

How do we know that we are assessing the “right things”?

J. R. Rudland, C. Lacey, K. Kenrick & M. Tweed

Abstract

Introduction: As assessment is perceived as a powerful tool for learning, it is important to ensure that assessments reflect the learning considered most significant. Assessment blueprinting offers the opportunity to ensure perceptions of what *should* be assessed align with what *is* assessed.

Methods: An expert panel was asked to determine the percentage of undergraduate assessment that should be devoted to broad domain areas previously agreed to define the outcomes of an undergraduate curriculum; this provided the blueprint. Staff who co-ordinated, implemented and assessed students on clinical runs indicated the percentage of their assessments allocated to the domain areas. Staff who constructed end-of-year summative assessments also analysed their assessments in terms of the domain areas. The “expert panel” blueprint values were then compared using Mann-Whitney U with the actual assessment conducted to determine variations between the ideal and actual assessment.

Results: What was considered important to assess closely aligned with what was assessed in most domain areas. The exceptions were the underassessment of two domains, Māori Health and Research and Information Literacy.

Conclusions: The chosen methodology identified areas that were under-represented in the actual student assessments. This prompted the school to consider whether this underrepresentation is problematic; if so, whether to redistribute or increase assessment, and whether the required increases should occur in-course or at the end of the year.

Keywords: assessment; medical education; blueprinting; subject representation.

University of Otago, Wellington

Correspondence

Joy R Rudland, Director
Education and Staff Support
University of Otago Medical School, Education Unit, University of Otago, Wellington
23A Mein Street
Newtown
Wellington 6242
New Zealand
Tel: +64 27 212 9305
Fax: +64 3 479 5459
Email: joy.rudland@otago.ac.nz

ASSESSMENT BLUEPRINTING

Introduction

The literature on assessment is well represented by studies exploring assessment tools and techniques, for example the validity, reliability, fairness and feasibility of specific tools. Quality assurance (QA) of items used in assessments has also been reported (Ware & Vik, 2009) as have course outcomes (Rexwinkel, Haenen, & Pilot, 2013). This research is to be applauded, but there is a danger of failing to ensure that what is assessed is what is deemed to be important. Assessment should follow what is expected to be learnt as opposed to driving learning.

What is learnt does not necessarily equate to what is formally taught. Many forms of evidence validate the interpretation of assessment results. One of these is content evidence based on an assessment blueprint (Downing, 2003). Assessment blueprinting is a process whereby the assessment can be conducted according to a reproducible plan (Hamby, 2006). It is unknown what percentage of medical schools currently adopt assessment blueprinting. In 2003, an audit found that only 15% of curriculum administrators at 144 United States and Canadian medical schools developed assessment blueprints (Bridge, Musial, Frank, Toe, & Sawilowsky, 2003).

A blueprint should be a necessary precursor to constructing assessment (Hamby, 2006). It may, but does not have to, align to the amount of teaching time in the curriculum; it should demonstrate the linkage between assessments and required learning outcomes; and it should ensure that sampling of the curriculum is even-handed. An assessment blueprint may be especially useful in an integrated curriculum (Hayes, 2008), where it can consider important generic aspects as opposed to specific specialty-derived learning.

Whilst the creation of a new medical school may afford the luxury of developing and working to a prescribed assessment blueprint, the authors argue that this may be difficult to achieve prospectively for existing schools. It is proposed that a more productive method determines what gap, if any, exists between an ideal assessment blueprint and assessment currently being conducted. From this more informed position, any requirements for change can be identified.

School managers should use the curriculum to determine how to blueprint. Assessment blueprints should be sufficiently detailed to describe subcategories and subclassifications (Downing, 2003), which may be by departments, discipline areas and/or key outcome attributes, etc. An independent expert panel can review the assessment content to ensure an appropriate balance across these categories and classifications.

Little has been published on how blueprints should be devised. This study describes one method—adopting a prospective vision of what a plan should look like and then comparing it with a retrospective analysis of what has actually occurred. It focuses on the process of blueprinting, not the selection of the parameters for blueprinting.

ASSESSMENT BLUEPRINTING

Method

Course background

Traditionally, for the University of Otago undergraduate curriculum, experts have developed assessment items (stations, questions, etc.), which are then reviewed by colleagues. The items are mapped to disciplines and tasks, and a committee is responsible for a final overview to look at the balance of the assessments and modify, where appropriate. However, the undergraduate curriculum has recently been defined by six broad programme outcome domains:

- Clinical (consultation) skills
- Medical sciences (such as anatomy, behavioural science, etc.)
- Ethics/professional development/legal
- Research and information literacy (relating to research skills)
- Population and community health
- Māori health (nation-specific indigenous health issues).

These domains are the key constructs to which the medical programme maps assessment and learning.

The school's medical curriculum is delivered in three discrete stages: early learning in medicine (Years 2 and 3), advanced learning in medicine (ALM) (Years 4 and 5) and the trainee intern (TI) (Year 6). In Years 4 and 5, relevant clinical departments each assess a series of modules/attachments to traditional specialties. There is also a Year 5, end-of-year (EOY) series of summative examinations, consisting of a short answer question exam, 200 extended matching items and an objective structured clinical exam (OSCE).

Proposed blueprint: Expert panel

Members of the Bachelor of Medicine and Bachelor of Surgery (MB ChB) Curriculum Committee, which oversees the content and delivery of the MB ChB degree course, were asked to use a paper-based survey to estimate the percentage of formal assessment conducted that should ideally be allocated to each domain area across the three stages of the course, including in-course and EOY assessment. This group, together with the deans of the component medical schools, comprised the "expert panel". The members of the expert panel each submitted their proposed value for each domain area, and these were then averaged.

This study uses only the Year 4 and 5 data. As this was a largely analytical exercise involving only academic staff with curriculum or assessment oversight responsibilities, ethics approval was not required.

ASSESSMENT BLUEPRINTING

Actual blueprint of Year 4/5***In-course assessment***

In mid-2010, documentation was developed and tested on a small group of module convenors representing the different locations and years of the course. This confirmed that the process was not too time consuming or counter-intuitive, so the exercise was rolled out to Year 4 convenors in 2010, Year 5 in 2011, and Year 6 (TI) in 2012.

Years 4 and 5 module convenors, responsible for delivering and assessing learning throughout the year, were asked to consider their current formal assessments and evaluate the percentage of each that pertained to each domain area. A written request, including a template listing their assessments, was sent to the convenors, and for each assessment tool, the convenors indicated what percentage of the assessment covered a specific domain area. A sample form is included in Appendix 1. For example, 90% of an OSCE may be allocated to the clinical skills domain.

End-of-year assessment

The convenors for the OSCE, short-answer paper and the paper for extended multiple-choice questions analysed the contents of all these examinations, assigning percentage values to the domain areas assessed in each.

The percentage value for actual in-course and EOY assessments were combined (assuming equivalence of value) and compared with the overall average values from the expert panel. Comparison was made using the Mann Whitney U test (using SPSS 19) to determine the difference between expected and actual. A non-parametric test was used, as a normal distribution could not be assumed.

Table 1
Proposed and Actual Blueprinting of Assessments for Year 4/5

Domain area	Proposed	Actual		
	Expert panel (n = 14) Mean % (Expert panel range)	Module assessments (n = 37) Mean %	EOY (2009–2011) Mean %	Module and EOY Combined Mean %
Clinical skills	45.71 (30–70)	46.85	43.97	46.39
Medical sciences	21.64 (5–40)	23.17	26.56	23.71
Ethics, professional development/legal	8.93 (5–15)	12.07	11.74	12.02
Research/information literacy	9.46 (5–20)	6.80	12.02	7.64*
Population health	8.54 (5–15)	9.70	3.71	8.75
Hauora Māori	5.71 (5–10)	2.67	1.06	2.51*

*Significant difference ($p < 0.05$) between the proposed percentage and actual total combined percentage.

ASSESSMENT BLUEPRINTING

Results

Fourteen of 20 (75%) members of the expert panel responded to the request for an “ideal” blueprint. The range of responses, as a percentage, is included in Table 1.

Convenors submitted assessment blueprints for 37 of 45 (82%) eligible modules. The convenors for the OSCE and short-answer paper analysed these examinations for the last 3 years (2009–2011) and the 2011 paper for extended multiple-choice questions, by domain areas. The results are presented in Table 1.

Significant difference existed only between the proposed percentage and actual total combined percentage for the Māori domain ($p < 0.001$) and research and information literacy ($p = 0.027$) areas. In both cases, the proposed percentages were higher than the actual percentage values for that stage of the course (Year 4/5).

Discussion

The close alignment of the values from the expert panel and actual assessment for most domains is reassuring. This indicates that most of the domain areas are being assessed to the level anticipated and considered appropriate.

However, the significant under-representation in assessment of two domain areas, Māori health and research and information literacy, raises questions for the medical school of whether more assessment is required in the under-represented areas. Repercussions from this decision will determine whether current assessments should be redistributed or additional assessments in the under-represented areas added. The quantity of assessment in a course is often a vexed issue. It goes beyond the scope of this study and attests often to the ethos of the medical school. The current data, however, supports a dialogue taking place regarding appropriate representation of specific domain areas, both learnt and assessed.

If an increase in assessments is required, it must be considered whether this is best achieved in EOY assessments or through in-course assessments throughout the year. Research and information literacy may be better assessed as a continuous in-course requirement, but any assessment not only has to fulfil certain assessment criteria (among them authenticity, reliability, fairness), but also be feasible. A complicating factor may be inadequate resources to ensure learning is achieved rather than inadequate assessment. This may result in a greater commitment to remedy underserved areas of learning and/or assessment or, conversely, an acknowledgment of the insolubility of the issue and an acceptance of the under-representation.

This study has started a process whereby staff involved in teaching and assessment can consider the balance of assessment in their courses. It allows the opportunity to debate the balance beyond what are often silos of practice. It may not be incumbent on all attachments to modify assessment practice but rather have an awareness of where some outcomes are assessed.

ASSESSMENT BLUEPRINTING

Finally, the study does not attend to the relationship between learning and assessment. It is not assumed that the focus on learning should be equally represented by the amount of assessment. However, it is acknowledged that assessment may drive some learning (McLachlan, 2006).

This study is limited as it relied on staff (module convenors and assessment convenors) appropriately allocating their assessments to outcome domain areas and appreciating how one assessment tool (in-course assessment) or item within an EOY assessment may cover a variety of domain areas. Although guidelines were given, some staff indicated difficulty in making these judgments. This attests to the artificial nature of taxonomy systems. The equivalence implied by combining the EOY and in-course assessment data is questionable.

An additional limitation is the retrospective design, which may contribute to the expert panel reporting an ideal distribution that simply reflects current practice rather than being a truly independent assessment of the ideal. Several members of the expert panel are members of the assessment committee that discussed the blueprinting project; two members are also module convenors. This slightly confounds their individual submissions of the overall ideal distribution, as they are inevitably somewhat informed by the experience of having reported on the actual module data. Finally, it is debatable whether the curriculum committee comprised an appropriate expert panel; it may have been advantageous to include students, community representatives, or other health professionals, such as nurses. While other professions had been consulted regarding the appropriateness of the outcomes, no attempt was made to determine the balance they felt was appropriate.

This study has demonstrated one method of blueprinting assessment that compared a judgement of actual practice with a proposed ideal and identified key areas of the curriculum that are currently under-represented in those assessments. This process has started a dialogue with staff regarding assessment and what is perceived to be valued. The process was relatively straightforward, and we recommend that all medical schools consider undertaking some form of blueprinting in order to appreciate assessment practice and to engender a useful dialogue regarding what is valued through assessment.

This retrospective quality assurance of what is assessed should occur routinely but not yearly. A 3 to 5-year cycle may be advantageous. Alternatively, a more responsive and proactive approach of prospective blueprinting could be adopted yearly. This could be occasionally compared to a “panel of experts” opinion on what should be learnt to reflect changes in societal needs. Ideally, community representation should be part of this panel.

Acknowledgements

Thanks to KW for co-ordinating and collating responses from module convenors, SJ for facilitating the completion of end-of-year assessment data and JB for proofreading the final draft.

ASSESSMENT BLUEPRINTING

References

- Bridge, P. D., Musial, J., Frank, R., Toe, T., & Sawilowsky, S. (2003). Measurement practices: Methods for developing content-valid student examinations. *Medical Teacher*, 25, 414–421.
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37, 830–837.
- Hamby, H. (2006). Blueprinting for the assessment of health care professionals. *The Clinical Teacher*, 3, 175–179.
- Hayes, R. (2008). Assessment in medical education: Roles for clinical teachers. *The Clinical Teacher*, 5, 23–27.
- McLachlan, J. C. (2006). The relationship between assessment and learning. *Medical Education*, 40, 716–777.
- Rexwinkel, T., Haenen, J., & Pilot, A. (2013). Quality assurance in higher education: Analysis of grades for reviewing course levels. *Quality and Quantity*, 47, 581–598.
- Ware, J., & Vik, T. (2009). Quality assurance of item writing: During the introduction of multiple choice questions in medicine for high stakes examinations. *Medical Teacher*, 31, 238–243.

ASSESSMENT BLUEPRINTING

Appendix**Assessment-Domain Mapping Template**

School
Year
Module name
Module convenor

Please note: we
don't expect that all
domains will apply
to all assessments.

What programme domains do the assessment tools assess?

Please complete the following grid. If the assessment tool no longer exists please delete (score through); if there are tools we have missed, please add to the chart using the blank rows. Please refer to page 6 of the accompanying booklet if you need further information.

Assessment tool	Programme Domains							Total
	Clinical Skills	Medical Science	Ethics/Prof Develop.	Population Health	Research & Information Literacy	Hauora Maori		
Eg - OSCE	60	15	15	10				100%
PASAF*								100%
Case write up								100%
Case presentation								100%
OSCE								100%
MCQ								100%
Logbook								100%

*PASAF: 'Collective opinion of professional attitudes': please describe how you obtain the information on which you base your PASAF assessment.

Who completes the assessment?

Please consider who does the assessing – just consider broad values.

Person/s	Percentage
Staff (academic/clinical)	
Self (the student)	
Student peers	
Patients	
Total	100%

If you have identified an assessor other than staff, please provide details.

Please return in the envelope provided. If you have any queries regarding completion of the form please contact your local MEA.