

MULTIPLE MEASURES - BACKGROUND PAPER

USING MULTIPLE MEASURES FOR BENCHMARKING ASSESSMENT DESIGN

Dr Kate Tregloan, Faculty of Art Design + Architecture, Monash University
Prof Kit Wise, Tasmanian College of the Arts, University of Tasmania
Dr Wendy Fountain, Tasmanian College of the Arts, University of Tasmania

These papers offer background reading and information for users of the Multiple Measures online tool and website, www.multiplemeasures.org.au.

For further detail, or with any questions, please contact the authors, listed above.

Introduction

This paper is developed to support benchmarking decisions for users of the Multiple Measures tool.

The key prompt to benchmark a developing or delivered interdisciplinary (ID) assessment design has been developed by the project team as follows:

How well is interdisciplinary learning supported by the assessment design?

This prompt is expanded further by additional questions related directly to the filters that are applied for the sorting of the Multiple Measures library collection – further detail is provided in this paper.

The project team investigated a wide range of approaches to assessment for interdisciplinarity in Australian higher education, and concluded that there is not currently a single effective model for interdisciplinary teaching and learning activities. By contrast, the team uncovered a wide variety of approaches with subtle variation. The tool, and this paper, considers this range of approaches to assessment design for ID learning in terms of three key perspectives on interdisciplinary education that were identified through the project. Provided background papers expand on these perspectives and relate them directly to the questions. 'Sliders' within the tool allow users to sort the library according to these themes, to select exemplars most relevant to their own intentions, and inform the benchmarking process.

The identified perspectives focus on:

- student cohort
- learning outcomes
- pedagogic approach

Constructive alignment is a well-known approach to education design, aiming to ensure that these three sets of concerns are addressed coherently. The challenge identified by this project, is the constructive alignment of the broad range of learning outcomes and pedagogic approaches to be effective for diverse cohorts.

Benchmarking ID assessment design with Multiple Measures

Benchmarking is a formal process of establishing relative performance, through systematic comparison of some aspect of an example with that of other relevant 'partners'.

Benchmarking processes may provide richer forms of evidence for development and improvement than quality assurance, which emphasises threshold standards. Benchmarking is ideally dialogical, involving evaluative exchange with chosen partners, based on iterative reviews of data, and includes consideration of processes as well as outputs/outcomes. The utility of a benchmarking process may regularly depend on the choice of benchmarking partners; a degree of humility and honesty in the partner relationship; inclusion of process elements in the analysis; a sound understanding of the indicators for the function/s of interest; and commitment to strategic use of the evidence it provides (Epper, 1999; Henderson-Smart et al., 2006). Recent audits of tertiary institutions have proposed benchmarking as a suitable framework for quality assurance and quality improvement activities (Henderson-Smart et al, p 146).

In the development of the Multiple Measures project, the range of ID assessment design approaches identified has been broad and varied. Very few educators approach assessment of ID engagements in the same way or for the same reasons, and the examples identified are distributed across (and therefore impacted by) different institutions. The project aimed to offer educators interested in ID education a way to benchmark their own approaches or past deliveries, as well as developing ideas for new ID assessment design.

The project team has developed the Multiple Measures tool and website to inform self-benchmarking by users, on the basis of the themes or perspectives identified by the team through the project. These themes cross a number of approaches and intentions for ID education, as well as institutional boundaries. They are expressed in the tool as a set of six paired questions and are outlined above and within the tool itself. These perspectives on ID assessment design allow a user to sort the collected library of MM projects, and to find those most relevant to their own purposes and approaches. The establishment of an evaluative framework is therefore the first step of the benchmarking process and is made of the basis of these themes.

Once a user has filtered the library using the sliders or tickboxes relating to the six questions, a 'short-list' is presented. Project summaries provide a brief description, coloured lozenges indicate the coding of each MM unit/course/subject according to the six themes identified, and an adjacent diagram shows the timing, weighting and type of assessment tasks. Main assessment criteria are listed. Users can assess the appropriateness of each project on the shortlist as a 'benchmarking partner' via these details, and also refer to the project summaries provided. Users select **three** benchmarking 'partners' before proceeding to the benchmarking screen. Although a developing or delivered ID assessment is self-benchmarked, these selected partners offer a simplified context for that decision-making.

The benchmarking screen within the Multiple Measures tool brings together the selected MM projects with summary details and links as before. Reminders of a user's response to the sliders for sorting is included, and the tool also provides space for self-benchmarking notes and ideas. This is an opportunity to focus on the qualities of the ID engagements presented and considered in terms of the questions below, these benchmarking questions relate directly to the filters used to sort the library to find suitable comparators. The first page of the summary for each of the MM units/courses/subjects includes notes that respond to each of the questions to assist users of the tool to benchmark their own ID assessment approaches or ideas in these terms.

Benchmarking prompts developed through the Multiple Measures project follow the main question presented above. The main question is further explored through six identified questions as follows:

How well is interdisciplinary learning supported by the assessment design?

For the unit/subject/course to be benchmarked or developed ...

How well does the assessment design fit the ID cohort?

- Does it fit the level of student expertise?
- Does it respond to the range and style of cohort learning expectations ?

How well does the assessment design align to intended ID learning outcomes?

- Do the tasks and criteria sufficiently support development of students' disciplinary practices?
- Do the tasks and criteria sufficiently support development of students' interdisciplinary skills?
- Do the student / staff roles influencing the direction / aims of the assessment tasks support this?
- Do the student / staff roles influencing the process / development of the assessment tasks support this?

Users of the Multiple Measures tool can refer to the background papers for each of the six paired questions, this paper, and the details provided in the MM project summaries to frame their own perspective on the value and approaches of ID education from these perspectives. Educators should consider the strengths of their own assessment design for benchmarking, as well as opportunities for learning and development. At the conclusion of the benchmarking exercise, a pdf report will bind together selected MM summaries and any background materials, the user's own benchmarking and development notes. Background papers can be downloaded separately.

Constructive alignment + ID assessment

Constructive alignment suggests that tasks, criteria, evaluation tools and feedback methods are implicitly inter-connected. According to Biggs' widely adopted concept: "[A] design for teaching in which what it is intended students should learn and how they should express their learning is clearly stated before teaching takes place. Teaching is then designed to engage students in learning activities that optimise their chances of achieving those outcomes, and assessment tasks are designed to enable clear judgments as to how well those outcomes have been attained" (Biggs, 2014, pp. 5-6).

In this project, key design considerations apparent in interdisciplinary teaching have related the pedagogic approach and the desired learning outcomes: the effective alignment of these constitutes 'good' assessment design. The student cohort is often decided for the educator by outside factors, however this project suggests that assessment design decisions that aim for constructive alignment must be made relative to this cohort. This paper and the benchmarking approach of the tool will focus on ID assessment design as it aligns to intended learning outcomes and pedagogic approaches, and then as the whole relates to the particulars of interdisciplinary cohorts.

Most assessment design evident in MM projects values both process and outcomes, with varying emphases on each. These are typically delivered through cumulative, scaffolded assessment tasks, and usually (but not always) over the course of a semester. The rich suite of assessment tasks evidences different conceptions of interdisciplinarity and the relationship and value of processes and outcomes of student learning.

At the same time, while 'constructive alignment' (Biggs, 2003, 2014) is now well established in the creative arts, design and humanities disciplines, the learning outcome/assessment criteria alignment may be less clear in some ID assessment tasks. This conflict between process and outcome focus in ID learning may result, for example, if intended learning outcomes foreground process, but assessment criteria and evaluation tools favour outcome (or vice versa). Similarly, cross-disciplinary research and contextual study may be highlighted in task specifications and learning outcomes, but may not be explicitly apparent in assessment criteria. This challenge for both tutors and students may result from the translation of usual assessment practices from educators' usual disciplines to a new ID realm.

In the context of designing for interdisciplinary assessment Biggs' (2003, 2014) 'constructive alignment' becomes further nuanced and challenging in application. The combination of learning outcomes achieved by students, the processes they follow, and the project outcomes they achieve may be (and arguably should be) more difficult to define in advance due to the constructivist nature of interdisciplinary inquiry and its evaluation (Burgett et al, 2011; Klein, 2008). When interdisciplinary approaches are promoted with the suggestion that they will enable the development of new perspectives and innovative outcomes, the inherent conflict seems self-evident.

Sadler (2009, 2015) argues that in higher education only a subset of likely learning outcomes can be anticipated and prescribed via assessment criteria and rubrics. The MM examples suggest that educators may extend 'alignment' to the interlinking of learning outcomes, the aspects of 'becoming more interdisciplinary' or 'learning interdisciplinarity' they target (Burgett et al., 2011, p.467), and the evidence provided via the process, product/outcome, or the combination of these.

The key task for alignment of interdisciplinary assessment design is the selection of evaluation principles, such that the value identified in students' work is effectively linked with the learning developed and assessed through the interdisciplinary engagement. The role of both educator and student in the reflective evaluation of project development and success is necessarily linked in this undertaking, as is the potential for the qualitative descriptor as a broad and flexible definition of 'success' for either process or outcome. The treatment of assessment as either formative or summative depends in part on the definition of context – a development of broad creative skills for lifelong engagement, or the completion of a specified unit/course/subject within a qualification that may lead to the same end.

Assessment design considerations

ID assessment design + intended learning outcomes

Two broad aims for assessment design in terms of ID learning outcomes were identified through the project. These include the extension or development of students' individual disciplinary or personal practice/s through their engagement with colleagues from other disciplines, and the development of skills that will assist working with others in

interdisciplinary ways. These are not mutually exclusive, nor necessarily in competition. Multiple Measures background papers relating to questions 3 and 4 provide more detail.

Assessment design considerations relevant to learning outcomes (Q3 and Q4):

- Appropriately aligned assessment criteria can be conceived as an extension of the evaluation principles Klein (2008) identified in her analysis of interdisciplinary research projects. Examples include goal variability (scope, scale and level of integration in student work); variable criteria and indicators (experimentation, aesthetic quality, explanatory power, unforeseen consequences); and effectiveness and impact (sensitivity to goals and responsiveness to criteria and indicators, long-term impacts, value added) (Klein, 2008, p. 122).
- Where students' ability to work with others from different disciplines is strongly valued (eg Q4), relevant evaluation principles adapted from Klein's (2008) work on interdisciplinary research may provide a useful reference for benchmarking or development. These include design for 'leveraging integration', 'interacting social and cognitive factors', and 'leadership and coaching'. In practice, assessment criteria may seek to address collaborative skills via students' ability to:
 - Weave multiple perspectives into a new whole;
 - Articulate and clarify differences among stakeholders, and mediate their differing approaches to the brief/problem;
 - Build consensus and share decision making; and
 - Play a role in 'coaching the process' (Klein, 2008, p. 122).

ID assessment design + pedagogy

Shifting perspective from learning outcomes to pedagogy in ID assessment highlights the role or agency available to students and staff to influence the project direction / brief / aims (Q5) and the project process / development (Q6). In this conception, the assessment tasks and criteria become mechanisms through which students may be both empowered and evaluated. Addressing Q5 and Q6 in assessment design calls for consideration of the interplay of collaborative, disciplinary and meta-learning skills (Klein, 2008; Winters, 2011). Metacognitive engagement is considered central to ID learning by Burgett et al. (2011).

Assessment design considerations relevant to Pedagogic Approach (Q5 and Q6):

- Assessment tasks with explicit roles – educator, peers, students – can help build skills of evaluation appraisal that connect with skills of critique; “Developing evaluative expertise through guided practice would also equip learners to become self-critical and able to self-monitor their own work while it is in production, which ultimately is the very point at which it can make a difference to the work's quality” (Sadler, 2009, p. 177).
- Where student direction of the brief and/or process is strongly valued for allowing students to 'participate in practice', as Boud and Falchikov (2006) urge, there may be a case for re-orienting assessment toward in-process meta-learning, and re-balancing the importance commonly assigned to product/outcome (Burgett et al., 2011).

ID assessment design and the **student cohort**

Across the MM assessment tasks it is often implied that ideation and exploration will be enriched by the diversity of disciplines represented, and the expertise students bring with them. Q1 and Q2 reflect this concern for the range of disciplinary cultures and students' expertise levels, and how they might inform assessment design.

Assessment design considerations relevant to student cohort (Q1 and Q2):

- In order to be inclusive of a range of disciplinary cultures, assessment tasks may provide the option to work individually and/or collaboratively e.g. MM2, reflected in assessment task weightings and criteria.
- Providing project options can also cater to diverse disciplinary cohorts (e.g. MM1 group project (50%)) so that students can assimilate the routine ways of working in their 'home' discipline.
- Similarly, assessment tasks can invite multiple modes of representation to be inclusive of multiple disciplines, but also prompt technical and communication skill development (e.g. MM3, MM16).

Notes: Key terms and types of MM assessment tasks and criteria

In the MM project, and our discussion of assessment approaches in this paper, we have adopted the following definitions in relation to assessment:

- **Assessment tasks:** high impact learning activities that “not only enable judgements to be made about what has been learned ... [but also] require substantial involvement over time ... [ideally] in an interlinked, constructive, organised and coherent sequence” (Boud, 2010, p. 2). Tasks may serve one or more functions of assessment as distinguished by Hickman (2007) – formative, ‘ipsative’ (emphasising personal development), and diagnostic – as well as being summative for grading purposes.
- **Assessment criteria:** targeted features of the outcomes of assessment tasks against which student performance is judged. Each criterion is typically described (in text) in terms of the extent to which it is evident in the student work along an attainment scale (e.g. high distinction through to fail) (Sadler, 2009).
- **Evaluation tools:** mechanisms through which value is framed and communicated to students, and through which student performance is characterised (and often weighted) e.g. rubrics, self- and peer appraisal guidelines. The assessor (whether educator, peer or student) nominates the descriptor of ‘best fit’ for each criterion (Sadler, 2009).
- **Feedback:** communication of performance to students (by educators and peers) in relation to targeted learning, typically including identification of strengths and weaknesses and advice on how to improve subsequent performance. Feedback ideally functions as ‘feedforward’ – or applied formative feedback – that is used actively by students to improve their learning (Boud, 2010; Sadler, 2010).

The following list captures the most common assessment tasks in MM projects:

Ideation, exploration and development (process)

- Preliminary presentations in / to groups (visual, verbal, digital, material); formative feedback (tutor, peer, self)
- Journals / visual diaries
- Contextual study and research writing e.g. place-based observations, responses

Production, presentation and representation (outcome)

- 2D or 3D artwork or design
- Performance
- Exhibition
- Portfolio
- Digital documentation e.g. of time-based or place-based work, video presentations

Self-appraisal and critical reflection (often spanning process and outcome)

- Reflective writing
- Personal learning statements

Useful References

- Biggs, J. (2003). *Teaching for quality learning at university* (2nd ed.). Maidenhead, UK: Open University Press.
- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1, 5-22.
- Boud, D. (2010). Assessment 2020: Seven propositions for assessment reform in higher education. In Australian Learning and Teaching Council (Ed.). Sydney.
- Boud, D. & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment and Evaluation in Higher Education* 31(4), 399-413.
- Burgett, B., Hillyard, C., Krabill, R., Leadley, S. & Rosenberg, B. (2011). Teaching interdisciplinarity. *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture*, 11(3), 465-491.
- de la Harpe, B. & Peterson, F. (2008). *A model for holistic studio assessment in the creative disciplines*. Paper presented at the Proceedings of the 2009 ATN Assessment Conference, Adelaide, Australia.
- Epper, Rhonda Martin. (1999). Applying benchmarking to higher education: Some lessons from experience. *Change*, 31(6), 24-31. doi: 10.2307/40177858
- Henderson-Smart, Cheryl., Winning, Tracey., Gerzina, Tania., King, Shalinie & Hyde, Sarah. (2006). Benchmarking learning and teaching: developing a method. *Quality Assurance in Education*, 14(2), 143-155. doi: doi:10.1108/09684880610662024
- Hickman, R. (2007). (In defence of) whippet-fancying and other vices: Re-evaluating assessment in art and design. In T. Rayment (Ed.), *The problem of assessment in art and design* (pp. 77-87). Bristol: Intellect.
- Klein, J.T. (2008). Evaluation of interdisciplinary and transdisciplinary research: A literature review. *American Journal of Preventive Medicine* 35(2S), 116-123.
- Mansilla, V.B. (Jan/Feb 2005). Assessing student work at disciplinary crossroads. *Change*, 37, 14-21.
- Mansilla, V.B. & Duraising, E.D. (2007). Targeted assessment of students' interdisciplinary work: An empirically grounded framework proposed. *The Journal of Higher Education*, 78(2), 215-237. doi: 10.2307/4501203
- Ostwald, Michael J, & Askland, Hedda H. (2012a). Models and Matrices. In H. H. Askland, M. J. Ostwald & A. Williams (Eds.), *Assessing Creativity - supporting learning in architecture and design* (pp. 63-80). Sydney: Australian Government Office for Learning and Teaching.
- Ostwald, Michael J, & Askland, Hedda H. (2012b). Assessment regimes: patterns of creative evaluation in architecture and design. In H. H. Askland, M. J. Ostwald & A. Williams (Eds.), *Assessing Creativity - supporting learning in architecture and design* (pp. 81-100). Sydney: Australian Government Office for Learning and Teaching.
- Sadler, D.R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughlin (Ed.), *Assessment, learning and judgement in higher education* (pp. 1-19). Dordrecht: Springer.
- Sadler, D.R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment and Evaluation in Higher Education*, 35(5), 535-550. doi: 10.1080/02602930903541015
- Sadler, D.R. (2015). Backwards assessment explanations: Implications for teaching and assessment practice. In D. Lebler (Ed.), *Assessment in music education: From policy to practice* (Vol. 16) (pp. 9-19). Switzerland: Springer.
- Winters, T. (2011). Facilitating meta-learning in art and design education. *International Journal of Art and Design Education*, 30(1), 90-101. doi: 10.1111/j.1476-8070.2011.01685.x