Workplace Based Assessments and Entrustable Professional Activities: Moving away from classroom assessments

Dr Fun-Gee Chen
Director Division of Graduate Medical Studies
Yong Loo Lin School of Medicine
National University of Singapore

Director of Critical Care
National University Health System
Population: 5 460 302 (July 2913)
GDP per capita: $61400
Infant Mortality Rate: 2.59 / 1000 life births
Maternal Mortality Rate: 3 / 100000 live births
Life expectancy: 84.04 years
Health expenditure: 4% of GDP
Doctor/Pop: 1.83 doctors to 1000 population
ACGME (I) Residency

• Duration of 5 to 6 years in Singapore

• Residents progress based on assessment of competencies by faculty, clinical competency committee

• On completion of training, resident is signed off by program director

• Specialty certification examinations conducted by DGMS and JCST

• Resident qualifies as specialist on completion of BOTH training and examination
Core Competencies
American College of Graduate Medical Education

- Patient Care
- Medical Knowledge
- Practice Based Learning and Improvement
- System Based Practice
- Professionalism
- Interpersonal Skills and Communication
Introduction by Program Director

The NUHS Anaesthesiology Residency Program strives for excellence in all aspects of residency training. The cornerstone of our Academic Mission is the training of high-quality, well-rounded specialists in Anaesthesia, Critical Care and Pain Medicine.

The mission of our Program is the dedicated education of our residents in all areas of knowledge, clinical & technical skills, and professional attitudes related to Anaesthesia. Graduate anaesthetists meet the requirements to become certified by the Master of Medicine (MMed), Anaesthesia Singapore and/or Fellowship of the Australian and New Zealand College of Anaesthetists (FANZCA) as specialist anaesthetists, who are proficient to practice anaesthesiology independently.

Our PD Raymond Goy (far left) with CA-1 Residents: Timothy Lam, Aileen Tan, Cheng Yi-Ling, Selene Tan, Teo Wei Wei, Vinod Chaubey and Eric Lee at the June 2011 OBS

A/Prof Raymond Goy
OUTLINE OF PRESENTATION

• The evolution of assessment in education
• Competency based medical education and assessments
• Assessment tools: is there a “best” tool?
• Programmatic assessment: what it means?
• Challenges in implementation
• Dundee University Experience / NUS pharmacy Experience
The Imperial Examination: The heart of the exam was a regurgitation of the Four Great Books and Five Classics, including Confucius's Analects, the Book of Mencius, the Great Learning, and the Doctrine of the Mean. This exam will determine if an applicant can enter the Imperial Civil Service.
IMPERIAL EXAMINATION: HOW TO CHEAT!
The Apprenticeship Model of Training

Professional Education where the master observes, guides, and conducts formative assessments of the student in various tasks to achieve MASTERY.
Medical Training: Classroom to Bedside
Competency based Education

Six critical components

1. Explicit learning outcomes: Required skills and concomitant proficiency
2. Flexible time frame to acquire and master these skills
3. Varied instructional activities to facilitate learning
4. Criterion referenced testing of the required outcomes
5. Certification based on demonstrated learning outcomes
6. Adaptable modules for optimal learning

Van der Horst & McDonald 1997
Competency based Education

<table>
<thead>
<tr>
<th></th>
<th>Traditional Instruction</th>
<th>Competency-based Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure</strong></td>
<td>Time-based</td>
<td>Learner-centered</td>
</tr>
<tr>
<td><strong>Teaching mode</strong></td>
<td>Group learning</td>
<td>Individualized</td>
</tr>
<tr>
<td><strong>Assessment Method</strong></td>
<td>Summative, high stakes</td>
<td>Mastery-learning, performance-based</td>
</tr>
<tr>
<td><strong>Pace</strong></td>
<td>Faculty-paced</td>
<td>Self-paced</td>
</tr>
<tr>
<td><strong>Program completion</strong></td>
<td>Finish when required courses are passed</td>
<td>Finish when mastery of courses is demonstrated</td>
</tr>
</tbody>
</table>
# The Two-Dimensional Matrix Relationship Between Entrustable Professional Activities (EPAs) and General Competencies*

<table>
<thead>
<tr>
<th>EPAs</th>
<th>Care of uncomplicated pregnancies</th>
<th>Normal delivery</th>
<th>Uncomplicated puerperium and neonate</th>
<th>The high risk complicated delivery</th>
<th>Perioperative care</th>
<th>Surgery estimated as low risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACGME competencies†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The ability to provide adequate patient care</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>The possession and ability to apply medical knowledge</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>The ability to learn from clinical practice and to improve it</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>The possession and ability to apply interpersonal and communication skills</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>The ability and commitment to carry out professional responsibilities</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>The awareness of and ability to operate optimally within the context, system, and resources of health care</td>
<td>•</td>
<td>•</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

EPAs are the focus of assessment, by observation, ratings or otherwise.

The overall assessment of competencies is not actually done. In stead, their presence is inferred from the assessment of sufficient EPAs.
The Miller’s Triangle

Miller’s Pyramid of Competency evaluation through Performance

- **Performance integrated into practice**: eg through direct observation, workplace based assessment
- **Demonstration of learning**: eg via simulations, OSCEs
- **Interpretation/Application**: eg through case presentations, essays, extended matching type MCQs
- **Fact gathering**: eg traditional true/false MCQs

Does
Shows
Knows how
Knows

Adapted from Burns and Mehay (2009) Miller’ Prism of clinical competency
* Multiple choice questions (MCQ)
Assessing professional competence: from methods to programmes

Cees P M van der Vleuten & Lambert W T Schuwirth

Table 1  Reliability estimates of different assessment instruments as a function of testing time

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Description</th>
<th>Reliability for different testing times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple choice*42</td>
<td>Short stem and short menu of options</td>
<td>0.62 0.76 0.93 0.93</td>
</tr>
<tr>
<td>Patient management problem*42</td>
<td>Simulation of patient, full scenarios</td>
<td>0.36 0.53 0.69 0.82</td>
</tr>
<tr>
<td>Key feature case (write-in)*43</td>
<td>Short patient case vignette followed by write-in answer</td>
<td>0.32 0.49 0.66 0.79</td>
</tr>
<tr>
<td>Oral examination†44</td>
<td>Oral examination based on patient cases</td>
<td>0.50 0.69 0.82 0.90</td>
</tr>
<tr>
<td>Long case examination†45</td>
<td>Oral examination based on previously unobserved real patient</td>
<td>0.60 0.75 0.86 0.90</td>
</tr>
<tr>
<td>OSCE*46</td>
<td>Simulated realistic encounters in round robin format</td>
<td>0.54 0.69 0.82 0.90</td>
</tr>
<tr>
<td>Mini-clinical exercise (mini-CEX)†47</td>
<td>Short follow-up oral examination based on previously observed real patient</td>
<td>0.73 0.84 0.92 0.96</td>
</tr>
<tr>
<td>Practice video assessment†16</td>
<td>Selected patient–doctor encounters from video recordings in actual practice</td>
<td>0.62 0.76 0.93 0.93</td>
</tr>
<tr>
<td>Incognito standardised patients†48</td>
<td>Real consultations scored by undetected simulated patients</td>
<td>0.61 0.76 0.82 0.86</td>
</tr>
</tbody>
</table>

* One-facet all random design with items crossed with persons (pxi).
† Two-facet all random design with judges (examiners) nested within items within persons (j:ip).
‡ One-facet all random design with items nested within persons (ip).
Utility Formulae for Assessment


Utility = R x V x EI x P x A x CE

(R = Reliability; V = Validity; EI = Educational impact; P = Practicability; A = Acceptability; CE = Cost-effectiveness)
Traditional Approach to Examinations

• End of course comprehensive ("finals") assessment: if the resident passes he is granted a license to practice as a specialist; if he fails he resits until he is successful within a agreed time frame (minimal feedback on why he failed)

• Assumption: the successful candidate has “mastered” the expected domains for the rest of his life; able to apply this knowledge to “real life”.

• Learning behavior: Passing the examination is of ultimate importance, rather than learning to be a competent specialist (focus on exam technique rather than competency)
Programmatic assessment: From assessment of learning to assessment for learning

LAMBERT W. T. SCHUWIRTH & CEES P. M. VAN DER VLEUTEN
Maastricht University, The Netherlands

Abstract

In assessment a considerable shift in thinking has occurred from assessment of learning to assessment for learning. This has important implications for the conceptual framework from which to approach the issue of assessment, but also with respect to the research agenda. The main conceptual changes pertain to programmes of assessment. This has led to a broadened perspective on the types of construct assessment tries to capture, the way information from various sources is collected and collated, the role of human judgement and the variety of psychometric methods to determine the quality of the assessment. Research into the quality of assessment programmes, how assessment influences learning and teaching, new psychometric models and the role of human judgement is much needed.
Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious effects on the curriculum and student learning.
Programmatic Assessment: What and Why?

Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious effects on the curriculum and student learning.

• It represents an attempt to consider assessment as constituting a programme as important as the curriculum itself and thus meriting being planned and reviewed in a similar way.
Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious effects on the curriculum and student learning.

- It represents an attempt to consider assessment as constituting a programme as important as the curriculum itself and thus meriting being planned and reviewed in a similar way.

- It underlines that while assessment is necessary as a basis for progress and award decisions, it may also be considered as a learning programme for students in itself.
Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious effects on the curriculum and student learning.

- It represents an attempt to consider assessment as constituting a programme as important as the curriculum itself and thus meriting being planned and reviewed in a similar way.
- It underlines that while assessment is necessary as a basis for progress and award decisions, it may also be considered as a learning programme for students in itself.
- It stresses that the data derived from assessment charts the progress and development of students throughout their programme.
Programmatic Assessment is a radical approach to assessment throughout the medical education programme devised to address endemic problems in assessment and its deleterious effects on the curriculum and student learning.

- It represents an attempt to consider assessment as constituting a programme as important as the curriculum itself and thus meriting being planned and reviewed in a similar way.

- It underlines that while assessment is necessary as a basis for progress and award decisions, it may also be considered as a learning programme for students in itself.

- It stresses that the data derived from assessment charts the progress and development of students throughout their programme.

- It reflects the belief that no one assessment point should determine progress or award but that such decisions should be based on an aggregation of points through the programme.
The Assessment Blue Print
Table 3. Suggestions for standardization of portfolio assessment.

1. Same portfolio’s units of evidence are assigned to all students
2. Tasks and criteria for assessment are defined and made clear
3. Instructions to students provide clear guidelines
4. The portfolio reading process and rating of material follow standardized guidelines—mainly through written instructions and training workshops for examiners
5. The probing in an oral review of the portfolio with the student follows standardized guidelines
A model for programmatic assessment fit for purpose

C. P. M. VAN DER VLEUTEN¹, L. W. T. SCHUWIRTH², E. W. DRIESSEN¹, J. DIJKSTRA¹, D. TIGELAAR³, L. K. J. BAARTMAN⁴ & J. VAN TARTWIJK⁵

¹Maastricht University, The Netherlands, ²Flinders Medical School, Australia, ³Leiden University Graduate School of Teaching, The Netherlands, ⁴Utrecht University of Applied Sciences, The Netherlands, ⁵Utrecht University, The Netherlands

---

**Diagram:**

- **Training Activities**
- **Assessment Activities**
- **Supporting Activities**

**Legend:**

- ○ = learning task
- ○ = learning artifact
- △ = single assessment data-point
- ▲ = single certification data-point for mastery-tasks
- — = learner reflection and planning
- — = social interaction around reflection (supervision, intervision)
- — = learning task being an assessment task also

**Periods:**

- **Period 1**
- **Period 2**
- **Period n**

**Time Arrow:**

---
# The Dundee University Portfolio

<table>
<thead>
<tr>
<th>Student work</th>
<th>Year</th>
<th>Number</th>
<th>Pre-marked*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student’s personal summary of progress towards each outcome</td>
<td>5</td>
<td>12</td>
<td>No</td>
</tr>
<tr>
<td>Patient presentations: short summaries of patients seen by the student,</td>
<td>4</td>
<td>10</td>
<td>Sometimes marked</td>
</tr>
<tr>
<td>selected from the 100 core clinical problems on which teaching and learning is based in this phase of the curriculum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case discussions: reports of approximately 1500 words, each analysing a patient’s history and findings in terms of one of the curriculum themes</td>
<td>5</td>
<td>7</td>
<td>Yes</td>
</tr>
<tr>
<td>Year 4 assignment: a project report with a grade awarded by the project supervisor together with feedback for the students</td>
<td>4</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Record of achievement: a record of procedures that the student was expected to have completed or observed during the phase</td>
<td>4 &amp; 5</td>
<td>1</td>
<td>Signed by faculty</td>
</tr>
<tr>
<td>GP special study module assessment form: a report on the student by his/her general practice supervisor with a grade awarded</td>
<td>5</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Clinical special study module assessment form</td>
<td>5</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>Theme special study module assessment form: a report on student performance during the module with a grade awarded for relevant outcomes</td>
<td>5</td>
<td>2</td>
<td>Yes</td>
</tr>
<tr>
<td>PRHO apprenticeship in medicine assessment form: a learning contract between the student and his/her educational supervisor with grades awarded for each learning outcome</td>
<td>5</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>PRHO apprenticeship in surgery assessment form: a learning contract between the student and his/her educational supervisor with grades awarded for each learning outcome</td>
<td>5</td>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Elective report: a report completed by the student after the elective period. This was read by one of the two members of staff responsible for elective studies and written feedback provided</td>
<td>5</td>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*The pre-marked student work is scored using the university grading system (A–G) and contains comments provided by faculty.*
Figure 1: The portfolio examination process: the examiners’ perspective
Figure 3: Flow diagram for the decision-making process
WEB PAPER

DOPS assessment: A study to evaluate the experience and opinions of trainees and assessors†

NATISH BINDAL1, HELEN GOODYEAR2, TARUNA BINDAL3 & DAVID WALL2
1Department of Anaesthesia, Queen Elizabeth Hospital Birmingham, UK, 2West Midlands Workforce Deanery, UK, 3West Midlands Workforce Deanery, Alexandra Hospital Redditch, UK

Abstract

Background: Workplace based assessments (WBAs) have been part of UK training for the last 3 years. Carrying out procedures efficiently and safely is of paramount importance in anaesthesia.

Aims: To explore opinions and experiences of Direct Observation of Procedural Skills (DOPS) assessments in a regional anaesthetic training programme.

Methods: 19 and 20-item questionnaires were distributed to trainees and consultants respectively.

Results: Questionnaire response rate was 76% (90/119) for trainees and 65% (129/199) for consultants. 43% of consultants and 33% of trainees were not trained in DOPS use. Assessments were usually not planned. 50% were ad hoc and the remainder mainly retrospective. Time spent on assessment was short with DOPS and feedback achieved in ≤15 minutes in the majority of cases with lack of suggestions for further improvement. Both trainees and consultants felt that DOPS was not a helpful learning tool (p = 0.001) or a reflection of trainee competency.

Conclusions: DOPS assessments are currently not valued as an educational tool. Training is essential in use of this WBA tool which needs to be planned and sufficient time allocated so as to address current negative attitudes.
Both trainees and consultants felt that DOPS was not a helpful learning tool \((p=0.001)\) or a reflection of trainee competency.
Workplace assessments in the operating room in anaesthesia residency:

A qualitative inquiry of its learning value at National University Health System

Fun-Gee Chen: Masters in Health Professions Education Thesis
Maastricht University
Experience with miniCEX and DOPS

“Most WBAs are opportunistic. It is not a planned thing like oh, I can do this procedure tomorrow. The supervisor does not know what miniCEX or DOPs I have done. Some supervisors have vague ideas (about WBAs) because they may have done it before. Some have completely no idea and we have to explain what is required”

“The miniCEX done here (NUHS) was really just to fill up the form, and the seniors approach to miniCEX was: give me the form and I will then sign……OK bye. I think in terms of learning, very…. to be honest, very minimal.”

Junior Resident Focus group Inquiry
Pros

• Neutralizes the limitations of traditional assessment
• High-stakes decision is not based on the outcome of a single assessment
• Informal assessments also find a place in the final decision
• Feedback is the backbone of the entire process
• Both quantitative and qualitative feedback are given equal weight-age
• Mentor-Mentee system plays an important role in improving the student performance
• Helps Assessors to take an evidence based high-stakes decision
Cons

- It requires extensive microplanning for success of this form of assessment
- The performance in each of the assessments has to be compiled for each student, which is a tedious task
- Difficult to make pass fail decision for inexperienced faculty members
- Feedback can be disheartening, if not delivered constructively
- Compilation remains a difficult task
- Often it takes a back-seat and faculty members do not find time for the tasks
- Requires a plan for the entire academic year that is well designed right at the start of the year
Summary of take home lessons

• Requirements of good summative assessments

• In order to make a good judgement regarding a resident’s competence you require
  - Understanding of the blueprint
  - Understanding of how to conduct the assessment
  - Understanding of the standard of performance required at the particular stage of training

• There no perfect assessments (portfolio examinations enable multiple sampling of the resident’s abilities, BUT if the assessments are not done appropriately it is difficult then to determine if the resident is sufficiently competence for practice
Can I leave the theatre? A key to more reliable workplace-based assessment

<table>
<thead>
<tr>
<th>Progression to Autonomy</th>
<th>Required Supervisor input for safe practice</th>
<th>Generally autonomous, some guidance required</th>
<th>Autonomous practice</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domains</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Patient assessment/investigation</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Preparation for anaesthesia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problem solving/decision making</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vigilance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organisation/efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professionalism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Can I leave the theatre? A key to more reliable workplace-based assessment

J. M. Weller¹,²*, M. Misur², S. Nicolson², J. Morris³, S. Ure⁴, J. Crossley⁵ and B. Jolly⁶

<table>
<thead>
<tr>
<th>Level of Independence</th>
<th>Supervisor required in the theatre suite</th>
<th>Supervisor required in hospital</th>
<th>Supervisor not required</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

What level of supervision did the trainee require for this case?
Entrustable Professional Activity (EPA)

“A critical part of professional work that can be identified as a unit to be entrusted to a trainee once sufficient competence has been reached.” Ten Cate, 2005

• The supervisor confirms when a trainee has reached the level where they can be trusted to perform each activity without direct supervision.

• A trainee should gain entrustment in performing these activities without direct supervision before progression to subsequent stages of training.
### Table 1  Promises and advantages of a programmatic approach to assessment. Adapted from: Dijkstra et al. 2010; Van der Vleuten et al. 2012

<table>
<thead>
<tr>
<th>Promises/purposes</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of what is and what is not being measured</td>
<td>Promote the validity of content and prevent emphasis on easy-to-measure elements (over- and underrepresentation)</td>
</tr>
<tr>
<td>Compensation for deficiencies of instruments by strengths of other instruments</td>
<td>Diverse spectrum of complementary measurement instruments capturing competence as a whole</td>
</tr>
<tr>
<td>Matching instruments to free space and time for the assessment of other subjects</td>
<td>Increase efficiency by reducing redundancy in information gathering</td>
</tr>
<tr>
<td>Combine information from different sources (tests/instruments) in high-stakes assessment</td>
<td>Reach better-informed and highly defensible high-stakes decisions</td>
</tr>
<tr>
<td>Multiple individual assessment points that are maximally informative to the learning process</td>
<td>Optimise the learning function of assessment (assessment for learning)</td>
</tr>
<tr>
<td>Aggregated data used for high-stakes pass/fail and remediation decisions</td>
<td>Optimise the certification function (assessment of learning)</td>
</tr>
<tr>
<td>Reducing bias in assessment of complex tasks through smart sampling strategies and procedural measures</td>
<td>Expert judgment of competence in performing daily tasks becomes valid and reliable</td>
</tr>
</tbody>
</table>
"When the cook tastes the soup, that's formative: When the guests taste the soup, that's summative."

"As the cook, or teacher, we need to stop and taste the soup before we move forward with instruction. We need to design instruction so students can press the reset button and go back to learn what they missed the first time. We can use many techniques to assess student achievement and understanding."

Debra Dicksen