. F. Waugh (Eds.), *Applications of Rasch Measurement in Learning Environments Research.* Rotterdam, The Netherlands: Sense Publishers.

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). *Let's stop marking exams.* Paper presented at the International Association for Educational Assessment Conference. Retrieved from http://www.cambridgeassessment.org.uk/ca/Our_Services/Research/Conference_Papers

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice, 19*(3), 281-300.

Ridgway, J., McCusker, S., & Pead, D. (2004). *Literature Review of E-assessment.* Bristol, UK: NESTA Futurelab.