

The new curriculum framework and assessment practices: current challenges for postgraduate years 1 and 2

Brian C Jolly

The launch of a national curriculum framework for junior doctors is a significant and necessary stride in the development of medical education in Australia.¹ However, it is not sufficient to create a workable curriculum for junior doctors. The definition of core content alone will only go part of the way to establishing “the curriculum” that will drive learning for our junior doctors. Michael Eraut, Professor of Education at the University of Sussex in the United Kingdom, has identified any curriculum as requiring at least four interacting components, of which content is only one (Box).² Additionally, a consistent and coherent set of aims and objectives related to that content, a teaching and learning strategy, and an assessment program focused on monitoring the outcomes are required. For example, in a functional medical curriculum (as opposed to a syllabus), the teaching and learning activities support the learners’ transformation of content into useable clinical expertise. The assessment processes underpin and make this personal development accountable. Without all these dynamic attributes of the curriculum, trainees or students would have to invent their own learning activities to comply with the demands that the curriculum makes upon them — they would have to “fill in the gaps” in the curriculum structure. These gaps — the “hidden curriculum” — can have deleterious effects on trainees’ development, especially if their ability to provide the missing components is constrained by the context in which they are working.³ In this article, I discuss the need to underpin the curriculum framework with other educational strategies, mainly focusing on the requirement for trainee-friendly but reliable and valid assessment, and I suggest appropriate tools. I conclude with a brief discussion and challenge about how the assessment continuum should align with the continuum of training.

Context of the curriculum

Trends in medical education, including the new prevocational curriculum, usually arise from complex political, social, scientific and educational interactions. Such developments are rarely completely evidence-based.³⁻⁵ Changes to (or creation of) curricula need to fit the context in which they will be used. For a curriculum to be maximally effective, all its components need to be explicitly aligned in the intended direction. These components also need to be monitored to ensure they symbiotically enhance progress towards the program’s goals.

The new prevocational curriculum must operate in a social and clinical context in which relationships between trainees and supervisors can be challenging. For example, in the following interview segment from a study I conducted as part of a Master of Education thesis (Sussex University) in 1989, a new clinical student describes his impressions of the ward-based learning environment:

... the thing is to blunderbuss your way through... [it’s] the complete syndrome of medical people, you’re taught the art of doing that well. The idea is to teach you self-confidence, but what it does do is to teach [you] to say things in confidence

ABSTRACT

- The new curriculum framework for doctors in postgraduate years 1 and 2 is a step towards seamless medical education.
- The framework will need additional components to make “the curriculum” deliverable.
- Assessment is an essential element of most curricula, and assessment systems should be carefully planned.
- Diligent observation and rating in the workplace may provide a suitable approach.
- In the future, Australia must also thoroughly engage with the debate on continuing validation of competence.

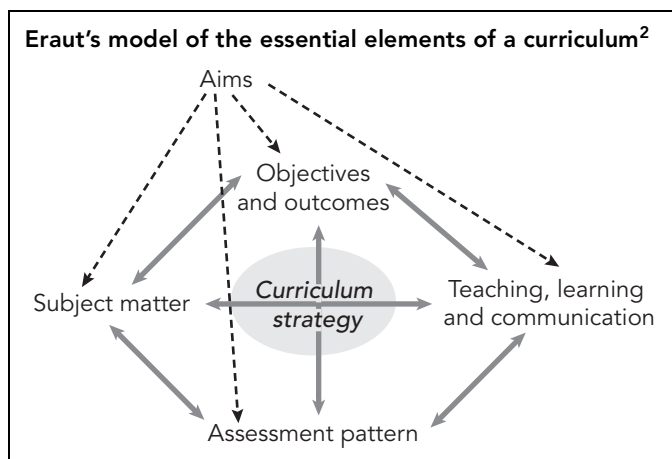
MJA 2007; 186: S33–S36

when [you] haven’t got it right. Everyone’s got to do this in front of patients.

There are numerous other similar examples in the literature on learning in clinical contexts over a 25-year period.⁶⁻⁸ Although this culture has changed, some values persist. This example may have arisen because of:

- the “person culture”⁹ of hospital environments allowing senior medical personnel considerable control;
- the widespread use of the Socratic method; and/or
- the perceived need to withhold some medical information from the patient, mainly by the use of technical jargon.

With increasing patient empowerment, this last cause has virtually disappeared, but the culture persists. Moreover, it is a collusive culture. Senior doctors and trainees will collaborate over bureaucratic workplace requirements to concentrate on the needs of the patient, and on their own. So a curriculum for the workplace needs to recognise this. All elements of the model in the Box will need careful deliberation, discussion, consultation and implemen-



tation to preserve the integrity of the curriculum without antagonising or being weakened by the clinical context.

For example, a salient phenomenon in the clinical learning environment when choosing educational and assessment strategies is the relatively poor relationship between volume of clinical experience and performance or competence.^{10,11} With more experienced trainees, there is a larger correlation between the volume of experience on procedures generally and their day-to-day competence with specific procedures.¹² Nevertheless, even for a group of surgical trainees, the best predictor of competence on a particular procedure is the volume of experience on that procedure, and this is significantly larger than the correlations between general surgical experience, time expired or other indices of experience, and competence.^{12,13}

This finding has several implications:

- hours of experience or numbers of clinical encounters are not the best measures for assessment;
- examination-style assessment may not reflect real differences between levels of clinical experience and competence obtained by trainees or students; and
- clinical skills are not necessarily generic — the best way to get good at a clinical skill or procedure is to practise that particular skill.

Trainees trying to respond to the curriculum framework will vary enormously in the amount of clinical experience they obtain. This is true across the undergraduate/postgraduate divide, and both within and between attachments and trainees.^{10,12} It also seems that clinical and educational supervisors consistently overestimate how much feedback they give,¹⁴⁻¹⁶ and cannot accurately predict how much experience, guidance and feedback trainees can obtain.

Because of these issues, curricula require infrastructure. However, previous analyses of the Australian vocational training context have suggested that important elements of infrastructure that need to be in place to deliver a vibrant curriculum may be missing. Essentially, there is:

- little cohesion across health and education sections, and national and state jurisdictions inhibit effective national policies in medical education and training;
- lack of coordination in Australia for medical education and training, and programs are under-resourced and under pressure;
- need for a coordinated national approach to assessment;
- need for improved links between prevocational and vocational training; and
- need for workforce and training agencies to work collaboratively.^{16,17}

Formal supervision arrangements, the training of supervisors, and the definition of and reward for their roles are all underdeveloped in Australia. Part of the solution lies in making the work of junior staff the focus of assessment, and integrating assessment methods into the training environment in a seamless fashion.

Assessment for the prevocational curriculum

There has been increasing interest in the development of work-based learning and assessment activities in the medical profession over the past two decades.^{18,19} Innovations have occurred across a wide front, from new methods such as the miniCEX (a clinical evaluation exercise)²⁰ to enhancements of older techniques such as “chart-stimulated recall”,²¹ used to investigate the practice of poorly performing doctors.²² In the UK, a comprehensive

approach to assessment of the new Foundation years has been attempted,²³ but such an approach in Australia will not be sustainable until both infrastructure is improved and there is general acceptance of and development around the notion of a training continuum from undergraduate to specialty levels.¹⁶ Any system will require extensive planning and discussion with stakeholders. Nevertheless, it is likely that effective assessment at prevocational level (either formative or summative) will include one or more of the validated and now increasingly accepted methods described below. The four methods share the following features:

- clinical work-based stimuli with which the trainee interacts — a patient, case records, or direct involvement in clinical work with the assessor;
- direct observation and rating of performance by the people most involved — peers, supervisors, or sometimes patients; and
- judgements captured on a rating form, across a number of different dimensions of the interaction — such as ethics, management, history, examination, and clinical decisions.

MiniCEX

The miniCEX was introduced by the American Board of Internal Medicine as a way of tapping into the behaviour formerly assessed by the long case; namely, history-taking and physical examination of the patient. It does this within a more focused and controlled framework, in a shorter time, and simultaneously achieves a larger sample of real work activities. It consists of a 15–20-minute sample of history or clinical examination activity, observed and rated unannounced by a senior staff member supervising a trainee. A summary of strengths and areas for improvement based on the case is signed off by both assessor and trainee. It is now in routine use in many training programs, and is being evaluated by specialties in Australia. A study involving 1228 encounters, 421 trainees, and 316 faculty members, in 21 training programs, has been reported.²⁴ A range of patient problems, settings, and types of visits were employed. Trainees’ ratings increased throughout the year, indicating evidence of construct validity, even though average assessor-rated patient complexity remained the same. Examiners were very satisfied with the format. Further evaluations of miniCEX-style assessments compared with specialty examinations indicate that the miniCEX performs well on content, construct and concurrent validity.²⁵

Peer or 360-degree assessment

A 360-degree assessment is a way of measuring and recording essential attributes of the professional clinician that may not be immediately observable in a one-off clinical encounter. Professionalism, patient management and self-management, teamwork skills and diligence are examples of such characteristics. A 360-degree assessment does this by collecting impressions about multiple attributes of trainees from a large group of peers or supervisors in an attempt to make these judgements more reliable.

Peer assessment was developed in the early 1990s by a group based in the United States.²⁶ In their study, a sample of peers rated the performance of a clinician on a series of nine-point scales. The data showed that about 11–13 peers could produce a reliable estimate of performance that also correlated well with other types of assessment in a credible pattern. The nature of the peers’ relationships with the rated participant did not make a difference

in the ratings, nor did the method of selection of peers. Peer assessment has developed into 360-degree or multisource feedback,²⁷ where other health professionals and patients²⁸ also contribute to the ratings.

Case-based discussions

Clinical decision making and safe patient management are vital elements of professional practice. Assessing these elements is difficult. Some researchers have developed techniques for this based on an approach, originally used as a research tool, called chart-stimulated recall.²¹ A patient record, or a videotape or audiotape of a consultation was used as a stimulus in a discussion between investigator and clinician about the diagnosis and management of the patient. Used as an assessment, the discussion usually focuses on choices that the clinician made (eg, selecting particular diagnoses or management decisions in preference to others) and the reasons for making them. This can then be compared with standard protocols, evidence or expert-based consensus to give a score. The technique has been shown to be especially useful in discussing clinical strategies with known or suspected poor performers.^{29,30} In a study of such doctors, several different assessment methods were applied to a group of volunteers and a group of doctors referred for assessment. The case-based discussion was highly correlated with a simulated patient-based examination and an oral examination, and was able to discriminate between the referred and volunteer groups. In the UK, assessors in the General Medical Council's performance procedures consistently described the case-based discussion as the most useful tool in the battery of tests used (Dame Lesley Southgate, Professor, St George's Medical School, University of London, London, UK, personal communication).

Directly observed procedures

Knowing that a doctor can do a particular task is usefully reassuring for a supervisor, and can allow them to devolve or delegate responsibilities. There has been recent interest in assessing trainees through directly observed procedures (DOPS). This work has burgeoned simultaneously in a number of countries.^{12,13,31,32} DOPS usually uses generic versions of rating scales, similar to objective structured clinical examination scales, applied to a real practical procedure in a work-based setting. In that sense, it is nothing new. Frequently, it is not convenient to have procedure-specific rating scales, although some researchers have worked with these.³³ In one of the early studies, researchers used a 120-item operation-specific checklist and a 10-item general global rating applied to a total of 41 theatre cases of three common operations: cholecystectomy (20 procedures), inguinal hernia (16) and bowel resection (5).³⁴ They found statistically significant differences based on year of training. Inter-rater reliability was good (0.78, 0.73). The Australian and New Zealand College of Anaesthetists is currently running some pilot projects on the use of DOPS and miniCEX.

Assessment and the continuum of training

The separate jurisdictions in Australia have contributed to the fragmentation of medical training and the compartmentalisation of its regulation. Recently, a university consortium made overtures, subsequently muted, about the consortium becoming engaged in

specialty training.³⁵ This was met with some concern by Colleges. For example, a report of the Royal Australasian College of Surgeons Council stated:

There was overt scepticism regarding the capacity of universities to deliver, particularly with the universal belief that most medical course curricula are inadequate in terms of surgical education.³⁶

However, there are good arguments for involving the nation's clinical academics in a phase of training that many see as the natural target of most of those who enter medical school. For example, it would help to make the continuum of training a reality. It also makes little sense, from a regulatory perspective, for the continuum of assessment to be apparently suspended during internship and again at specialty accreditation. This is particularly so when Australia has recently experienced challenging events, concerning individual and collective responsibility for poor clinical performance, without the concepts of revalidation or relicensing even being raised in the public and professional consciousness.³⁷ Work-based and 360-degree assessment might have avoided some of these problems. While the rest of the world grapples with these major issues,³⁸ Australia is largely silent on patient involvement in assessment of doctors, and on revalidation. To draw a thoroughly Australian metaphor, it is helpful to view such assessment as professional back-burning — a preventive measure that comes with initial discomfort, but might avert much bigger disasters. You can never tell when the next bushfire is going to threaten communities, but you can do everything possible to prevent them happening.

Competing interests

None identified.

Author details

Brian C Jolly, BSc, MA, PhD, Director
Centre for Medical and Health Science Education, Monash University,
Melbourne, VIC.
Correspondence: brian.jolly@med.monash.edu.au

References

- 1 Confederation of Postgraduate Medical Education Councils. Australian Curriculum Framework for Junior Doctors. November 2006. <http://www.cpmec.org.au/curriculum> (accessed Nov 2006).
- 2 Eraut M. Developing professional knowledge and competence. London: Falmer Press, 1994.
- 3 Bennett N, Lockyer J, Mann K, et al. Hidden curriculum in continuing medical education. *J Cont Educ Health Prof* 2004; 24: 145-152.
- 4 Norman G. Beyond PBL. *Adv Health Sci Educ Theory Pract* 2004; 9: 257-260.
- 5 Norman GR. Research in medical education: three decades of progress. *BMJ* 2002; 324: 1560-1562.
- 6 Arluke A. Roundsmanship: inherent control on a medical teaching ward. *Soc Sci Med [Med Psychol Med Sociol]* 1980; 14A: 297-302.
- 7 Seabrook MA. Clinical students' initial reports of the educational climate in a single medical school. *Med Educ* 2004; 38: 659-669.
- 8 Seabrook MA. Medical teachers' concerns about the clinical teaching context. *Med Educ* 2003; 37: 213-222.
- 9 Handy CB. Understanding organizations. 3rd ed. Harmondsworth: Penguin Books, 1985.
- 10 Jolly BC, Jones A, Dacre JE, et al. Relationship between students' clinical experiences in introductory clinical courses and their performance on an objective structured clinical examination (OSCE). *Acad Med* 1996; 71: 909-916.

SUPPLEMENT

- 11 Chatenay M, Maguire T, Skakun E, et al. Does volume of clinical experience affect performance of clinical clerks on surgery exit examinations? *Am J Surg* 1996; 172: 366-372.
- 12 Beard JD, Jolly BC, Newble DI, et al. Assessing the technical skills of surgical trainees. *Br J Surg* 2005; 92: 778-782.
- 13 Beard JD, Jolly BC, Southgate LJ, et al. Developing assessments of surgical skills for the GMC performance procedures. *Ann Roy Coll Surg Eng* 2005; 87; 242-247.
- 14 Grant J, Kilminster S, Jolly B, Cottrell D. Clinical supervision of SpRs: where does it happen, when does it happen and is it effective? *Med Educ* 2003; 37: 140-148.
- 15 Cottrell D, Kilminster S, Jolly B, Grant J. What is effective supervision and how does it happen? *Med Educ* 2002; 36: 1042-1049.
- 16 Paltridge D. Prevocational medical training in Australia: where does it need to go? *Med J Aust* 2006; 184: 349-352.
- 17 McGrath BP, Graham IS, Crotty BJ, Jolly BC. Lack of integration of medical education in Australia: the need for change. *Med J Aust* 2006; 184: 346-348.
- 18 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; 65 (9 Suppl): S63-S67.
- 19 Rethans JJ, Norcini JJ, Baron-Maldonado M, et al. The relationship between competence and performance: implications for assessing practice performance. *Med Educ* 2002; 36: 901-909.
- 20 Holmboe ES, Huot S, Chung J, et al. Construct validity of the mini clinical evaluation exercise (miniCEX). *Acad Med* 2003; 78: 826-830.
- 21 Tugwell P, Dok C. Medical record review. In: Neufeld VR, Norman GR, editors. *Assessing clinical competence*. New York: Springer, 1985: 142-182.
- 22 Norman GR, David DA, Painvin A, et al. Comprehensive assessment of clinical competence of family/general physicians using multiple measures. In: *Proceedings of the Association of American Medical Colleges' Research in Medical Education (RIME) Conference*. Washington, DC: AAMC, 1989: 75-80.
- 23 Beard J, Strachan A, Davies H, et al. Developing an education and assessment framework for the Foundation Programme. *Med Educ* 2005; 39: 841-851.
- 24 Norcini JJ, Blank LL, Duffy D, et al. The mini-CEX: method for assessing clinical skills. *Ann Intern Med* 2003; 138: 476-481.
- 25 Hatala R, Ainslie MO, Kassen B, et al. Assessing the mini-Clinical Evaluation Exercise in comparison to a national specialty examination. *Med Educ* 2006; 40: 950-956.
- 26 Ramsey PG, Wenrich MD, Carline JD, et al. Use of peer ratings to evaluate physician performance. *JAMA* 1993; 269: 1655-1660.
- 27 Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. *BMJ* 2003; 326: 546-548.
- 28 Greco M. Raising the bar on consumer feedback — improving health services. *Aust Health Consumer* 2005-06; (3): 11-12.
- 29 Cunnington JPW, Hanna E, Turnbull J, et al. Defensible assessment of the competency of the practicing physician. *Acad Med* 1997; 72: 9-12.
- 30 Southgate L, Cox J, David T, et al. The assessment of poorly performing doctors: the development of the assessment programmes for the General Medical Council's performance procedures. *Med Educ* 2001; 35 Suppl 1: 2-8.
- 31 Griffiths CEM. Competency assessment of dermatology trainees in the UK. *Clin Exp Derm* 2004; 29: 571-575.
- 32 Morris A, Hewitt J, Roberts CM. Practical experience of using directly observed procedures, mini clinical evaluation examinations, and peer observation in pre-registration house officer (FY1) trainees. *Postgrad Med J* 2006; 82: 285-288.
- 33 Darzi A, Mackay S. Assessment of surgical competence. *Qual Safety Health Care* 2001; 10 Suppl 2: ii64-ii69.
- 34 Winckel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. *Am J Surg* 1994; 167: 423-427.
- 35 Brooks P. Submission to the Productivity Commission on Australia's health workforce. 2006. <http://www.pc.gov.au/study/healthworkforce/subs/sub051.pdf> (accessed Dec 2006).
- 36 Stitz R. President, Royal Australasian College of Surgeons. Council highlights July 2006. http://www.surgeons.org/Content/Navigation-Menu/WhoWeAre/Council/CouncilHighlights/Council_Highlights_July_2006.pdf (accessed Mar 2007).
- 37 Dunbar JA, Reddy P, Beresford B, et al. In the wake of hospital inquiries: impact on staff and safety. *Med J Aust* 2007; 186: 80-83.
- 38 Dauphinee D. Self-regulation must be made to work. *BMJ* 2005; 330: 1385-1387.

(Received 20 Dec 2006, accepted 14 Feb 2007)

□